

ND = 12982
KOE: 12852

2001

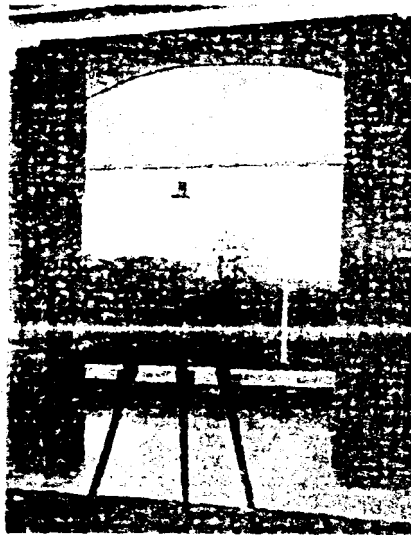
ΔΙΑΔΙΑΝΕΠΙΣΤΗΜΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
"ΙΣΤΟΡΙΑ ΚΑΙ ΦΙΛΟΣΟΦΙΑ ΤΩΝ ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΤΗΣ ΤΕΧΝΟΛΟΓΙΑΣ"

ΤΟΜΕΑΣ Α.Κ.Ε.Α.
ΤΜΗΜΑ Ε.Μ.Φ.Ε.
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΤΟ ΠΡΟΒΛΗΜΑ ΤΩΝ ΝΟΗΤΙΚΩΝ ΑΝΑΠΑΡΑΣΤΑΣΕΩΝ
ΣΤΗ ΓΝΩΣΙΑΚΗ ΕΠΙΣΤΗΜΗ:
ΠΡΟΣ ΜΙΑ ΜΗ-ΑΝΑΠΑΡΑΣΤΑΣΙΑΚΗ ΠΕΡΙΓΡΑΦΗ
ΤΩΝ ΝΟΗΤΙΚΩΝ ΦΑΙΝΟΜΕΝΩΝ

ΚΩΣΤΑΣ ΠΑΓΩΝΑΙΩΤΗΣ



ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ
ΔΡΙΣΤΕΙΑΣ ΜΗΛΑΤΑΣ
ΑΝΑΠΑΗΡΩΤΗΣ ΚΑΘΗΓΗΤΗΣ Ε.Μ.Π.

ΑΘΗΝΑ 2001

Αφιερώνεται στους γονείς μου
Κυριάκο και Μαργαρίτα Παγωνδιώτου
και στην αδελφή μου
Ευαγγελία Παγωνδιώτου



ΠΕΡΙΕΧΟΜΕΝΑ

ΕΙΣΑΓΩΓΗ	1
-----------------------	----------

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ: Ιστορικές καταβολές της γνωσιακής επιστήμης	17
1.1 Εισαγωγή.....	18
1.2 Εξελίξεις στην ψυχολογία που οδήγησαν στην ανάπτυξη της γνωσιακής ψυχολογίας.....	18
1.2.1 Η κριτική της ενδοσκοπικής ψυχολογίας και η ανάπτυξη του μεθοδολογικού συμπεριφορισμού.....	18
1.2.2 Η κριτική του συμπεριφορισμού και η εμφάνιση της γνωσιακής ψυχολογίας.....	20
1.3 Εξελίξεις στη μαθηματική λογική που οδήγησαν στην ανάπτυξη της τεχνητής νοημοσύνης.....	23
1.3.1 Η συμβολή του Frege.....	23
1.3.2 Η Θεωρία Αποδείξεων.....	26
1.3.3 Η Τυπική Σημασιολογία.....	27
1.3.4 Η πληρότητα του κατηγορηματικού λογισμού.....	28
1.3.5 Η ανάπτυξη της θεωρίας του υπολογισμού.....	29
1.3.6 Πώς λειτουργεί μια μηχανή Turing;.....	31
1.3.7 Η μηχανή von Neumann.....	35
1.3.8 Η ανάδυση της Τεχνητής Νοημοσύνης.....	36
1.4 Εξελίξεις στην φιλοσοφία που συνέβαλαν στην ανάπτυξη της γνωσιακής επιστήμης.....	37
1.4.1 Ο Λογικός Συμπεριφορισμός.....	38
1.4.2 Η Θεωρία Ταυτότητας.....	40
1.4.3 Ο Λειτουργισμός.....	42
1.5 Επίλογος.....	43

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ: Η Αναπαραστασιακή-Υπολογιστική Θεωρία του Νου (ΑΥΘΝ)	45
2.1 Εισαγωγή.....	46
2.2 Αφετηριακές θέσεις της Γνωσιακής Επιστήμης.....	46
2.3 Ιστορική τοποθέτηση του προβλήματος της φυσικοποίησης της νόησης στο πλαίσιο της φιλοσοφίας του 20ου αιώνα.....	49
2.3.1 Το πρόβλημα της αποβλεπτικότητας.....	49
2.3.2 Το πρόβλημα των προτασιακών στάσεων.....	50
2.4 Η Αναπαραστασιακή Θεωρία του Νου (ΑΘΝ).....	51
2.4.1 Η υπόθεση της γλώσσας της σκέψης.....	51
2.4.2 Σε τι συνίσταται η γλωσσικότητα των προτάσεων του νου;.....	56
2.5 Η Υπολογιστική Θεωρία του Νου (ΥΘΝ).....	59
2.6 Η Αρχιτεκτονική του Νου.....	63

2.6.1	Τα τρία είδη γνωσιακών συστημάτων	63
2.6.2	Επιχειρήματα υπέρ της ύπαρξης αυτόνομων υπομηχανισμών.69	
2.7	Επίλογος.....	71

ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ: Μια πρώτη προσέγγιση της έννοιας της αναπαράστασης

3.1	Εισαγωγή.....	74
3.2	Προβλήματα με τη δυνατότητα πρόσβασης στις νοητικές αναπαραστάσεις.....	76
3.3	Το πιο οικείο είδος αναπαραστάσεων: εξωτερικές αναπαραστάσεις.....	77
3.3.1	Βασικές διακρίσεις.....	78
3.3.1.1	Η μορφή της αναπαράστασης και ο φορέας της αναπαράστασης.....	78
3.3.1.2	Το περιεχόμενο της αναπαράστασης και η διάκριση νοήματος/αναφοράς	80
3.3.1.3	Η προέλευση της αναπαραστασιακής σχέσης.....	101
3.3.1.3.1	Θεωρίες προέλευσης του νοήματος.....	102
3.3.1.3.2	Θεωρίες προέλευσης της αναφοράς	105
3.3.1.4	Είδη αναπαραστάσεων	106
3.4	Επίλογος.....	114

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ: Το αδιέξοδο του παραδοσιακού προγράμματος αναπαράστασης της γνώσης και η αναγκαιότητα ύπαρξης μη προτασιακής γνώσης.....

4.1	Εισαγωγή.....	118
4.2	Το πρόγραμμα αναπαράστασης της γνώσης στην κλασική τεχνητή νοημοσύνη	119
4.2.1	Από τη συντακτική επεξεργασία της φυσικής γλώσσας στους μικρόκοσμους	120
4.2.2	Αναπαράσταση της γνώσης.....	123
4.2.2.1	Πλαίσια (frames).....	124
4.2.2.2	Εννοιολογικές εξαρτήσεις (conceptual dependencies) και σενάρια (scripts).....	125
4.2.2.3	Προβλήματα που αντιμετώπισαν οι πρώτες τεχνικές αναπαράστασης της γνώσης.....	126
4.2.2.3.1	Το πρόβλημα της κοινότοπης γνώσης	126
4.2.2.3.2	Το πρόβλημα του πλαισίου	131
4.2.2.4	CYC : Το πρόγραμμα κωδικοποίησης της κοινότοπης γνώσης	133
4.3	Υπάρχει λύση στα προβλήματα του πλαισίου και της κοινότοπης γνώσης;	135
4.3.1	Βασικές παραδοχές στις οποίες στηρίζεται η κλασική TN	135
4.3.2	Εναλλακτικές θεωρήσεις της έννοιας της κατάστασης και η αναγκαιότητα ύπαρξης μη-προτασιακής γνώσης.....	137
4.3.3	Μη-προτασιακή γνώση που αφορά τη δομή και τη λειτουργία του σώματος.....	144

4.3.4	Μη-προτασιακή γνώση που προέρχεται από την ένταξη σε ένα κοινωνικό περιβάλλον.....	145
4.4	Επίλογος.....	146

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ: Νοητικά Είδηλα 149

5.1	Εισαγωγή.....	150
5.2	Η υπόθεση περί ύπαρξης νοητικών ειδώλων.....	151
5.2.1	Προϊστορία της υπόθεσης περί ύπαρξης νοητικών ειδώλων..	151
5.2.2	Η διατύπωση της υπόθεσης μέσα στο πλαίσιο της γνωσιακής επιστήμης.....	152
5.2.3	Φαινομενολογικά τεκμήρια.....	155
5.2.3.1	Το επιχείρημα με βάση τη φαινομενολογία των ειδώλων της φαντασίας 156	
5.2.3.2	Το επιχείρημα με βάση την ομοιότητα της φαινομενολογίας της φαντασίας με εκείνη της αντίληψης.....	158
5.2.4	Συμπεριφοριστικά τεκμήρια.....	160
5.2.5	Νευροφυσιολογικά τεκμήρια.....	166
5.2.5.1	Χρήση κοινών εγκεφαλικών μηχανισμών κατά την αντίληψη και τη χρήση νοητικών ειδώλων.....	167
5.2.5.2	Η ανακάλυψη εγκεφαλικών χαρτών που διατηρούν την τοπολογία του αμφιβληστροειδή.....	168
5.3	Τι συνιστά εικόνα;.....	170
5.3.1	Η κυρίαρχη άποψη.....	170
5.4	Επίλογος.....	174

ΚΕΦΑΛΑΙΟ ΕΚΤΟ: Κατανεμημένες Αναπαραστάσεις..... 177

6.1	Εισαγωγή.....	178
6.2	Γιατί τα νευρωνικά δίκτυα αφορούν τη γνωσιακή επιστήμη; 179	
6.3	Τι είναι ένα νευρωνικό δίκτυο;.....	183
6.4	Είδη νευρωνικών δικτύων.....	185
6.5	Πώς μαθαίνουν τα νευρωνικά δίκτυα;.....	187
6.6	Μελέτη ενός νευρωνικού δικτύου αναγνώρισης προσώπων.....	189
6.6.1	Ανάλυση του προβλήματος.....	189
6.6.2	Περιγραφή του δικτύου.....	190
6.6.3	Εκπαίδευση του δικτύου.....	191
6.6.4	Απόδοση του δικτύου.....	191
6.6.5	Πώς κωδικοποιούνται οι πληροφορίες στο εσωτερικό του δικτύου;191	
6.6.5.1	Η σημασιολογία του χώρου φάσεων.....	193
6.7	Τι είναι κατανεμημένη αναπαράσταση;.....	195
6.7.1	Πέντε προσεγγίσεις των κατανεμημένων αναπαραστάσεων..	195
6.7.2	Κριτική αποτίμηση των πέντε προσεγγίσεων της έννοιας της κατανομής.....	201
6.7.3	Τι είδους αναπαραστασιακή μορφή προϋποτίθεται στην έννοια της ισχυρής υπέρθεσης.....	210

6.7.3.1	Αναπαραστασιακή μορφή: ο βαθμός ενεργοποίησης των κόμβων	210
6.7.3.2	Αναπαραστασιακή μορφή: τα βάρη των συνδέσεων μεταξύ των κόμβων	212
6.7.4	Τι είδους περιεχόμενο έχουν οι κατανεμημένες αναπαραστάσεις;	213
6.7.5	Τι είδους περιεχόμενα αναπαρίστανται στις εισόδους και τις εξόδους ενός νευρωνικού δικτύου;	218
6.8	Επίλογος	220

ΚΕΦΑΛΑΙΟ ΕΒΔΟΜΟ: Κριτική των προγραμμάτων φυσικοποίησης του περιεχομένου των νοητικών αναπαραστάσεων

7.1	Εισαγωγή	224
7.2	Ψυχοσημασιολογία: Το πρόβλημα της προέλευσης του περιεχομένου των αναπαραστάσεων	226
7.2.1	Ιντερναλισμός	231
7.2.1.1	Μεθοδολογικός Σολιγισμός	233
7.2.1.2	Η Αρχή της Αυτονομίας	234
7.2.2	Εξτερναλισμός	235
7.2.3	Η αιτιακή - ιστορική θεωρία της αναφοράς ως αφηγηρικό σημείο της διαμάχης ιντερναλισμού - εξτερναλισμού	236
7.2.4	Εξτερναλιστικές προσεγγίσεις	245
7.2.4.1	Σημασιολογία ενδεικτών	246
7.2.4.1.1	Χονδροειδής αιτιακή θεωρία	247
7.2.4.1.2	Η πληροφοριακή σημασιολογία του Dretske	251
7.2.4.1.3	Η θεωρία της ασύμμετρης εξάρτησης	254
7.2.4.2	Σημασιολογία της επιτυχίας (Success semantics)	256
7.2.4.3	Τελεολογική σημασιολογία	262
7.2.4.3.1	Η έννοια της ίδιας λειτουργίας	262
7.2.4.3.2	Μετριοπαθής τελεολογική σημασιολογία	266
7.2.5	Ιντερναλιστικές προσεγγίσεις	274
7.2.5.1	Σημασιολογία εννοιολογικού ρόλου	276
7.2.6	Θεωρία διπλής όψης	281
7.3	Μια κοινή αποδεκτή διατύπωση του ιντερναλισμού και του εξτερναλισμού	281
7.4	Επίλογος	288

ΚΕΦΑΛΑΙΟ ΟΓΔΟΟ: Κοινές παραδοχές του ιντερναλισμού και του εξτερναλισμού και ο δρόμος για μια μη αναπαραστασιακή θεώρηση των νοητικών φαινομένων

8.1	Εισαγωγή	292
8.2	Πρώτη κοινή παραδοχή: Η οπτική του τρίτου προσώπου	292
8.2.1	Υποκειμενικότητα	295
8.2.1.1	Μπορεί η γνώση της πατρότητας να εξηγηθεί από την επιστημονική οπτική του τρίτου προσώπου;	296

8.2.1.1.1	Αποτελεί η γνώση της πατρότητας μια αναπαραστασιακή μορφή γνώσης;	296
8.2.1.1.2	Αποτελεί η γνώση της πατρότητας μια μορφή άμεσης αντίληψης; 301	
8.2.1.2	Πώς μπορεί να εξηγηθεί από την οπτική του πρώτου προσώπου η άμεση γνώση της πατρότητας;.....	305
8.2.1.2.1	Η θεωρία της ενδοσκόπησης	305
8.2.1.2.2	Η θεωρία του αναστοχασμού.....	306
8.2.1.2.3	Η άμεση γνώση της πατρότητας δεν οφείλεται σε ένα πρόσθετο ενέργημα.....	310
8.2.1.3	Ποια είναι η χρησιμότητα της συναίσθησης της πατρότητας; 313	
8.2.1.4	Η συναίσθηση της πατρότητας δεν είναι πρωταρχικά μια συναίσθηση της εσωτερικότητας αλλά μια συναίσθηση της ικανότητας του πράττειν	315
8.2.2	Άμεση Αντίληψη	316
8.2.2.1	Μπορεί η άμεση αντίληψη να εξηγηθεί από την επιστημονική οπτική του τρίτου προσώπου;.....	316
8.2.2.2	Υπάρχει δυνατότητα άμεσης αντίληψης από την οπτική του πρώτου προσώπου;.....	320
8.2.2.2.1	Στο φαινομενολογικό επίπεδο δεν εμφανίζονται νοητικές αναπαραστάσεις.....	322
8.2.2.2.1.1	Κατά την κανονική αντίληψη δεν εμφανίζονται νοητικές αναπαραστάσεις.....	322
8.2.2.2.1.2	Κατά την παραίσθηση δεν εμφανίζονται νοητικές αναπαραστάσεις.....	324
8.2.2.2.2	Στο φαινομενολογικό επίπεδο δεν εμφανίζονται εσωτερικά αντικείμενα	325
8.2.2.2.2.1	Η φαινομενολογία της παραίσθησης διαφέρει από τη φαινομενολογία της κανονικής αντίληψης	326
8.2.2.2.2.2	Κατά την παραίσθηση δεν εμφανίζονται εσωτερικά αντικείμενα.....	334
8.2.2.2.2.3	Στην φαντασία δεν εμφανίζονται εσωτερικά αντικείμενα 335	
8.2.2.3	Τα νοητικά φαινόμενα ως τρόποι φανέρωσης του κόσμου στον άνθρωπο	338
8.3	Δεύτερη κοινή παραδοχή: ταυτότητα μεταξύ νοητικών και λειτουργικών/νευροφυσιολογικών φαινομένων	339
8.4	Επίλογος.....	348

ΕΠΙΛΟΓΟΣ..... 351

ΠΑΡΑΡΤΗΜΑ 1: Η θεωρία σημειογραφίας του Nelson Goodman

.....	363
Εισαγωγή	363
Η επίθεση κατά της έννοιας της ομοιότητας.....	363
Η Θεωρία Σημειογραφίας.....	365

Συντακτικές συνθήκες.....	366
Σημασιολογικές συνθήκες	371
Σε τι διαφέρουν οι εικονιστικές από τις γλωσσικές αναπαραστάσεις;	374
Αναλογικοί και Ψηφιακοί Υπολογιστές.....	376
ΠΑΡΑΡΤΗΜΑ 2: Η προσέγγιση των γνωσιακών διαδικασιών μέσω δυναμικών συστημάτων.....	379
ΒΙΒΛΙΟΓΡΑΦΙΑ	387

ΠΡΟΛΟΓΟΣ

Η παρούσα εργασία ξεκίνησε το 1992. Ωστόσο, τα ερωτήματα που με ώθησαν στη φιλοσοφία είχαν δημιουργηθεί πολύ νωρίτερα ήδη από τα χρόνια του σχολείου. Το βασικό ερώτημα που είχε τραβήξει το ενδιαφέρον μου ήταν το πρώτης όψης γνωσιολογικό. Με δυο λόγια, θα μπορούσα να το συνοψίσω ως εξής: τι ακριβώς γνωρίζει κάποιος που έχει αποκτήσει μια δεξιότητα σε έναν τομέα και πώς φτάνει να την αποκτήσει; Είχα παρατηρήσει ότι δεν αρκεί να μάθει κανείς τη σχετική θεωρία, κάτι παραπάνω χρειαζόταν. Μάλιστα, πολλές φορές παρατηρούσα ανθρώπους που ήταν επιδέξιοι σε κάποιον τομέα, χωρίς να έχουν τη συναφή θεωρητική κατάρτιση. Τα παραδείγματα εκτεινόταν ακόμα και στον χώρο της ηθικής: άνθρωποι με πολύ συγκροτημένες αρχές και απόψεις περί ηθικής έκαναν συχνά τη λάθος κίνηση όταν έπρεπε να αποφασίσουν εν θερμώ για το πως να πράξουν, ενώ, αντίθετα, άλλοι, εντελώς ανυποψίαστοι για όλες αυτές τις αρχές και τις απόψεις, επέλεγαν εκείνο που ήταν αρμόζον κάθε φορά. Έτσι, η "θέση" στην οποία είχα καταλήξει και την οποία επιβεβαίωνα με κάθε ευκαιρία ήταν ότι υπάρχει μια γνώση που χαρακτηρίζει όσους έχουν πείρα και κρίση, η οποία είναι πέραν της γλωσσικά εκφρασμένης γνώσης. Η διάκριση μεταξύ θεωρητικής και πρακτικής γνώσης καθώς και το πρωτείο της δεύτερης έναντι της πρώτης χαράσσονταν όλο και πιο έντονα στο μυαλό μου. Όμως αυτή η πρακτική γνώση ήταν για μένα ένα μυστήριο που δεν μπορούσα να αναλύσω. Εν τω μεταξύ, είχα αρχίσει τις σπουδές μου στο Πολυτεχνείο και ήταν αυτονόητο για μένα εκείνη την εποχή ότι δεν μπορούσα παρά να γίνω μηχανικός. Ωστόσο, ποτέ δεν κατάφερα να νιώσω οικεία σε αυτό το ρόλο. Τα φιλοσοφικά ερωτήματα τραβούσαν πάντα πολύ περισσότερο το ενδιαφέρον μου. Χρειάστηκε, όμως, να φτάσω στο τέλος των φοιτητικών μου χρόνων για να πάρω την απόφαση ότι η φιλοσοφία είναι ο "τόπος" στον οποίο ανήκω. Σε αυτή τη μεταστροφή σημαντικό ρόλο έπαιξε η Κιάρινα Κορδέλα και θέλω να την ευχαριστήσω για αυτό. Έτσι, το 1988, έχοντας σχεδόν τελειώσει με τα μαθήματα στο Τμήμα Μηχανολόγων Μηχανικών του Ε.Μ.Π. και μια εβδομάδα πριν αρχίσω την ήδη συμφωνημένη διπλωματική εργασία στην Αεροδυναμική, αποφάσισα να αθετήσω το λόγο μου και να ξεκινήσω διπλωματική στη φιλοσοφία με τον Αριστείδη Μπαλτά. Αυτή η εργασία πήρε δύο χρόνια για να ολοκληρωθεί, καθώς ένοιωθα ότι έπρεπε με κάθε τρόπο να αναπληρώσω την έλλειψη συστηματικών σπουδών στη φιλοσοφία. Το θέμα της διπλωματικής ήταν η κριτική των φυσικών επιστημών και του εξαντικειμενικευτικού-παραστασιακού τρόπου σκέπτεσθαι από τον Heidegger. Τον επόμενο χρόνο, το 1991, έφυγα στην Αμερική για μεταπτυχιακές σπουδές φιλοσοφίας στο πανεπιστήμιο Johns Hopkins. Για προσωπικούς λόγους διέκοψα τις εκεί σπουδές μου και επέστρεψα στην Ελλάδα όπου το 1992 άρχισα το διδακτορικό στο Πρόγραμμα Φιλοσοφίας και Ιστορίας των Επιστημών και της Τεχνολογίας (ΠΦΙΕΤ) του Ε.Μ.Π.. Με εντελώς γενικούς όρους, το θέμα που με ενδιέφερε αρχικά ήταν η ανάλυση των νοητικών φαινομένων με βάση τον εξαντικειμενικευτικό-παραστασιακό τρόπο σκέπτεσθαι και η

κριτική αυτής της προσέγγισης. Τα πρώτα δύο χρόνια το ενδιαφέρον μου εστιάστηκε κυρίως στην κριτική που ασκεί ο Dreyfus στην τεχνητή νοημοσύνη και τη γνωσιακή επιστήμη. Κατά τη διετία 1994-1995 υπηρέτησα τη στρατιωτική μου θητεία και, αναγκαστικά, η έρευνά μου διακόπηκε. Όταν ξαναρχισα το 1996 είχα πλέον την αίσθηση ότι η κριτική που ασκεί ο Dreyfus είναι μια "εξωτερική" κριτική που προσεγγίζει πολύ σχηματικά τη γνωσιακή επιστήμη και γι' αυτό δεν καταφέρνει να καταστεί πειστική για όσους δουλεύουν μέσα στο πλαίσió της. Αποφάσισα, λοιπόν, ότι έπρεπε να ασχοληθώ με τη γνωσιακή επιστήμη "από τα μέσα". Αυτή ήταν μια επώδυνη απόφαση γιατί ήταν σαν να ξαναρχίζω το διδακτορικό μου από την αρχή και κυρίως γιατί αναγκάστηκα να διαβάσω ένα πολύ μεγάλο αριθμό βιβλίων και άρθρων τα οποία όχι μόνο ήταν αντίθετα στην προσέγγιση που ήθελα να αναπτύξω αλλά μου φαινόταν και ότι διαπράττουν όλα τα λάθη που είχαν εντοπίσει καλύτερα μεγάλοι φιλόσοφοι όπως ο Frege, ο Husserl, ο Heidegger και ο Wittgenstein. Φυσικά τα πράγματα δεν ήταν τόσο απλά. Σιγά-σιγά άρχισα να ανακαλύπτω την αξία αρκετών από τα διαβάσματά μου και να εξοικειώνομαι με το διεπιστημονικό χώρο της γνωσιακής επιστήμης. Χρειάστηκε τελικά πάρα πολύς κόπος για να καταφέρω να βάλω μια τάξη στις τρέχουσες απόψεις αναφορικά με το πώς υποτίθεται ότι αναπαρίσταται η γνώση στο νου και για να συνθέσω τις διάφορες κριτικές. Το εάν αυτή η προσπάθεια αντανακλάται στο τελικό κείμενο είναι κάτι που θα το κρίνει ο αναγνώστης.

Πολλοί είναι αυτοί που συνέβαλλαν, ο καθένας με τον δικό του τρόπο, στην ολοκλήρωση της διατριβής. Θα ξεκινήσω από τους "συναδέλφους" μου, ορισμένοι από τους οποίους έχουν ήδη ενταχθεί στο διδακτικό προσωπικό του πανεπιστημίου. Θα ήθελα να ευχαριστήσω θερμά τους Πάνο Θεοδώρου, Γιώργο Μαλάμη, Κατερίνα Μπαντινάκη, Δημήτρη Παταγιαννάκο και Γιάννη Σκουμπούρη για τις συζητήσεις μας και τα σχολιά τους σε κείμενά μου. Ο Γιώργος Φουρτούνης, που τον γνώρισα το 1988 στο σεμινάριο επιστημολογίας του Πολυτεχνείου, είναι ένας από τους πιο διεισδυτικούς συνομιλητές και αναγνώστες που έχω συναντήσει γι' αυτό οι φιλοσοφικές συζητήσεις μας μου έδιναν πάντα πάρα πολλά, μακάρι να ήταν περισσότερες. Η Λάρα Σκουρλά εμφανίστηκε στο μεταπτυχιακό γύρω στα 1993 και μου συμπαραστάθηκε σε όλες σχεδόν τις φάσεις του διδακτορικού μου. Θα ήθελα να την ευχαριστήσω θερμά για την εμπιστοσύνη και τη στήριξη που μου παρείχε καθώς και για την ενέργεια που αφιέρωσε στην ανάγνωση και το σχολιασμό διαδοχικών εκδοχών της διατριβής μου. Η εξαιρετική κρίση της με βοήθησε πάρα πολύ όλα αυτά τα χρόνια. Ο Βασίλης Ραϊσης εντάχθηκε την ίδια περίπου περίοδο στο μεταπτυχιακό πρόγραμμα. Είναι ο άνθρωπος που έχει ακούσει μάλλον τις περισσότερες εκδοχές του διδακτορικού μου. Στις συχνές συζητήσεις μας τα τελευταία χρόνια, όντας υπέρμαχος - όταν κάνει φιλοσοφία - μιας απομαγευμένης φύσης, προέβαλλε σχεδόν πάντα ένα πλήθος αντιρρήσεων που με βοήθησαν να ξεκαθαρίσω τις θέσεις μου. Σημαντικότερη, επίσης, ήταν η βοήθεια του Σπύρου Πετρονάκου που τον γνώρισα το 1997. Στις τακτικές συζητήσεις μας, ειδικότερα τον τελευταίο χρόνο, με βοήθησε να

επεξεργαστώ τα θέματα του όγδου κεφαλαίου προσφέροντας διακριτικά λύσεις. Η Δήμητρα Σπανού με βοήθησε με τη συμπεριφορά της και τις συνομιλίες μας να εξετάσω στην πράξη ορισμένα ζητήματα που αφορούν τη διατριβή μου.

Θα ήθελα επίσης να ευχαριστήσω τους Πέτρο Δαμιανό, Φωτεινή Θεοδώρου-Βασιλείου, Μάριο Καλιακάτσο, Ιπποκράτη Κλήμη, Βασίλη Κόλλια, Σωτήρη Κότσανη, Βίκυ Νασούλη, Βάιο Παπαθεοχάρη, Πέτρο Πολυμένη, Μαρία Πουρνάρη, Μαρία Ρεντετζή, Δήμητρα Σγουρούδη, Χρήστο Σταματέλο, Άννα Τηγάνη, Αναστασία Χριστοπόλου και Χάρη Χρόνη για τις συζητήσεις μας και τα σχόλια τους πάνω στη διατριβή μου.

Ένα από τα μεγάλα πλεονεκτήματα του μεταπτυχιακού προγράμματος Φιλοσοφίας και Ιστορίας των Επιστημών και της Τεχνολογίας είναι ότι κάθε χρόνο προσκαλεί σημαντικότερους καθηγητές του εξωτερικού για διαλέξεις και σεμινάρια. Αρκετοί από αυτούς με βοήθησαν στην εργασία μου και θέλω να τους ευχαριστήσω. Πρόκειται για τους Peter Machamer, John Haugeland, Adrian Cussins, Michael Wheeler, John McDowell, Pascal Engel, Michael Frede, James Conant, C.A. Hooker. Θα ήθελα, ιδιαίτερα να σταθώ στον Peter Machamer ο οποίος μέσα από τις διαλέξεις και τα σεμινάρια του και κυρίως μέσα από τις έντονες συνομιλίες μας μου παρείχε πάντα με τη διαύγεια που τον διακρίνει την αντίσταση που χρειαζόμουν για να προχωρήσω. Ευχαριστώ, επίσης, τους Álvaro Moreno και Ruth Millikan για τα σχόλιά τους σε ορισμένα ζητήματα της διατριβής.

Ο διεπιστημονικός χαρακτήρας της γνωσιακής επιστήμης δημιούργησε την ανάγκη να έρθω σε επαφή με επιστήμονες διάφορων ειδικοτήτων. Ένας χώρος όπου προωθείται η διεπιστημονική ανταλλαγή ιδεών είναι η "Ομάδα Γουδί" με τις εβδομαδιαίες διαλέξεις που γίνονται σταθερά τα τελευταία χρόνια. Η παρακολούθηση αρκετών από αυτές τις διαλέξεις από το 1997 έως σήμερα με έφερε σε επαφή με διαφορετικές προσεγγίσεις μέσα στον χώρο της γνωσιακής επιστήμης. Ευχαριστώ τους βασικούς διοργανωτές της συγκεκριμένης σειράς διαλέξεων καθηγητές Θανάση Τζαβάρια και Γιώργο Παπαδόπουλο για τις συζητήσεις μας τα τελευταία χρόνια καθώς και για την πρόσκληση που μου απηύθυναν δύο φορές να μιλήσω στο πλαίσιο της σειράς. Θα ήθελα, επίσης, να ευχαριστήσω το Γιάννη Μαϊστρο και τη Στέλλα Μαρκαντιανάτου που από το 1996 μέχρι σήμερα με βοήθησαν να κατανοήσω ζητήματα που αφορούν την αναπαράσταση γνώσης στη τεχνητή νοημοσύνη και την υπολογιστική γλωσσολογία. Η συνεργασία μας ξεκίνησε με το ερευνητικό πρόγραμμα "Αναπαράσταση γνώσης στον υπολογιστή: προτασιακές και μη προτασιακές μορφές αναπαράστασης", στη συνέχεια μου ζήτησαν να συμβάλω στο σχεδιασμό και τη διδασκαλία του σεμιναρίου "Παράσταση Γνώσης στον Υπολογιστή" που έγινε το χειμερινό εξάμηνο του 1999 στο πλαίσιο των μεταπτυχιακών προγραμμάτων "Βασική και Εφαρμοσμένη Γνωσιακή Επιστήμη" και "Φιλοσοφία και Ιστορία των Επιστημών και της Τεχνολογίας" και, τέλος, το 2000 με επέλεξαν ως μέλος του προγράμματος βασικής έρευνας "Μεταφραστική Βάση Γνώσης για Ρηματικά Ζεύγη της

Ελληνικής και Αγγλικής Γλώσσας". Ευχαριστίες οφείλω, επίσης, στον καθηγητή νευρωνικών δικτύων του Τμήματος Ηλεκτρολόγων Μηχανικών και Ηλεκτρονικών Υπολογιστών του Ε.Μ.Π. Ανδρέα Σταφυλοπάτη καθώς και στο Γιώργο Μαγουλά για τα σχόλιά τους στο έκτο κεφάλαιο της διατριβής. Η Δήμητρα Κατή με προσκάλεσε το 1997 να μιλήσω στη διεπιστημονική διημερίδα με τίτλο "Νόηση και Γλώσσα" και στη συνέχεια συμπεριέλαβε την ανακοίνωσή μου με τίτλο "Υπάρχει Γλώσσα της σκέψης;" στο τόμο που δημοσίευσε με την Μαριάννα Κονδύλη και την Κυκή Νικηφορίδου. Τμήματα αυτού του άρθρου έχουν συμπεριληφθεί στη διατριβή. Οι συζητήσεις που είχα με τη Δήμητρα και τα σχόλιά της στο άρθρο μου με βοήθησαν να ξεκαθαρίσω ζητήματα που αφορούν διάφορες προσεγγίσεις στη ψυχολογία περί γνωστικών αρχιτεκτονικών.

Ευχαριστώ το Δρακούλη Νικολινάκο ο οποίος με προσκάλεσε να συμβάλλω στο σχεδιασμό και τη διδασκαλία του σεμιναρίου "Αναπαράσταση και Προθετικότητα" που έγινε το εαρινό εξάμηνο του 1999 στο πλαίσιο των δύο προαναφερθέντων μεταπτυχιακών προγραμμάτων. Μέσα από τις συζητήσεις μας και την επαφή μου με την ενημερωμένη βιβλιοθήκη του ωφελήθηκα σημαντικά. Ευχαριστίες οφείλω επίσης στους κ.κ. Μιχάλη Ασημακόπουλο, Κατερίνα Ιεροδιακόνου, Βασιλή Καρασμάνη, Βαγγέλη Καφετζόπουλο, Γιώργο Κολέτσο, Κώστα Κρμπά, Παντελή Μπασάκο, Πανταζή Τσελεμάνη και Παύλο Χριστοδουλίδη οι οποίοι με τα σεμινάρια τους και τις συζητήσεις μας με βοήθησαν να διευκρινίσω ζητήματα που αφορούν τη διατριβή μου. Ο Άρης Αραγιώργης διάβασε από το πρώτο κεφάλαιο την ενότητα για τη μαθηματική λογική. Τον ευχαριστώ θερμά για τις παρατηρήσεις του. Ευχαριστίες οφείλω και στη Φαίη Ζήκα η οποία με βοήθησε σημαντικά λίγο πριν το τέλος της διατριβής μου.

Με τον Άρη Κουτούγκο κάναμε αρκετές συζητήσεις τα τελευταία χρόνια στον ΑΚΕΔ και η συμφωνία μας σε πολλά ζητήματα μου έδινε θάρρος να συνεχίσω. Θα ήθελα να ευχαριστήσω ιδιαίτερα το Γιώργο Ξηροπαίδη ο οποίος με τα συστηματικά και υψηλότατου επιπέδου σεμινάρια που έκανε από το 1993 έως το 2000 έπαιξε καθοριστικό ρόλο στη φιλοσοφική μου συγκρότηση και με επηρέασε βαθιά. Ο Γιώργος μού συμπαραστάθηκε και με στήριξε όλα αυτά τα χρόνια όχι απλά ως ένας δάσκαλος. Η Βάσω Κιντή με έφερε σε επαφή με το έργο του Wittgenstein και με βοήθησε άμεσα με πολλούς τρόπους στη συγγραφή της διατριβής. Η γενικότερη στάση της απέναντί μου από το 1988 που τη γνώρισα μέχρι σήμερα είναι πολύ σημαντική για μένα.

Ευχαριστίες οφείλω στην επταμελή μου επιτροπή. Θα ξεκινήσω από τα τέσσερα μέλη που συμπληρώνουν την τριμελή συμβουλευτική επιτροπή. Ευχαριστώ θερμά το Στέλιο Βιρβιδάκη ο οποίος ήδη από το 1988 με κατεύθυνε στην ανάγνωση του Kant και λίγο αργότερα με τα σεμινάρια του στο Deree με εισήγαγε στην αγγλοσαξονική γνωσιολογία. Ο Στέλιος συνέχισε σταθερά μέχρι σήμερα να με βοηθά με τις ανεξάντλητες γνώσεις και την κρίση του σε όλες τις φάσεις του διδακτορικού. Η Βούλα Τσινόρεμα διάβασε διάφορα κείμενά μου και με βοήθησε πολύ με τα σχόλιά της τα τελευταία χρόνια. Η καθηγήτρια Στέλλα Βοσνιάδου οργάνωσε το

μεταπτυχιακό πρόγραμμα Βασικής και Εφαρμοσμένης Γνωστικής Επιστήμης στο πλαίσιο του οποίου έγιναν πολλές διαλέξεις και σεμινάρια της ίδιας και άλλων που με βοήθησαν σημαντικά. Την ευχαριστώ θερμά για τις παρατηρήσεις που έκανε στη διατριβή καθώς και για την εμπιστοσύνη και την ευθύτητά της απέναντί μου. Τέλος, θα ήθελα να ευχαριστήσω τον καθηγητή Διονύσιο Αναπολιτάνο για τα διεισδυτικά σχόλιά του στη διατριβή μου. Η αδιάκοπη προσπάθειά του όλα αυτά τα χρόνια να βελτιώσει το διαπανεπιστημιακό μεταπτυχιακό πρόγραμμα Φιλοσοφίας και Ιστορίας των Επιστημών και της Τεχνολογίας σίγουρα επηρέασε και μένα ως μέλος του προγράμματος.

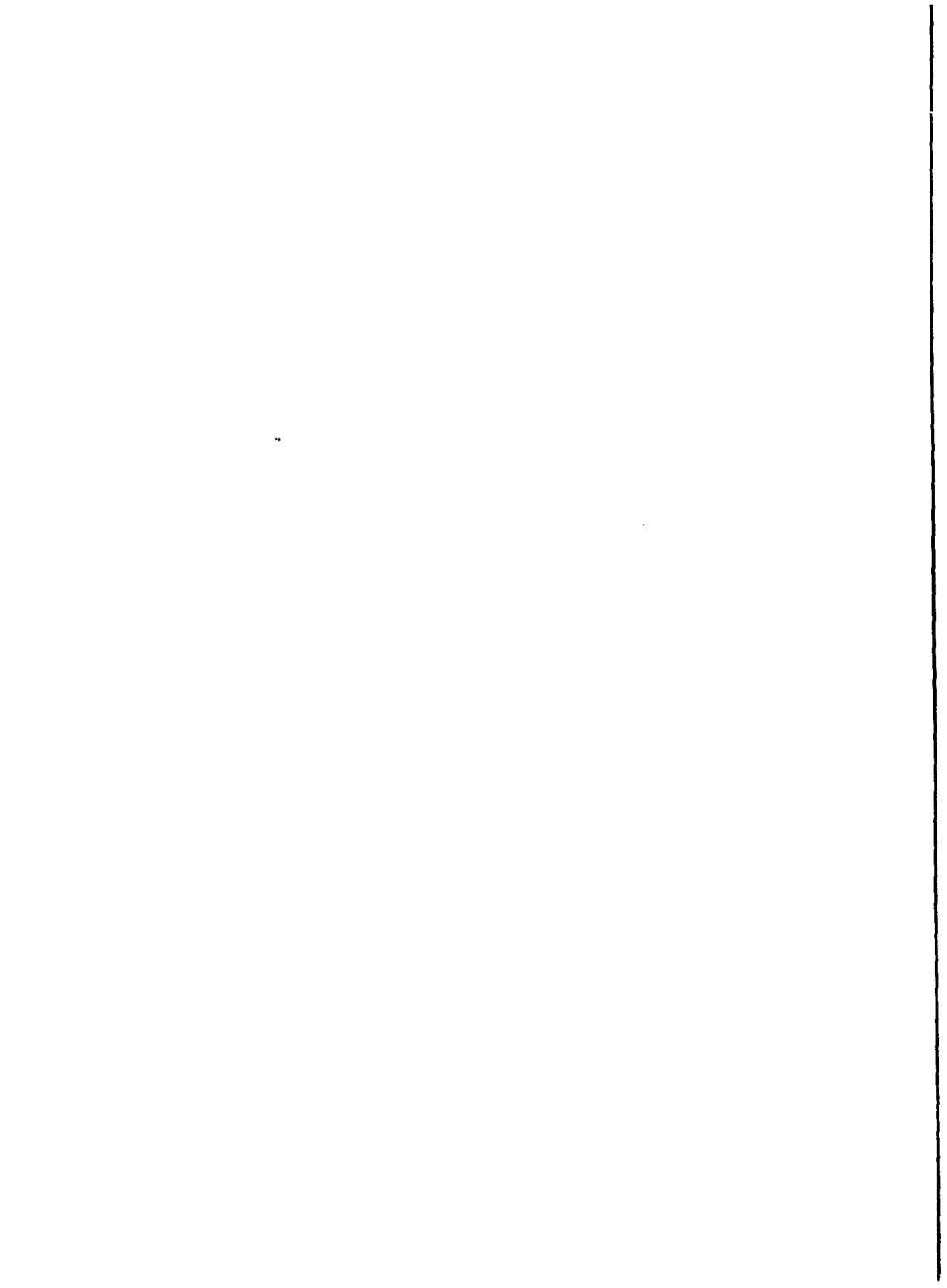
Η τριμελής συμβουλευτική μου επιτροπή απαρτίζεται από τον Αριστείδη Μπαλτά, επιβλέποντα της διατριβής, τον Παντελή Νικολακόπουλο και τον Κώστα Γαβρόγλου. Ο Παντελής Νικολακόπουλος, που τόσο αναπάντεχα έφυγε από κοντά μας, με βοήθησε πάρα πολύ από τα πρώτα χρόνια που εμφανίστηκα στο μεταπτυχιακό. Τα σεμινάρια του στη νεώτερη φιλοσοφία ήταν πηγή σκέψης και έμπνευσης. Ειδικότερα, τα τελευταία χρόνια που άρχισα να του δίνω κείμενα, οι επαφές μας εντατικοποιήθηκαν. Η καθαρή του σκέψη και οι καίριες παρατηρήσεις του μου έδιναν πάντα πολύ μεγάλη βοήθεια και ικανοποίηση. Ο Παντελής μερμινούσε μέχρι το τέλος για την πορεία της διατριβής μου παρά την κατάσταση της υγείας του. Η στάση του αυτή θα μείνει για πάντα στη μνήμη μου. Το δεύτερο μέλος της συμβουλευτικής μου επιτροπής είναι ο Κώστας Γαβρόγλου. Θα ήθελα να τον ευχαριστήσω για τη σταθερή υποστήριξή του και όλα όσα μου προσέφερε.

Τέλος θα ήθελα να αναφερθώ στον επιβλέποντα καθηγητή μου Αριστείδη Μπαλτά. Δεν είναι εύκολο να βρω λόγια να ευχαριστήσω τον Αριστείδη ο οποίος έπαιξε καθοριστικό ρόλο στη ζωή μου από την πρώτη στιγμή που τον γνώρισα το 1988. Χωρίς την παρουσία του σίγουρα δεν θα είχα διανοηθεί να ξεκινήσω διδακτορικό φιλοσοφίας στην Ελλάδα. Οι αμέτρητες συζητήσεις μας όλα αυτά τα χρόνια με διαμόρφωσαν και με επηρέασαν σε πολλά διαφορετικά επίπεδα που δεν αφορούν μόνο τη φιλοσοφία αλλά και τη γενικότερη στάση ζωής. Για όλα αυτά και πολλά άλλα θέλω να τον ευχαριστήσω θερμά.

Ευχαριστώ τους φίλους μου Θρασύβουλο Παπαδόπουλο, Ολυμπία Λαμπούση και Βάλια Δημοπούλου για τη στήριξη που μου παρείχαν όλα αυτά τα χρόνια. Ο Φραγκίσκος Σπανός με βοήθησε υπομονετικά σε πολλά τεχνικά ζητήματα και τον ευχαριστώ θερμά.

Η Ελευθερία Ασρινάκη με έκανε να πιστέψω ότι η διατριβή μου έχει ένα τέλος. Την ευχαριστώ για όλα.

Τέλος θα ήθελα να ευχαριστήσω ολόψυχα τους γονείς μου Κυριάκο και Μαργαρίτα Παγωνιδιώτου καθώς και την αδελφή μου Ευαγγελία Παγωνιδιώτου για την ακαταπόνητη στήριξη που μου παρείχαν όλα αυτά τα χρόνια, μια στήριξη που ήταν τόσο βαθιά ώστε να εμφανίζεται ως αυτονόητη.



ΕΙΣΑΓΩΓΗ

Όποιος ασχοληθεί έστω και λίγο με τη σύγχρονη φιλοσοφία του νου διαπιστώνει ότι αποτελεί ένα εξαιρετικά ετερογενές πεδίο με επιρροές τόσο από την παλαιότερη φιλοσοφία όσο και από τις σύγχρονες επιστήμες του νου, του εγκεφάλου και της συμπεριφοράς, επιστήμες που όλο και περισσότερο εντάσσονται σε ό,τι αποκαλείται "γνωσιακή επιστήμη". Η σύγχρονη φιλοσοφία του νου βρίσκεται σε αναβρασμό και υπό διαμόρφωση. Στο εσωτερικό της μαινονται ένα πλήθος από διαμάχες που συγκροτούνται γύρω από εννοιολογικά δίπολα του τύπου: έμφυτο / επίκτητο, συμβολικές / κατανεμημένες αναπαραστάσεις, συμβολικές / εικονιστικές αναπαραστάσεις, σημασιολογία ενδεικτών / τελεολογική σημασιολογία, κλασική αρχιτεκτονική / δυναμικά συστήματα, ιντερναλισμός / εξτερναλισμός, εννοιολογικό / μη εννοιολογικό περιεχόμενο, φαινομενική συνείδηση / συνείδηση πρόσβασης κ.λπ. Κάθε μια από αυτές τις διαμάχες, τις περισσότερες φορές, διεξάγεται ανεξάρτητα από τις υπόλοιπες, γεγονός που συμβάλλει στην εικόνα πολυδιάσπασης που παρουσιάζει η σύγχρονη φιλοσοφία του νου. Αυτό συμβαίνει, ως ένα βαθμό, γιατί οι διαμάχες βρίσκονται εν εξελίξει και δεν υπάρχουν παγιωμένες απαντήσεις. Ένας βαθύτερος λόγος, όμως, που οι εμπλεκόμενοι σε μια διαμάχη συχνά δεν λαμβάνουν υπόψη τα πορίσματα και τα αδιέξοδα που προκύπτουν από τις άλλες είναι η υιοθέτηση μιας βασικής παραδοχής που διαπνέει μεγάλο μέρος της αναλυτικής φιλοσοφίας. Πρόκειται για την παραδοχή ότι μεγάλα προβλήματα μπορούν να επιλυθούν με επιμερισμό σε μικρότερα και ανεξάρτητα μεταξύ τους προβλήματα. Αυτή η μέθοδος αποτέλεσε μια από τις βάσεις για την ανάπτυξη και επιτυχία των φυσικών επιστημών. Μάλιστα στις μέρες μας, ορισμένοι γνωσιακοί επιστήμονες έχουν φτάσει να υποστηρίξουν ότι ο ίδιος ο εγκέφαλος λειτουργεί με τον παραπάνω τρόπο: Ο εγκέφαλος απαρτίζεται από ένα σύνολο αυτόνομων γνωσιακών υπομηχανισμών (cognitive modules) που καθένας επιλύει συγκεκριμένα προβλήματα.

Ανεξάρτητα από το πώς τελικά λειτουργεί ο ανθρώπινος εγκέφαλος, θεωρούμε ότι η ανάλυση ενός σύνθετου προβλήματος σε επιμέρους προβλήματα αποτελεί ένα εντελώς απαραίτητο μέσο για την κατανόηση των νοητικών φαινομένων. Ωστόσο, αυτή η μέθοδος οδηγεί σε μια σκέψη σχηματική και αποσπασματική στο χώρο της φιλοσοφίας, αν δεν καταβληθεί εξ' ίσου προσπάθεια να αρθρωθεί μια συνολικότερη "εικόνα" μέσα από την αξιολόγηση της συνεκτικότητας των απαντήσεων που δίδονται στα επιμέρους προβλήματα αλλά και των αδιεξόδων στα οποία οδηγούνται ορισμένοι τρόποι ερωτηματοθεσίας. Η σύγχρονη φιλοσοφία του νου και, ειδικότερα, εκείνο το κυρίαρχο τμήμα της που ασχολείται με τη γνωσιακή επιστήμη υστερεί σημαντικά σε αυτό το συνθετικό έργο για έναν πρόσθετο λόγο: η ίδια η γνωσιακή επιστήμη, λόγω του ότι βρίσκεται ακόμα υπό διαμόρφωση, χαρακτηρίζεται από μια εκρηκτική συσσώρευση νέων θεωριών και τεχνικών οι οποίες συχνά δεν είναι συμβατές μεταξύ τους. Η γνωσιακή επιστήμη απέχει ακόμα πολύ από

ΤΟ ΠΡΟΒΛΗΜΑ ΤΩΝ ΝΟΗΤΙΚΩΝ ΑΝΑΡΑΓΜΑΤΩΝ ΣΤΗ ΓΥΝΕΣΙΑΚΗ ΕΠΙΣΤΗΜΗ

Ελβετικής και Αγγλικής Γλώσσας". Ευχαριστίες οφείλω, επίσης, στον διευθυντή νευρωνικών δικτύων του Τμήματος Ηλεκτρολόγων Μηχανικών και Ηλεκτρονικών Υπολογιστών του Ε.Μ.Π. Ανδρέα Σταφυλοπούλη καθώς και στο Γιώργο Μανουσά για τα σχόλιά τους στο έκτο κεφάλαιο στη διατριβή. Η Δήμητρα Κατή με προσκάλεσε το 1997 να μιλήσω στη διεπιστημονική διημερίδα που με τίτλο "Νόηση και Γλώσσα" και την Κική Νικηφορίδου. Τμήματά μου με τίτλο "Μαρίνα Κονδύλη στη διατριβή μου. Η αδιάκοπη προσπάθειά του όλα αυτά τα χρόνια να βελτιώσει το διαπαντασχηματικό μεταπτυχιακό πρόγραμμα Φιλοσοφίας και Ιστορίας των Επιστημών και της Τεχνολογίας σίγουρα επηρέασε και μένα ως μέλος του προγράμματος.

Η τριμελής συμβουλευτική μου επιτροπή αποτελείται από τον Αριστοτέλη Μπαλά, επιβλέπωντα της διατριβής, τον Παντελή Νικολακόπουλο και τον Κώστα Γαβρόγλου. Ο Παντελής Νικολακόπουλος που τόσο αντανάγχε έφινε από κοντά μας, με βοήθησε πάρα πολύ από τη ζωή χρόνια που εμφανίστηκα στο μεταπτυχιακό. Τα σεμινάρια, τα νεότερη φιλοσοφία ήταν πηγή σκέψης και έμπνευσης. Ειδικότερα, τα τελευταία χρόνια που άρχισα να του δίνω κείμενα, οι επαφές μας εντατικοποιήθηκαν. Η καθαρή του σκέψη και οι καιρικές παρατηρήσεις του μου έδιναν πάντα πολύ μεγάλη βοήθεια και ικανοποίηση. Ο Παντελής παρακίνησε μέχρι το τέλος για την πορεία της διατριβής μου παρά τη κατάσταση της υγείας του. Η στάση του αυτή θα μείνει για πάντα χαραγμένη μου. Το δεύτερο μέλος της συμβουλευτικής μου επιτροπής είναι ο Κώστας Γαβρόγλου. Θα ήθελα να τον ευχαριστήσω για τη σταθερή υποστήριξή του και όλα όσα μου προσέφερε.

Τέλος θα ήθελα να αναφερθώ στον επιβλέποντα καθηγητή μου Αριστέδη Μπαλά. Δεν είναι εύκολο να ορίσω τον ρόλο μου από την πρώτη στιγμή που τον γνώρισα το 1988. Χωρίς την παρουσία του σίγουρα δεν θα είχα διανοηθεί να ξεκινήσω διδακταρικό φιλοσοφίας στην Ελλάδα. Οι αμέτρητες συζητήσεις μας όλα αυτά τα χρόνια με διαμόρφωσαν και με επηρέασαν σε πολλά διαφορετικά στάδια ζωής. Για όλα αυτά και πολλά φιλοσοφία αλλά και τη γενικότερη κατάσταση Ολυμπία Ευχαριστώ τους φίλους μου Θρασύβουλο Παπαδόπουλο, Ολυμπία Ζαμπούση και Βάλια Δημοπούλου για τη στήριξη που μου παρέχουν όλα αυτά τα χρόνια. Ο Φραγκίσκος Στανός με βοήθησε υπομονετικά σε πολλά τεχνικά ζητήματα και τον ευχαριστώ ιδιαίτερα για τη βοήθειά μου Ευαγγελία Η Έλευθερία Ασρινάκη για όλα.

Η Έλευθερία Ασρινάκη και την αδελφή μου Ευαγγελία Τέλος θα ήθελα να ευχαριστήσω ολόψυχα τους γονείς μου Κυριάκο και Μαργαρίτα Παγωνιδιώτη για την ακαταπόνητη στήριξη που μου παρέχουν όλα αυτά τα χρόνια, ώστε να εμφανίζεται ως Παγωνιδιώτου για την ακαταπόνητη στήριξη που ήταν τόσο βαθιά ώστε να εμφανίζεται ως Γλόνια, μια στήριξη που ήταν τόσο βαθιά ώστε να εμφανίζεται ως αυτονόητη.

Αθήνα 13 Οκτωβρίου 2001

Επικυρώθηκε και ημερομηνία και υπογραφή

το να καταστεί γνήσια διεπιστημονική έρευνα. Οι συμβαλλόμενες επιστήμες, κυριότερες από τις οποίες είναι η ψυχολογία, η γλωσσολογία, η νευροεπιστήμη και η πληροφορική - τεχνητή νοημοσύνη δεν συνεργάζονται, ως επί το πλείστον, σοβαρά μεταξύ τους. Εξαιρέση αποτελεί ένα τμήμα της πληροφορικής που παρέχει την τεχνολογία για την προσομοίωση των μοντέλων που αναπτύσσονται από τις υπόλοιπες επιστήμες. Οι τομείς της ψυχολογίας, της γλωσσολογίας και της νευροεπιστήμης που αυτοχαρακτηρίζονται "γνωσιακοί" δεν συνεργάζονται συνήθως στενά μεταξύ τους, παρά μόνο σε ορισμένα εξειδικευμένα ερευνητικά κέντρα, κυρίως των Η.Π.Α.. Έτσι, ο χαρακτηρισμός μιας επιστήμης ως "γνωσιακής" δεν συνεπάγεται, κατ' ανάγκη, ότι πρόκειται για διεπιστημονική προσέγγιση. Γι' αυτό είναι ίσως ορθότερο να ισχυρισθούμε ότι, επί του παρόντος, υπάρχουν πολλές γνωσιακές επιστήμες και όχι μία. Δεν είναι εξάλλου τυχαίο ότι η πληρέστερη και πιο πρόσφατη εγκυκλοπαίδεια που έχει εκδοθεί για το συγκεκριμένο πεδίο ερευνών τιτλοφορείται *Η Εγκυκλοπαίδεια του MIT για τις Γνωσιακές Επιστήμες*¹.

Η μη επαρκής ανάπτυξη της διεπιστημονικής συνεργασίας μεταξύ των γνωσιακών επιστημών δεν σημαίνει φυσικά ότι δεν υπάρχουν κάποιοι κοινοί θεματικοί και μεθοδολογικοί άξονες μεταξύ αυτών των επιστημών που δικαιολογούν τον προσδιορισμό τους ως "γνωσιακών". Η ανάδειξη και η συσχέτιση των κοινών παραδοχών που διέπουν τις γνωσιακές επιστήμες - του πυρήνα που συνιστά τη γνωσιοκρατία (cognitivism) - αποτελεί κατεξοχήν αρμοδιότητα της φιλοσοφίας. Η φιλοσοφία των γνωσιακών επιστημών στο βαθμό που παίζει αυτόν το συνθετικό ρόλο λειτουργεί "συγκολλητικά" μεταξύ των συναφών επιστημών και συμβάλλει στην ανάπτυξη της διεπιστημονικότητας. Ωστόσο, όπως προαναφέραμε, τόσο ο κατακερματισμός των προβλημάτων μέσα στον χώρο της φιλοσοφίας του νου όσο και η εκρηκτική συσσώρευση νέων θεωριών από τις διάφορες γνωσιακές επιστήμες καθιστούν αυτό το εγχείρημα εξαιρετικά δύσκολο. Γι' αυτό, η συνολικότερη "εικόνα" που κυριαρχεί στη φιλοσοφία του νου είναι αρκετά αδρή. Πρόκειται για την Αναπαραστασιακή - Υπολογιστική Θεωρία του Νου (AYΘN) η οποία σχηματικά μπορεί να προσδιοριστεί με την ακόλουθη θέση: τα νοητικά φαινόμενα δεν είναι τίποτε άλλο παρά η κατοχή και ο υπολογιστικός χειρισμός νοητικών αναπαραστάσεων (mental representations). Παρότι ονομάζεται "θεωρία", η AYΘN αποτελεί περισσότερο ένα προσχέδιο θεωρίας που σκιαγραφεί το φιλοσοφικό πλαίσιο των γνωσιακών επιστημών και τις κοινές παραδοχές στις οποίες στηρίζονται. Αυτή η "εικόνα" εγκαθιδρύθηκε τα πρώτα χρόνια ανάπτυξης της γνωσιοκρατικής προσέγγισης, αλλά έχει πάψει πλέον να ενημερώνεται από τις απαντήσεις και να διδάσκεται από τα αδιέξοδα των επιμέρους φιλοσοφικών και επιστημονικών προβλημάτων και έχει μετατραπεί

¹ Wilson & Keil 1999.

περισσότερο σε ένα δόγμα. Πρόκειται για ένα δόγμα, ωστόσο, που εξακολουθεί να διαπνέει και να καθοδηγεί το ερευνητικό έργο της πλειονότητας των φιλοσόφων του νου και των γνωσιακών επιστημόνων. Ελλείψει εναλλακτικής "εικόνας" οι ερευνητές επαφίονται στην ΑΥΘΝ όπως ο καθημερινός άνθρωπος επαφίεται στον κοινό νου.

Στην παρούσα διατριβή έχουμε δύο κύριους στόχους: πρώτον, να αμφισβητήσουμε την ορθότητα της ΑΥΘΝ ασκώντας κριτική στην κεντρική της παραδοχή περί ύπαρξης νοητικών αναπαραστάσεων και, δεύτερον, να σκιαγραφήσουμε μια διαφορετική συνολικότερη "εικόνα" για τα νοητικά φαινόμενα η οποία δεν βασίζεται στην έννοια των νοητικών αναπαραστάσεων και, κατά συνέπεια, δρα διαλυτικά όσον αφορά τα υπάρχοντα προβλήματα γύρω από αυτήν την έννοια. Πρόκειται για μια "εικόνα" η οποία έχει αρχίσει τα τελευταία χρόνια να κερδίζει έδαφος μέσα στον χώρο της φιλοσοφίας του νου και να χρησιμοποιείται κυρίως ως φιλοσοφικό συμπλήρωμα μη αναπαραστασιακών προσεγγίσεων που αναπτύσσονται στο χώρο της γνωσιακής επιστήμης. Ωστόσο, αυτή η χρήση είναι περισσότερο συνθηματολογική και δεν δικαιολογείται από τα επιτεύγματα των συγκεκριμένων επιστημονικών προσεγγίσεων. Εντελώς σχηματικά, η συνολικότερη "εικόνα" που υποστηρίζουμε είναι ότι τα νοητικά φαινόμενα δεν πρέπει να κατανοηθούν ως εσωτερικές καταστάσεις του εγκεφάλου, αλλά ως τρόποι φανέρωσης του κόσμου στον άνθρωπο. Ο νους δεν είναι ένα απομονώσιμο όργανο αλλά ένα σύνολο ικανοτήτων του ανθρώπου². Η ΑΥΘΝ στην προσπάθειά της να αποφύγει τον δυϊσμό νου/σώματος σπεύδει να φυσικοποιήσει τα νοητικά φαινόμενα χωρίς προηγουμένως να έχει προσδιορίσει τι ακριβώς είναι αυτό που θέλει να αναγάγει στο φυσικό. Αποτέλεσμα αυτής της σπουδής είναι να υιοθετεί μια καρτεσιανή εικόνα των νοητικών φαινομένων η οποία υπολείπεται κατά πολύ από τον τρόπο που δίδονται πρωταρχικά στο φαινομενολογικό επίπεδο. Ο καρτεσιανισμός της ΑΥΘΝ διαφαίνεται στην διατήρηση ενός χωρισμού μεταξύ νου και κόσμου. Δεδομένου βέβαια του φυσικοκρατικού προσανατολισμού της ΑΥΘΝ, δεν πρόκειται πλέον για οντολογικό χωρισμό σε *res cogitans* και *res extensa*, αλλά για *τοπολογικό* χωρισμό: ο νους βρίσκεται *μέσα* στο υποκείμενο και πιο συγκεκριμένα ταυτίζεται ως έκτυπο (*token-identified*) με τον εγκέφαλο. Οι νοητικές αναπαραστάσεις που προϋποθέτει η ΑΥΘΝ αποτελούν ουσιαστικά αναδιπλασιασμό του κόσμου εντός του υποκειμένου, μια συμβολική περιγραφή του κόσμου μέσα στο κεφάλι. Γι' αυτό η ΑΥΘΝ, σε τελική ανάλυση, αποτελεί έναν φυσικοποιημένο καρτεσιανισμό: αντί να υπερβεί το χάσμα μεταξύ νου και κόσμου, απλώς επιχειρεί να το φυσικοποιήσει.

Ο κύριος όγκος της διατριβής είναι αφιερωμένος στον πρώτο στόχο, δηλαδή στην κριτική της παραδοχής περί ύπαρξης νοητικών αναπαραστάσεων, γιατί η συγκεκριμένη παραδοχή είναι πολύ βαθιά

² Βλ. McDowell 1992 και Putnam 1994, 1994a.

εγκαθιδρυμένη και έχει αποκτήσει ένα χαρακτήρα αυτόνοτου για την πλειονότητα των ερευνητών. Ας δούμε, όμως, λίγο πιο συγκεκριμένα τι είναι οι νοητικές αναπαραστάσεις για τις οποίες κάνει λόγο η ΑΥΘΝ. Η έννοια των νοητικών αναπαραστάσεων δεν χρησιμοποιείται μόνο μέσα στο πλαίσιο της σύγχρονης φιλοσοφίας του νου, αλλά απαντά τόσο στον καθημερινό λόγο όσο και στο λόγο της υπόλοιπης φιλοσοφίας. Αυτό αποτελεί μια πηγή σύγχυσης για όποιον προσεγγίζει για πρώτη φορά την ΑΥΘΝ γιατί η συγκεκριμένη θεωρία χρησιμοποιεί με έναν ιδιαίτερο τρόπο την έννοια της νοητικής αναπαράστασης.

Προκειμένου να εντοπίσουμε σε τι έγκειται αυτή η ιδιαιτερότητα, είναι σκόπιμο να αναφερθούμε εν συντομία στον τρόπο που χρησιμοποιείται η έννοια στην καθημερινή ζωή και στην υπόλοιπη φιλοσοφία εκτός της ΑΥΘΝ. Καταρχάς, η έννοια της νοητικής αναπαράστασης θεωρείται ότι συνδέεται πολύ στενά με την έννοια των νοητικών φαινομένων και, πιο συγκεκριμένα, ότι μπορεί να χρησιμοποιηθεί για την εξήγηση των νοητικών φαινομένων. Για παράδειγμα, στον καθημερινό λόγο πολλοί θα συμφωνούσαν ότι η έκφραση "θυμάμαι το λιοντάρι που είδα στο ζωολογικό κήπο" είναι ισοδύναμη με την έκφραση "φέρνω στο νου μου την εικόνα του λιονταριού που είδα στο ζωολογικό κήπο". Το ίδιο μπορεί να ειπωθεί και για την περίπτωση της φαντασίας. Η θέση ότι η μνήμη και η φαντασία εμπλέκουν κάποιου είδους νοητικές αναπαραστάσεις είναι μια θέση που, σύμφωνα τουλάχιστον με μια διαδεδομένη αντίληψη³, έχει υποστηριχτεί ήδη στην αρχαία φιλοσοφία. Μια πρώτη εκδοχή της συναντάμε στο *Θεαίτητο* του Πλάτωνα, ο οποίος προτείνει ότι η μνήμη θα μπορούσε να είναι κάτι ανάλογο με μια πλάκα κεριού πάνω στην οποία αποτυπώνονται οι αισθήσεις και οι σκέψεις. Οτιδήποτε αποτυπώνεται το διατηρούμε στη μνήμη και το γνωρίζουμε για όσο χρόνο υφίσταται εκεί το αποτύπωμα: "Και ὁ μὲν ἂν ἐκμαγῆ, μνημονεύειν τε καὶ ἐπίστασθαι, ἕως ἂν ἐνῆ τὸ εἶδωλον αὐτοῦ" (Πλάτων 1975a, 191 c-d). Ο Αριστοτέλης περιγράφει αυτά τα αποτυπώματα της μνήμης - τα νοητικά είδωλα (mental images) με σημερινή ορολογία - ως ένα είδος εικόνων (*De Memoria* 450 a-b). Στο *Περὶ Ψυχῆς* (431a 15-20) γενικεύει αυτή την ιδέα λέγοντας ότι η ψυχή ποτέ δεν νοεῖ χωρίς νοητικά είδωλα ή, με τη δική του ορολογία, φαντάσματα. Σύμφωνα με μια αντίληψη⁴, στους μεσαιωνικούς χρόνους, συναντάμε μια ανάλογη άποψη, μεταξύ άλλων, στον Θωμά τον Ακινάτη

³ Βλ., ενδεικτικά, Putnam 1988, Cummins 1989 και Tye 1991. Πρόκειται για μια αντίληψη που υποστηρίζεται όχι τόσο από τους ειδικούς στην αρχαία φιλοσοφία, όσο από φιλοσόφους της σύγχρονης φιλοσοφίας του νου και γι' αυτό κινδυνεύει ίσως περισσότερο να αποτελεί μια αναχρονιστική προβολή καρτεσιανών θέσεων στην αρχαία σκέψη.

⁴ Beakley & Ludlow 1992.

(1992)⁵. Ωστόσο, ήταν η νεώτερη φιλοσοφία εκείνη που έδωσε στην έννοια των νοητικών αναπαραστάσεων έναν εντελώς κεντρικό ρόλο στην εξήγηση των νοητικών φαινομένων και διαμόρφωσε το πρόβλημα της σχέσης νου/σώματος που απασχολεί τη σύγχρονη φιλοσοφία του νου. Μέσα στη νέα παράδοση που δημιουργείται, όροι όπως, "ιδέα", "εντύπωση" (impression), εποπτεία (αγγλικά: "intuition", γερμανικά: "Anschauung"), "έννοια" (αγγλικά: "concept", γερμανικά: "Begriff"), "νοητικό είδωλο" (mental image) ονομάζουν ειδή νοητικών αναπαραστάσεων. Ειδικότερα, στον Βρετανικό Εμπειρισμό κάθε είδους νοητικό φαινόμενο συναρτάται με την κατοχή ή τον χειρισμό ιδεών⁶. Έτσι, η επαφή του υποκειμένου με τον κόσμο κατανοείται ως μια έμμεση πρόσβαση στον κόσμο που πραγματοποιείται πάντα με τη διαμεσολάβηση νοητικών αναπαραστάσεων⁷. Αυτή η θέση συνέχισε να υποστηρίζεται και μέσα στο πλαίσιο της αναλυτικής φιλοσοφίας από αρκετούς φιλοσόφους μέχρι περίπου τα μέσα του εικοστού αιώνα⁸. Οι νοητικές αναπαραστάσεις, δηλαδή οι οντότητες που μεσολαβούν μεταξύ του υποκειμένου και του

⁵ Για μια διαφορετική ερμηνεία βλέπε O'Callaghan 1997.

⁶ Για παράδειγμα, ο Locke (1984, βιβλίο II, κεφάλαιο VIII, σελ. 8) δηλώνει: "Οτιδήποτε αντιλαμβάνεται ο νους ή σιδήποτε αποτελεί άμεσα αντικείμενο αισθητηριακής αντίληψης, σκέψης ή νόησης, το αποκαλώ ιδέα". Βλ., επίσης, Tye 1991, σελ. 5: "... αντίθετα προς τον Καρτέσιο και τον Hobbes, οι βρετανοί εμπειριστές υποστήριζαν ότι κάθε είδους σκέψη συνίσταται στον χειρισμό είτε απλών νοητικών ειδώλων που πηγάζουν από την αισθητηριακή εμπειρία είτε σύνθετων νοητικών ειδώλων που κατασκευάζονται από αυτά τα απλά νοητικά είδωλα".

⁷ Εκτός από την αντίληψη περί νοητικών αναπαραστάσεων (ιδεών) ως εσωτερικών αντικειμένων που κυριαρχούσε στον Βρετανικό Εμπειρισμό, σύμφωνα με τον McRae (1965, σελ. 175) υπάρχουν δύο ακόμα εναλλακτικές αντιλήψεις στους επιγόνους του Descartes κατά τον 17^ο αιώνα: "Οι τρεις αντιλήψεις είναι (α) ότι μια ιδέα είναι ένα αντικείμενο (Malebranche, Locke, Berkeley): (β) ότι μια ιδέα είναι ένα ενέργημα (Spinoza, Arnauld): (γ) ότι μια ιδέα είναι μια προδιάθεση (Leibniz)". Επίσης, για την έννοια της αναπαράστασης στον Descartes και τον Leibniz βλέπε Απαρολιτανος 1998 (πρώτο κεφάλαιο). Το πώς ακριβώς κατανοούνται οι νοητικές αναπαραστάσεις στο σύνολο της νεώτερης φιλοσοφίας αποτελεί ένα ανοικτό ερευνητικό ζήτημα το οποίο υπερβαίνει το πλαίσιο της διατριβής.

⁸ Βλ., ενδεικτικά, Ράσσελ 1963, σελ. 11: "Έτσι, καθίσταται φανερό ότι το πραγματικό τραπέζι, αν υπάρχει κάτι τέτοιο, δεν είναι το ίδιο με αυτό του οποίου έχουμε άμεσα την εμπειρία μέσω της όρασης ή της αφής ή της ακοής. Το πραγματικό τραπέζι, αν υπάρχει τέτοιο, δεν μας γίνεται με κανέναν τρόπο άμεσα γνωστό, αλλά πρέπει να το συναγάγουμε από ό,τι μας είναι άμεσα γνωστό ... Ας δώσουμε το όνομα 'αισθητηριακά δεδομένα' στα πράγματα εκείνα που μας είναι άμεσα γνωστά στις αισθήσεις: στα χρώματα, τους ήχους, τις οσμές, τη σκληρότητα, την τραχύτητα κ.ο.κ.". Βλ., επίσης, Moore 1913, Ayer 1956 κ.α.

κόσμου, αποκτούν εδώ ένα νέο όνομα: "αισθητηριακά δεδομένα" (sense data)⁹.

Αυτή είναι εντελώς σχηματικά η αντίληψη περί νοητικών αναπαραστάσεων στην παραδοσιακή φιλοσοφία. Στο πλαίσιο της ΑΥΘΝ, η έννοια των νοητικών αναπαραστάσεων διαφοροποιείται σε τρία τουλάχιστον επίπεδα από την παραδοσιακή αντίληψη:

1. *Μορφικά*: Οι νοητικές αναπαραστάσεις της ΑΥΘΝ είναι ρητά γλωσσικής και όχι εικονιστικής ή άλλης μορφής.
2. *Γνωσιολογικά*: Οι νοητικές αναπαραστάσεις της ΑΥΘΝ δεν είναι προσβάσιμες από την οπτική του πρώτου προσώπου, δηλαδή δεν εμφανίζονται στη συνείδηση του εκάστοτε υποκειμένου που τις φέρει.
3. *Οντολογικά*: Η οντολογία του νου δεν διαφέρει από την οντολογία του φυσικού κόσμου. Οι νοητικές αναπαραστάσεις της ΑΥΘΝ είναι υλικές και επιδρούν αιτιακά επηρεάζοντας την ανθρώπινη συμπεριφορά.

Από τα παραπάνω, καθίσταται φανερό ότι οι νοητικές αναπαραστάσεις της ΑΥΘΝ διαφέρουν σημαντικά από τις νοητικές αναπαραστάσεις της υπόλοιπης φιλοσοφίας. Γι' αυτό το λόγο, στη διατριβή χρησιμοποιούμε τον όρο "αναπαράσταση" αποκλειστικά για να αναφερόμαστε στις πρώτες, ενώ για τις δεύτερες χρησιμοποιούμε τον όρο "παράσταση". Ο λόγος που επιλέγουμε αυτή τη σύμβαση είναι ότι ο όρος "παράσταση" έχει ήδη επικρατήσει στην ελληνική βιβλιογραφία της νεότερης φιλοσοφίας¹⁰. Να επισημάνουμε, επίσης, ότι στη διατριβή χρησιμοποιούμε τον όρο "αναπαράσταση" για να αναφερόμαστε σε κάθε είδους σημείο, όπως, για παράδειγμα είναι, σύμφωνα με την ταξινόμηση του Peirce, τα σύμβολα, οι εικόνες και οι ενδείκτες. Αυτό δεν αποτελεί προσωπική ιδιοτροπία αλλά υιοθέτηση της καθιερωμένης χρήσης του όρου "representation" στη σύγχρονη φιλοσοφία του νου¹¹.

Η διατριβή, ως ένα βαθμό, ακολουθεί την αναλυτική μέθοδο που περιγράψαμε στην αρχή της Εισαγωγής: Κάθε φορά εξετάζουμε ένα μεμονωμένο πρόβλημα γύρω από την έννοια της νοητικής αναπαράστασης θεωρώντας *προσωρινά* όλα τα υπόλοιπα ως *λυμένα*. Για κάθε ένα από αυτά τα προβλήματα μέσα στο πλαίσιο της γνωσιακής επιστήμης έχουν αναπτυχθεί διαφορετικές απαντήσεις-θεωρίες. Προκειμένου να

⁹ Βέβαια τα αισθητηριακά δεδομένα είναι κάτι πολύ διαφορετικό από μια εσωτερική εικόνα που δίδεται στο "μάτι" του νου, ωστόσο αυτό δεν εμπόδισε τους υποστηρικτές των αισθητηριακών δεδομένων να τα εκλάβουν ως νοητικές αναπαραστάσεις στο βαθμό που λειτουργούν ως *υποκατάστατα* των εξωτερικών αντικειμένων και των ιδιοτήτων τους.

¹⁰ Για παράδειγμα, ο καντιανός όρος "Vorstellung" που στα αγγλικά μεταφράζεται "representation", στα ελληνικά έχει επικρατήσει να μεταφράζεται με τον όρο "παράσταση" (βλ. Καντ 1979 –μετάφραση Α.Γιανναρά).

¹¹ Βλ., ενδεικτικά, Fodor 1981 και Haugeland 1991.

εξετάσουμε αυτές τις θεωρίες, αρχικά τις προσεγγίζουμε "από τα μέσα", δηλαδή χωρίς να επερωτούμε τα βασικά μεταφυσικά θεμέλια πάνω στα οποία στηρίζονται, και ελέγχουμε το πώς απαντούν στα ζητήματα που οι ίδιες εντοπίζουν και επιχειρούν να επιλύσουν. Αυτή η "εσωτερική" εξέταση υιοθετεί μια ευμενή (charitable) στάση έναντι των θεωριών με στόχο την καλύτερη κατανόηση υπό τους δικούς τους όρους. Ταυτόχρονα, δημιουργεί ένα στέρεο έδαφος για μια "εσωτερική" κριτική, δηλαδή μια κριτική που δεν έχει στόχο να καταρρίψει συνολικά την εκάστοτε εξεταζόμενη θεωρία, αλλά να αναδειξει ορισμένες επιμέρους αδυναμίες της και να συνεισφέρει ενδεχομένως στην βελτίωσή της. Σε ένα δεύτερο στάδιο, προβαίνουμε σε σύγκριση και σχετική αξιολόγηση των εναλλακτικών θεωριών που έχουν προταθεί για το ίδιο πρόβλημα. Αυτή η σύγκριση παρέχει μια συνολικότερη εποπτεία του φάσματος των προτεινόμενων λύσεων και των κοινών αδιεξόδων. Η εσωτερική εξέταση και κριτική των εκάστοτε θεωριών καθώς και η μεταξύ τους σύγκριση και αξιολόγηση δημιουργεί μια ισχυρότερη νομιμοποίηση οποιασδήποτε επακόλουθης "εξωτερικής" κριτικής, δηλαδή οποιασδήποτε κριτικής που δεν στέκεται στις λεπτομέρειες των θεωριών που έχουν προταθεί για το εκάστοτε πρόβλημα, αλλά επερωτά τις κοινές μεταφυσικές προϋποθέσεις στις οποίες αυτές στηρίζονται. Έτσι, θεωρούμε ότι αποφεύγεται ο αφοριστικός και σχηματικός χαρακτήρας που παρουσιάζει συνήθως μια "εξωτερική" κριτική όταν είναι αποκομμένη από την εσωτερική δυναμική των εξεταζόμενων θεωριών.

Τόσο η "εσωτερική" όσο και η "εξωτερική" κριτική που ασκούμε δεν βασίζεται στο έργο κάποιου σημαντικού φιλοσόφου του οποίου τις έννοιες και τις θεωρίες επιχειρούμε να εφαρμόσουμε στα επιμέρους προβλήματα. Μια τέτοια επιλογή, παρότι θα μας απέλλασε από μεγάλο μέρος της ευθύνης των υιοθετούμενων απόψεων μεταθέτοντάς την στο εν λόγω φιλοσοφικό έργο, θα ήταν, όχι μόνο απ' εαυτής δογματική, αλλά και πολύ περιοριστική για την εξέταση ενός πεδίου τόσο πολύμορφου και ετερογενούς όσο η σύγχρονη φιλοσοφία του νου. Αυτό δεν σημαίνει, φυσικά, ότι δεν βασιζόμαστε καθόλου στο έργο σημαντικών φιλοσόφων του παρελθόντος και του παρόντος - εξάλλου, κάτι τέτοιο είναι μάλλον απαραίτητο για όποιον στοχεύει να κάνει φιλοσοφία. Έτσι, όταν προβαίνουμε στην "εξωτερική" κριτική, η οποία στοχεύει στην ανάδειξη του μη αναπαραστασιακού χαρακτήρα των νοητικών φαινομένων, ανατρέχουμε κυρίως στο έργο της φαινομενολογίας και του ύστερου Wittgenstein. Ωστόσο, η χρήση που κάνουμε είναι εσκεμμένα αποσπασματική με στόχο να λειτουργήσει ως πηγή έμπνευσης και γενικότερου προσανατολισμού παρά ως εφαρμογή κάποιας θεωρίας. Γι' αυτό και όλες σχεδόν οι συναφείς παραπομπές έχουν τοποθετηθεί στις υποσημειώσεις και όχι στο κυρίως κείμενο.

Η διατριβή αποτελείται από οκτώ κεφάλαια, τον Επίλογο και δύο παραρτήματα. Στο πρώτο κεφάλαιο γίνεται μια αναδρομή στις κυριότερες ιστορικές καταβολές της γνωσιακής επιστήμης. Εξετάζονται βασικές

εξελίξεις και εννοιολογικές διερευνήσεις που οδήγησαν στην ανάδυση της γνωσιακής ψυχολογίας, της τεχνητής νοημοσύνης και του λειτουργισμού. Όπως θα δούμε, η γνωσιακή ψυχολογία αναπτύχθηκε ως αντίδραση στον εξοβελισμό των εσωτερικών διαδικασιών από τον μεθοδολογικό συμπεριφορισμό. Ωστόσο, η επιστροφή στο τι συμβαίνει εντός του υποκειμένου δεν συνιστά επιστροφή στην υποκειμενική σφαίρα που διερευνούσε η ενδοσκοπική ψυχολογία, διότι οι νοητικές διαδικασίες τις οποίες μελετά η γνωσιακή ψυχολογία είναι, κατά κύριο λόγο, μη-συνειδητές. Η ανάπτυξη της μαθηματικής λογικής, από τη μεριά της, κατέστησε δυνατή την κατασκευή των ψηφιακών υπολογιστών και την ανάπτυξη της τεχνητής νοημοσύνης. Ένα σημαντικότερο επίτευγμα πίσω από αυτή την εξέλιξη είναι ότι, σε ορισμένα συστήματα, οι σημασιολογικές σχέσεις μπορούν να "κατοπτριστούν" (mirrored) στις συντακτικές σχέσεις. Έτσι, δίνεται η δυνατότητα να ελεγχθεί η λογική εγκυρότητα ενός επιχειρήματος (δηλαδή, η διατήρηση της αλήθειας από τις προκείμενες στο συμπέρασμα) με καθαρά συντακτικό τρόπο. Με την ανάπτυξη της τεχνητής νοημοσύνης, μια νέα προσέγγιση για την κατανόηση των νοητικών διαδικασιών αρχίζει να κυριαρχεί, σύμφωνα με την οποία η σκέψη αποτελεί μια διαδικασία επεξεργασίας πληροφοριών. Τέλος, οι εξελίξεις στην φιλοσοφία οδήγησαν σε μια νέα πρόταση για τη φύση του νου. Τα νοητικά φαινόμενα κατανοήθηκαν ως λειτουργικές καταστάσεις του εγκεφάλου. Σύμφωνα με τον λειτουργισμό, οι νοητικές καταστάσεις περιγράφονται στο λειτουργικό επίπεδο. Πρόκειται για ένα επίπεδο που διαφοροποιείται τόσο από το νευροφυσιολογικό όσο και από το φαινομενολογικό-ενδοσκοπικό επίπεδο, δηλαδή το επίπεδο στο οποίο ορίζεται η έρευνα της ενδοσκοπικής ψυχολογίας. Συνοψίζοντας, μπορούμε να πούμε ότι η σύγκλιση τεχνητής νοημοσύνης και γνωσιακής ψυχολογίας μέσα και από τις επεξεργασίες της φιλοσοφίας του νου υπήρξε καθοριστική συνιστώσα στην ανάδυση της γνωσιακής επιστήμης.

Στο δεύτερο κεφάλαιο εξετάζεται η Αναπαραστασιακή - Υπολογιστική Θεωρία του Νου (AYΘN) η οποία αποτελεί το φιλοσοφικό πλαίσιο της κλασικής γνωσιακής επιστήμης. Κεντρική θέση της AYΘN είναι ότι ο νους συνιστά ένα είδος ψηφιακού υπολογιστή: οι νοητικές διαδικασίες είναι μετασχηματισμοί ασυνειδητων νοητικών αναπαραστάσεων μέσω κανόνων ευαίσθητων στη μορφή των αναπαραστάσεων. Οι νοητικές αναπαραστάσεις είναι γλωσσικής μορφής και ανήκουν σε ένα έμφυτο σύστημα -τη γλώσσα της σκέψης. Η υπόθεση της γλώσσας της σκέψης προτείνεται ως αφετηρία για τη λύση του προβλήματος της αποβλεπτικότητας (intentionality), δηλαδή του βασικού γνωρίσματος κάθε νοητικού φαινομένου να κατευθύνεται προς κάτι εκτός αυτού - υπάρχον ή μη. Αποδεκτή λύση, εν προκειμένω, θεωρείται η φυσικοποίηση (naturalization) της αποβλεπτικότητας, δηλαδή η εξήγησή της με φυσικούς μη νοητικούς όρους. Σύμφωνα με τους υποστηρικτές της AYΘN, οι αναπαραστάσεις είναι το μόνο είδος φυσικών αντικειμένων που διαθέτουν σημασιολογικές ιδιότητες γι' αυτό είναι εύλογο να υποθέσει κανείς ότι οι νοητικές καταστάσεις βασίζονται με κάποιο τρόπο σε

εσωτερικές αναπαραστάσεις. Στο σώμα του κεφαλαίου αναπτύσσεται το πλαίσιο της ΑΥΘΝ και παρουσιάζονται τα βασικότερα επιχειρήματα που έχουν προταθεί υπέρ αυτού. Από αυτή την παρουσίαση συνάγεται ότι η έννοια της νοητικής αναπαράστασης παίζει εντελώς κεντρικό ρόλο στην ΑΥΘΝ. Κατά συνέπεια, οποιαδήποτε αμφισβήτηση αυτής της έννοιας διαταράσσει ολόκληρη την ΑΥΘΝ.

Προκειμένου κανείς να ασκήσει κριτική στην έννοια της νοητικής αναπαράστασης πρέπει να γνωρίζει προηγουμένως τι είναι αναπαράσταση. Γι' αυτό, στο τρίτο κεφάλαιο επιχειρείται ένας πρώτος προσδιορισμός των βασικών χαρακτηριστικών της αναπαράστασης εν γένει. Καταρχάς, διευκρινίζεται ότι οι νοητικές αναπαραστάσεις που θέτει η γνωσιακή επιστήμη δεν πρέπει να συγχέονται με τις παραστάσεις (Vorstellungen) για τις οποίες μιλά ο Kant ή τις ιδέες του Βρετανικού Εμπειρισμού διότι οι νοητικές αναπαραστάσεις δεν δίδονται από την οπτική του πρώτου προσώπου. Ωστόσο, οι νοητικές αναπαραστάσεις δεν είναι άμεσα προσβάσιμες ούτε από την οπτική του τρίτου προσώπου, διότι ακόμα και αν είχαμε εγκεφαλοσκόπια δεν θα γνωρίζαμε, τουλάχιστον επί του παρόντος, τι ακριβώς είναι εκείνο που πρέπει να ψάξουμε. Γι' αυτό το λόγο, το ερώτημα "τι συνιστά μια αναπαράσταση εν γένει;" προσεγγίζεται με αφετηρία τις εξωτερικές αναπαραστάσεις, οι οποίες, όπως είδαμε και στο δεύτερο κεφάλαιο, αποτελούν το βασικό πρότυπο για την κατανόηση των νοητικών αναπαραστάσεων. Μια αναπαράσταση, σε πρώτη προσέγγιση, αναλύεται σε αναπαραστασιακή μορφή και αναπαραστασιακό περιεχόμενο. Η αναπαραστασιακή μορφή είναι όλα εκείνα τα φυσικά χαρακτηριστικά του αναπαραστασιακού φορέα που χρησιμοποιούν στον προσδιορισμό του αναπαραστασιακού περιεχομένου. Κατόπιν, εξετάζουμε διαδοχικά δύο ερωτήματα: α) τι σημαίνει για κάτι να έχει αναπαραστασιακό περιεχόμενο και β) πώς κάτι τέτοιο καθίσταται δυνατό; Καταρχάς, δείχνουμε ότι ορισμένες αρκετά διαδεδομένες προσεγγίσεις που επιχειρούν να εξηγήσουν την έννοια του αναπαραστασιακού περιεχομένου με αναδρομή στις έννοιες της αντικατάστασης και της υποκατάστασης δεν ευσταθούν. Το γεγονός ότι το Α αποτελεί υποκατάστατο ή αντικατάστατη του Β δεν καθιστά το Α αναπαράσταση του Β. Για παράδειγμα, η γλυκόζη όντας υποκατάστατο της ζάχαρης δεν αποτελεί αναπαράσταση της ζάχαρης. Ακολουθώντας αλλά και επεκτείνοντας εν προκειμένω τον Frege, υποστηρίζουμε ότι το περιεχόμενο κάθε είδους αναπαράστασης (γλωσσικά σύμβολα, εικόνες ή ενδείκτες) αναλύεται σε δύο "συστατικά", ήτοι την αναφορά και το νόημα. Το νόημα αφορά τον τρόπο που ο χρήστης της αναπαράστασης σχετίζεται με την αναπαραστασιακή μορφή. Υποστηρίζουμε ότι συνθήκη δυνατότητας για να κατανοεί κανείς αναπαραστάσεις είναι να μπορεί να αποσπάται από το άμεσο περιβάλλον (στο οποίο, μεταξύ άλλων, ανήκει ο φορέας της αναπαράστασης) και να συμπεριφέρεται προς κάτι ως εάν να είναι μη σωματικά παρόν. Αυτή η απόσπαση από το άμεσο περιβάλλον "διανοίγει" τον χρήστη στο χώρο του νοήματος. Μόνο όντα που έχουν αυτή την ικανότητα μπορούν να χρησιμοποιούν κάτι ως αναπαράσταση. Η αναπαραστασιακή σχέση είναι

τριμελής σχέση μεταξύ αναπαριστώμενου, αναπαράστασης και χρήστη. Η αναφορικότητα των αναπαραστάσεων δεν μπορεί να αναχθεί ούτε στη σχέση αιτιότητας ούτε στη σχέση ομοιότητας μεταξύ αναπαράστασης και αναπαριστώμενου, όπως θα δούμε ότι επεχείρησε να κάνει η γνωσιακή επιστήμη προκειμένου να επιλύσει το πρόβλημα της προέλευσης του περιεχομένου των νοητικών αναπαραστάσεων. Δείχνουμε ότι αυτά τα είδη διμελών σχέσεων δεν αποτελούν ούτε επαρκείς ούτε αναγκαίες συνθήκες για την ύπαρξη αναφορικής σχέσης. Οι σχέσεις ομοιότητας και αιτιότητας καθορίζουν είδη προσομοιώσεων -ή, με την ορολογία που θα χρησιμοποιήσουμε, είδη εγγραφών (records)- και όχι είδη αναπαραστάσεων. Στο υπόλοιπο της διατριβής δεν προϋποθέτουμε την παραπάνω "εξωτερική" κριτική, αλλά την αναπτύσσουμε εκ νέου πολύ πιο συστηματικά και λεπτομερειακά για την περίπτωση των νοητικών αναπαραστάσεων.

Βασίζόμενοι στις διακρίσεις που εισαγάγαμε στο τρίτο κεφάλαιο, στα επόμενα τρία κεφάλαια εξετάζουμε ποια είδη νοητικών αναπαραστάσεων προτείνονται και τι προβλήματα δημιουργούνται σε σχέση με αυτά. Προκειμένου να εξετάσουμε αυτό το ζήτημα, θεωρούμε προσωρινά ως δεδομένο ότι εσωτερικές καταστάσεις του οργανισμού μπορούν με κάποιο τρόπο να αποκτούν αναπαραστασιακό περιεχόμενο, δηλαδή να λειτουργούν ως αναπαραστάσεις - μια παραδοχή που θα τεθεί υπό αμφισβήτηση στο έβδομο κεφάλαιο.

Συγκεκριμένα, στο τέταρτο κεφάλαιο μελετούμε λεπτομερέστερα τις γλωσσικές-συμβολικές αναπαραστάσεις που προϋποθέτει η ΑΥΘΝ μέσα από το πρόγραμμα αναπαράστασης της γνώσης στην κλασική Τεχνητή Νοημοσύνη. Το αδιέξοδο στο οποίο οδηγήθηκε αυτό το πρόγραμμα αναδεικνύει ότι τόσο η κοινότυπη όσο και η περιστασιακή γνώση δεν είναι δυνατόν να αποτελούν προτασιακή γνώση (δηλαδή, γνώση που αφορά αντικείμενα και ιδιότητες). Τα δύο αυτά είδη γνώσης εντάσσονται σε ό,τι, γενικότερα, ονομάζεται "πρακτική γνώση δεξιοτήτων" (know-how). Πρόκειται για μια γνώση που επειδή ακριβώς δεν αφορά αντικείμενα και ιδιότητες δεν έχει, υπ' αυτή την έννοια, αναφορική διάσταση και, κατά συνέπεια, δεν μπορεί να εκφραστεί με τη βοήθεια αναπαραστάσεων. Η θέση που υποστηρίζουμε είναι ότι αναγκαία προϋπόθεση για την επιδέξια (skillful) εφαρμογή της προτασιακής γνώσης είναι η κατοχή πρακτικής γνώσης δεξιοτήτων. Συνεπώς, ακόμα και αν δεχθούμε προσωρινά ότι οι συμβολικές αναπαραστάσεις επαρκούν για να εκφράσουν την προτασιακή γνώση, η γνώση της επιδέξιας εφαρμογής της προτασιακής γνώσης δεν μπορεί να συλληφθεί μέσω συμβολικών αναπαραστάσεων και κανόνων χειρισμού τους.

Στο πέμπτο κεφάλαιο, εξετάζουμε τα νοητικά είδωλα (mental images), μια εναλλακτική μορφή νοητικής αναπαράστασης που προτείνεται από ορισμένους γνωσιακούς επιστήμονες για την ανάλυση της φαντασίας και της αντίληψης. Η θεωρία περί νοητικών ειδώλων συγκροτείται μέσα από την αντιπαράθεσή της με την κυρίαρχη συμβολική προσέγγιση. Το κύριο μέρος του κεφαλαίου αφιερώνεται στην κριτική

τριών ειδών επιχειρημάτων υπέρ των νοητικών ειδώλων: φαινομενολογικών, συμπεριφοριστικών και νευροφυσιολογικών. Στο τέλος του κεφαλαίου τίθεται το θεμελιακό ερώτημα "τι συνιστά εικόνα;". Τόσο οι υποστηρικτές όσο και οι επικριτές των νοητικών ειδώλων συμφωνούν ότι η διαμάχη τους αφορά τη *μορφή* των αναπαραστάσεων που υπόκεινται της εμπειρίας νοητικών ειδώλων και όχι το ίδιο το περιεχόμενο της εμπειρίας. Υπ' αυτή την έννοια, δέχονται ότι η διαφορά των εικόνων από τα σύμβολα έγκειται στη διαφορά των αντίστοιχων αναπαραστασιακών μορφών και πιο συγκεκριμένα στον τρόπο που το αναπαραστασιακό περιεχόμενο κατοπτρίζεται στην αναπαραστασιακή μορφή. Έτσι, όμως αποσιωπούνται οι διαφορές που παρουσιάζει το ίδιο το περιεχόμενο των εικόνων από το περιεχόμενο συμβόλων. Επίσης, η όλη διαμάχη περιστρέφεται γύρω από την κοινότοπη αντίληψη ότι μια εικόνα μοιάζει με ό,τι απεικονίζει - μια αντίληψη που έχει δεχθεί δριμύτατη κριτική. Υποστηρίζουμε ότι οι υπέρμαχοι των νοητικών ειδώλων, προκειμένου να διευκρινίσουν τι ακριβώς εννοούν με τον όρο "νοητικό είδωλο", θα πρέπει να λάβουν υπόψη τις διάφορες προσεγγίσεις για το τι συνιστά εξωτερική εικόνα οι οποίες έχουν αναπτυχθεί ανεξάρτητα, κυρίως, στον χώρο της αισθητικής. Σε αυτό το πνεύμα, στο πρώτο παράρτημα της διατριβής παρουσιάζουμε τη θεωρία σημειογραφίας του Nelson Goodman, η οποία αποτελεί, ίσως, την πιο επεξεργασμένη αλλά και πιο δύσκολη θεωρία για το τι συνιστά εικονιστική αναπαράσταση και σε τι διαφέρει από τις γλωσσικές αναπαραστάσεις.

Στο έκτο κεφάλαιο, εξετάζουμε τις κατανεμημένες αναπαραστάσεις, δηλαδή τις αναπαραστάσεις που υποτίθεται ότι φέρουν τα τεχνητά νευρωνικά δίκτυα. Τα τελευταία χρόνια, παρά την αλματώδη πρόοδο στην κατασκευή και μελέτη όλο και πιο σύνθετων μορφών νευρωνικών δικτύων, το ερώτημα για το τι είναι μια κατανεμημένη αναπαράσταση έχει αντιμετωπιστεί πολύ επιφανειακά. Το γεγονός αυτό οφείλεται κυρίως στην πολυπλοκότητα του ζητήματος και την έλλειψη επαρκών θεωρητικών εργαλείων. Στο παρόν κεφάλαιο, αφού γίνει μια παρουσίαση του τρόπου κατασκευής, εκπαίδευσης και λειτουργίας ενός νευρωνικού δικτύου, εξετάζονται οι πέντε βασικότερες προσεγγίσεις που προτείνονται για το τι είναι κατανεμημένη αναπαράσταση. Κοινό χαρακτηριστικό και των πέντε είναι ότι προσδιορίζουν τις κατανεμημένες αναπαραστάσεις ανατρέχοντας σε κάποια ειδικού τύπου σχέση μεταξύ αναπαραστασιακής μορφής και αναπαραστασιακού περιεχομένου. Μέσω αναλυτικής επιχειρηματολογίας δείχνουμε ότι και οι πέντε προσεγγίσεις είναι ανεπαρκείς. Στη συνέχεια, προτείνουμε μια διαφορετική σχέση μεταξύ αναπαραστασιακής μορφής και αναπαραστασιακού περιεχομένου για την οποία δείχνουμε ότι συλλαμβάνει τις ιδιότητες που παρουσιάζει ένα νευρωνικό δίκτυο. Πρόκειται για τη σχέση που αποκαλούμε "ισχυρή υπέρθεση" και η οποία, σε απλουστευμένη μορφή, μπορεί να διατυπωθεί ως εξής: Κάθε επαρκώς μεγάλο τμήμα της αναπαραστασιακής μορφής αναπαριστά όλα τα περιεχόμενα στα οποία έχει εκπαιδευτεί το δίκτυο. Κατόπιν διερευνούμε ποιο συγκεκριμένο είδος μορφής και περιεχομένου

μπορεί να ικανοποιήσει τη σχέση ισχυρής υπέρθεσης για να καταλήξουμε στα ακόλουθα συμπεράσματα α) στα νευρωνικά δίκτυα η μορφή στην οποία "κατοπτρίζεται" το περιεχόμενο στα κρυμμένα επίπεδα είναι τα βάρη των συνδέσεων και όχι ο βαθμός ενεργοποίησης των κόμβων, β) το "κατοπτριζόμενο" περιεχόμενο στα κρυμμένα επίπεδα δεν είναι προτασιακό αλλά κάποιο είδος πρακτικής γνώσης δεξιότητων -άρα δεν έχει αναφορική διάσταση (όπως δείχνουμε στο τέταρτο κεφάλαιο) και ο κατοπτρισμός του δεν συνιστά αναπαράσταση και γ) οι αναπαραστάσεις στα επίπεδα εισόδου και εξόδου των νευρωνικών δικτύων δεν αποτελούν κατανοημένες αναπαραστάσεις αλλά κωδικοποιημένες συμβολικές ή εικονιστικές αναπαραστάσεις.

Στο έβδομο κεφάλαιο, εξετάζουμε το πρόβλημα της προέλευσης του νοητικού περιεχομένου, δηλαδή με ποιο τρόπο μια εσωτερική κατάσταση του οργανισμού αποκτά περιεχόμενο και καθίσταται νοητική αναπαράσταση. Οι υπάρχουσες προσεγγίσεις είναι δύο ειδών, ιντερναλιστικές (internalist) και εξτερναλιστικές (externalist). Τόσο τα ιντερναλιστικά όσο και τα εξτερναλιστικά προγράμματα φυσικοποίησης του περιεχομένου των νοητικών αναπαραστάσεων αποδεικνύονται ανεπαρκή στο να εξηγήσουν την αποβλεπτικότητα των νοητικών αναπαραστάσεων. Χωρίς μια τέτοια εξήγηση, ωστόσο, η ίδια η έννοια της νοητικής αναπαράστασης καθίσταται μετέωρη: η γνωσιακή επιστήμη βρίσκεται να υιοθετεί και να χρησιμοποιεί μια θεωρητική οντότητα για την οποία δεν μπορεί να εξηγήσει τη δυνατότητα ύπαρξής της. Καμία από τις υπάρχουσες προσεγγίσεις δεν επιτυγχάνει να δείξει πώς μια εσωτερική κατάσταση του εγκεφάλου μπορεί να αποκτήσει αναπαραστασιακό περιεχόμενο, δηλαδή να καταστεί νοητική αναπαράσταση. Μπροστά σε αυτό το αδιέξοδο, καταλήγουμε στο συμπέρασμα ότι πρέπει να εγκαταλειφθεί η υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων και να αναζητηθεί μια μη αναπαραστασιακή προσέγγιση των νοητικών φαινομένων. Δεδομένου ότι ο ιντερναλισμός και ο εξτερναλισμός εμφανίζονται να εξαντλούν το φάσμα των δυνατών λύσεων, ο τρόπος που επιλέγουμε να προχωρήσουμε είναι να αναζητήσουμε κοινές παραδοχές των δύο προσεγγίσεων που μπορούν να τεθούν υπό αμφισβήτηση. Μέσα από αυτή τη μέθοδο θα προκύψει μια περιγραφή των νοητικών φαινομένων που είναι πιο πιστή στον τρόπο που μας "δίνονται" πρωταρχικά. Ωστόσο, πριν προχωρήσουμε προς αυτήν την κατεύθυνση αναζητούμε μια διατύπωση του ιντερναλισμού και του εξτερναλισμού η οποία είναι συμβατή με την υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων και επίσης χαρακτηρίζει όλες τις εξετασθείσες θεωρίες φυσικοποίησης του νοητικού περιεχομένου.

Στο όγδοο κεφάλαιο, με αφετηρία τις διατυπώσεις του ιντερναλισμού και του εξτερναλισμού στις οποίες καταλήξαμε, εντοπίζουμε δύο κοινές παραδοχές τους και ασκούμε κριτική σε αυτές. Πρόκειται α) για την υιοθέτηση της οπτικής του τρίτου προσώπου ως αποκλειστικού τρόπου προσέγγισης των νοητικών φαινομένων και β) για την αποδοχή κάποιου είδους ταυτότητας μεταξύ νοητικών και

νευροφυσιολογικών ή λειτουργικών καταστάσεων. Όσον αφορά την πρώτη παραδοχή, καταρχάς διακρίνουμε δύο είδη οπτικών του τρίτου προσώπου: την καθημερινή και την επιστημονική οπτική του τρίτου προσώπου. Η πρώτη αποδίδει νοητικά περιεχόμενα σε ολόκληρο το υποκείμενο, δηλαδή αφορά το προσωπικό επίπεδο ανάλυσης. Αντίθετα, η επιστημονική οπτική του τρίτου προσώπου, η οποία ταυτίζεται με την οπτική ενός υποθετικού εγκεφαλοσκόπου, αποδίδει νοητικά περιεχόμενα σε τμήματα του εγκεφάλου και κατά συνέπεια προϋποθέτει την ισχύ της αναπαραστασιακής θεωρίας. Στη συνέχεια συγκρίνονται η οπτική του εγκεφαλοσκόπου με την οπτική του πρώτου προσώπου και αναδεικνύονται δύο ουσιαστικά χαρακτηριστικά δύο νοητικών φαινομένων που διαφεύγουν από την πρώτη. Το πρώτο χαρακτηριστικό είναι η υποκειμενικότητα και, πιο συγκεκριμένα, το ότι κάθε νοητικό φαινόμενο ανήκει αναγκαία σε κάποιον. Αυτό το χαρακτηριστικό το ονομάζουμε "συναίσθηση της πατρότητας" του νοητικού φαινομένου. Το δεύτερο χαρακτηριστικό των νοητικών φαινομένων είναι ότι μας παρέχουν μια άμεση πρόσβαση στον κόσμο. Η εξέταση των προαναφερθέντων χαρακτηριστικών προσθέτει δύο νέες κριτικές στην έννοια της νοητικής αναπαράστασης. Καταρχάς, δείχνουμε ότι η συναίσθηση της πατρότητας δεν είναι αναπαραστασιμη. Κατά συνέπεια, ακόμα κι αν δεχόμασταν, προσωρινά, ότι τα "συστατικά" των νοητικών φαινομένων που αφορούν τον αντικειμενικό τους "πόλο" είναι αναπαραστάσιμα, σίγουρα η έννοια της νοητικής αναπαράστασης δεν θα επαρκούσε για να συλλάβει τα νοητικά φαινόμενα στην ολότητά τους. Η δεύτερη ριζικότερη κριτική που προκύπτει είναι ότι η έννοια της νοητικής αναπαράστασης είναι, τελικά, ασυνάρτητη. Ειδικότερα, δείχνουμε ότι, με δεδομένο ότι η έννοια της αναπαράστασης προϋποθέτει πάντα κάποιον χρήστη (κεφάλαιο 7) καθώς και τη δυνατότητα άμεσης πρόσβασης του χρήστη είτε στον αναπαραστασιακό φορέα είτε στο αναπαρασιτώμενο αντικείμενο (δηλαδή χωρίς τη μεσολάβηση της αναπαράστασης), οι εσωτερικές καταστάσεις του εγκεφάλου δεν είναι δυνατόν να συνιστούν αναπαραστάσεις.

Από την εξέταση της πρώτης παραδοχής προκύπτει μια μη αναπαραστασιακή περιγραφή των νοητικών φαινομένων. Αναλυτικότερα δείχνουμε ότι αν η οπτική του πρώτου προσώπου κατανοηθεί καρτεσιανά ως ένα είδος ενδοσκόπησης, τότε τα δύο προαναφερθέντα χαρακτηριστικά διαφεύγουν και από αυτή την οπτική. Η συναίσθηση της πατρότητας δεν "βρίσκεται" κάπου στο εσωτερικό της υποκειμενικής σφαιρας για να εντοπιστεί με κάποιου είδους εσωτερική αντίληψη, γι' αυτό δεν πρέπει να περιγραφεί καρτεσιανά, ως μια συναίσθηση της εσωτερικότητας, αλλά ως μια συναίσθηση της ικανότητας του πράττειν (*sense of agency*). Αυτή η θέση ενισχύεται από το γεγονός ότι στην περίπτωση ορισμένων παραισθήσεων ένα υποκείμενο μπορεί να αισθάνεται ότι έχει σκέψεις "μέσα" στο νού του χωρίς ταυτόχρονα να τις αντιλαμβάνεται ως δικές του. Ο υποκειμενικός και ο αντικειμενικός "πόλος" των νοητικών φαινομένων αποτελούν μια πρωταρχική ενότητα: Η ίδια η επιτέλεση ενός νοητικού ενεργήματος επιφέρει και τη συναίσθηση της πατρότητάς του. Το

ενέργημα και αυτός που το επιτελεί δεν είναι δύο διακριτά φαινόμενα που συνδέονται εξωτερικά, αλλά αποτελούν ένα ενιαίο φαινόμενο – μια οργανική ενότητα. Επίσης, δείχνουμε ότι η συναίσθηση της ικανότητας του πράττειν δεν είναι απλό επιφαινόμενο γιατί η απουσία της από ένα νοητικό φαινόμενο (όπως, για παράδειγμα, συμβαίνει στην περίπτωση ορισμένων παραισθήσεων) προκαλεί ριζικές αλλαγές στο νοητικό βίο και την ανθρώπινη συμπεριφορά. Εκτός από την αμεσότητα με την οποία έχουμε επίγνωση ότι ο κόσμος εμφανίζεται σε εμάς, τα νοητικά φαινόμενα χαρακτηρίζονται και από μια δεύτερη αμεσότητα - διαφορετικού τύπου - που αφορά τον αντικειμενικό τους "πόλο": με τα νοητικά φαινόμενα μάς εμφανίζεται άμεσα ο κόσμος, χωρίς διάμεσους. Δείχνουμε ότι στο φαινομενολογικό επίπεδο δεν εμφανίζονται νοητικές αναπαραστάσεις ούτε εσωτερικά αντικείμενα, αλλά ο ίδιος ο κόσμος. Κατά συνέπεια, εφόσον στο φαινομενολογικό επίπεδο ούτε ο αντικειμενικός ούτε ο υποκειμενικός "πόλος" των νοητικών φαινομένων δίδονται υπό μορφή αναπαραστάσεων ή εσωτερικών αντικειμένων, η πρόσβαση από την οπτική του πρώτου προσώπου δεν πρέπει να κατανοηθεί καρτεσιανά ως κάποιου είδους ενδοσκοπήση: τα νοητικά φαινόμενα δεν βρίσκονται κάπου μέσα στο υποκείμενο αλλά αποτελούν τους τρόπους εμφάνισης του κόσμου στον άνθρωπο. Τέλος, δείχνουμε ότι σημαντικό ρόλο στη διαφοροποίηση μεταξύ ειδών νοητικών φαινομένων παίζει η αντίστοιχη τροποποίηση της συναίσθησης της ικανότητας του πράττειν: κάθε είδος νοητικού φαινομένου (αντίληψη, ανάμνηση, φαντασία κ.λπ.) χαρακτηρίζεται από ένα ιδιαίτερο σύνολο προδιαθέσεων για συμπεριφορά και δεσμεύσεων εκ μέρους του υποκειμένου. Γι' αυτό άλλωστε, ενώ κατά την παραίσθηση και τη φαντασία μπορεί να έχουμε πρόσβαση στο ίδιο περιεχόμενο, αυτό συνοδεύεται στις δύο περιπτώσεις από ένα πολύ διαφορετικό σύνολο δεσμεύσεων και προδιαθέσεων για συμπεριφορά: όταν φανταζόμαστε ένα λιοντάρι, σε αντίθεση με την περίπτωση που έχουμε την παραίσθηση ενός λιονταριού, δεν τείνουμε να το βάλουμε στα πόδια.

Τα νοητικά φαινόμενα, εφόσον ουσιασώς ανήκουν πάντα σε κάποιον, δεν μπορούν να αναλυθούν σε υποπροσωπικό επίπεδο χωρίς να χάσουν την ταυτότητά τους. Αν παραμείνουμε εγκλωβισμένοι σε μια έννοια ταυτότητας (τύπου (type) ή εκτύπου (token)) μεταξύ νοητικών και εγκεφαλικών φαινομένων, τότε αναγκαζόμαστε να υποστασιοποιήσουμε τα νοητικά φαινόμενα και, εξαλείφοντας τον υποκειμενικό τους "πόλο", μιλούμε απλά για νοητικές καταστάσεις, όπως πεποιθήσεις, επιθυμίες κ.λπ. Έτσι, δημιουργείται το "ανόητο" ερώτημα "που βρίσκονται οι νοητικές καταστάσεις;" το οποίο καθοδηγεί την έρευνα στο υποπροσωπικό επίπεδο. Υποστηρίζουμε ότι η έρευνα της γνωσιακής επιστήμης δεν αφορά τα ίδια τα νοητικά φαινόμενα αλλά ορισμένες αναγκαίες προϋποθέσεις για την εμφάνισή τους. Η κριτική της δεύτερης κοινής παραδοχής του ιντερναλισμού και του εξτερναλισμού, δηλαδή της αποδοχής κάποιου είδους ταυτότητας μεταξύ νοητικών καταστάσεων και νευροφυσιολογικών ή λειτουργικών καταστάσεων, μας οδηγεί στην υιοθέτηση μιας

χαλαρότερης σχέσης μεταξύ φυσικών και νοητικών φαινομένων, τη συνολική επιγένεση (global supervenience).

Στον Επίλογο υποστηρίζουμε, με βάση τα προηγούμενα κεφάλαια, ότι η γνωσιακή επιστήμη, στο βαθμό που διεξάγει την έρευνά της στο υποπροσωπικό επίπεδο, είναι καθ' όλα χρήσιμη και πληροφοριακή υπό τον όρο ότι οι επονομαζόμενες "νοητικές αναπαραστάσεις" δεν κατανοούνται ως αναπαραστάσεις αλλά ως εγγραφές. Εδώ δεν πρόκειται απλά για μια ορολογική αλλαγή, αλλά για την παραδοχή ότι η γνωσιακή επιστήμη δεν μελετά αυτά καθ' εαυτά τα νοητικά φαινόμενα αλλά ορισμένες αναγκαίες συνθήκες για την εμφάνισή τους. Οι εγγραφές δεν έχουν αναφορικότητα και δεν μπορούν να εξηγήσουν την μη-θεματική συναίσθηση που έχει κανείς από την οπτική του πρώτου προσώπου ότι ο κόσμος εμφανίζεται άμεσα σε αυτόν τον ιδιον. Οι εγγραφές είναι απλώς προσομοιώσεις (ή "κατοπτρισμοί") συγκεκριμένων ειδών νοητικού περιεχομένου. Υπ' αυτό το πρίσμα, τα κεφάλαια 4 - 6 πρέπει να ξαναδιαβαστούν ως μια διερεύνηση διαφορετικών ειδών εγγραφών. Από την άλλη μεριά, η έρευνα της γνωσιακής επιστήμης που διεξάγεται στο προσωπικό επίπεδο δεν πρέπει να ανατρέχει στην έννοια της νοητικής αναπαράστασης του φαινομενολογικού επιπέδου γιατί όπως δείξαμε στο όγδοο κεφάλαιο δεν εμφανίζονται νοητικές αναπαραστάσεις σε αυτό το επίπεδο. Στο χώρο της γνωσιακής επιστήμης υπάρχουν, ωστόσο, και επιστημονικές προσεγγίσεις που δεν στηρίζονται στην έννοια της νοητικής αναπαράστασης. Οι κυριότερες από αυτές βασίζονται στα νευρωνικά δίκτυα και τη θεωρία δυναμικών συστημάτων. Όσον αφορά τα νευρωνικά δίκτυα, ήδη στο έκτο κεφάλαιο δείξαμε ότι δεν φέρουν αναπαραστάσεις στα κρυμμένα επίπεδα, όχι επειδή οι νοητικές αναπαραστάσεις που προϋποθέτει η γνωσιακή επιστήμη δεν μπορούν παρά να είναι απλώς εγγραφές, αλλά διότι το περιεχόμενο που κατοπτρίζεται στα κρυμμένα επίπεδα των νευρωνικών δικτύων δεν είναι προτασιακό αλλά κάποιο είδος πρακτικής γνώσης δεξιοτήτων. Δεδομένου δε ότι η πρακτική γνώση δεξιοτήτων, όπως δείξαμε στο τέταρτο κεφάλαιο, δεν έχει αναφορική λειτουργία, η μορφή στην οποία κατοπτρίζεται το συγκεκριμένο περιεχόμενο δεν συνιστά αναπαράσταση για κάποιον επαρκή χρήστη. Ωστόσο, το ερώτημα που δημιουργείται είναι το εξής: επιτυγχάνουν τα νευρωνικά δίκτυα κάτι παραπάνω από την προσομοίωση της πρακτικής γνώσης δεξιοτήτων; Η απάντηση που δίνουμε είναι αρνητική διότι η πρακτική γνώση δεξιοτήτων ανήκει πάντα σε κάποιον, δηλαδή χαρακτηρίζεται ουσιαστικά από μια συναίσθηση της ικανότητας του πράττειν. Χάρη σε αυτή τη συναίσθηση, όπως δείξαμε στο τέταρτο κεφάλαιο, ο κάτοχος της πρακτικής γνώσης δεξιοτήτων είναι σε θέση να διακρίνει αν η εκάστοτε δραστηριότητά του είναι επιδέξια ή όχι, δηλαδή είναι σε θέση να κάνει κανονιστικές διακρίσεις. Στα νευρωνικά δίκτυα δεν υπάρχει αυτή η διάσταση και γι' αυτό αποτελούν απλώς προσομοιωτές της πρακτικής γνώσης δεξιοτήτων. Η προσέγγιση των γνωσιακών διαδικασιών μέσω δυναμικών συστημάτων αποτελεί από πολλές πλευρές μια γενίκευση της θεωρίας νευρωνικών δικτύων. Στο δεύτερο παράρτημα κάνουμε μια σύντομη παρουσίαση των

βασικών αρχών της συγκεκριμένης προσέγγισης, ενώ στον Επίλογο δείχνουμε ότι ισχύει και για αυτήν ό,τι υποστηρίξαμε για τα νευρωνικά δίκτυα: Τα δυναμικά συστήματα προσομοιώνουν κάποιο είδος πρακτικής γνώσης δεξιοτήτων που αφορά κυρίως την προσαρμοστικότητα του οργανισμού στο φυσικό περιβάλλον. Το γεγονός ότι οι δυναμικιστές δεν εντοπίζουν το νου μέσα στο κεφάλι αλλά στην φυσική αλληλεπίδραση δύο δυναμικών συστημάτων όπως το ανθρώπινο σώμα και ο κόσμος, δεν αρκεί για να τους μεταφέρει στο προσωπικό επίπεδο ανάλυσης που προσιδιάζει στη μελέτη των νοητικών φαινομένων. Ο Επίλογος τελειώνει με τη γενικότερη παρατήρηση ότι το προσωπικό επίπεδο ανάλυσης δεν συλλαμβάνει όλα τα συγκροτητικά χαρακτηριστικά των νοητικών φαινομένων αν περιοριστεί στην καθημερινή οπτική του τρίτου προσώπου και την βιτγκεσταϊνική έννοια της χρήσης. Όπως προκύπτει από τη διατριβή η συγκεκριμένη ανάλυση απαιτείται να συμπληρωθεί από μια φαινομενολογική περιγραφή των νοητικών φαινομένων που είναι αποκαθαρμένη από καρτεσιανές παραδοχές. Για μια επαρκή περιγραφή των νοητικών φαινομένων απαιτείται εκτός από την έννοια της χρήσης και η έννοια της φανέρωσης ή εμφάνισης του κόσμου. Το ίδιο ισχύει και για την ανάλυση των εξωτερικών αναπαραστάσεων: εκείνο που τις διακρίνει από όλα τα άλλα εργαλεία είναι ότι καθιστούν κάτι άλλο φανερό. Πολλά προβλήματα παραμένουν ανοικτά όσον αφορά την καθημερινή οπτική του τρίτου προσώπου και την μη-καρτεσιανή οπτική του πρώτου προσώπου καθώς και την μεταξύ τους σχέση. Το κυριότερο από αυτά είναι, ίσως, να αναλυθεί περαιτέρω η έννοια της συναίσθησης της ικανότητας του πράττειν.

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ

Ιστορικές καταβολές της γνωσιακής επιστήμης

1.1 Εισαγωγή

Η γνωσιακή επιστήμη είναι ένας νέος κλάδος που μελετά το νου κυρίως από την άποψη των γνωσιακών του λειτουργιών. Πρόκειται για μια διεπιστημονική προσέγγιση που αναδύθηκε σταδιακά από την αλληλεπίδραση της γνωσιακής ψυχολογίας με την τεχνητή νοημοσύνη, τη γλωσσολογία, τις νευροεπιστήμες και τη φιλοσοφία του νου. Σε αυτό το πρώτο κεφάλαιο θα παρακολουθήσουμε εν συντομία τις κυριότερες ιστορικές καταβολές της γνωσιακής επιστήμης. Κατά σειρά, θα εξετάσουμε τις εξελίξεις στη ψυχολογία, τη μαθηματική λογική και τη φιλοσοφία που συνέβαλαν στην ανάπτυξη της γνωσιακής επιστήμης.

1.2 Εξελίξεις στην ψυχολογία που οδήγησαν στην ανάπτυξη της γνωσιακής ψυχολογίας

1.2.1 Η κριτική της ενδοσκοπικής ψυχολογίας και η ανάπτυξη του μεθοδολογικού συμπεριφορισμού

Πριν από τη γνωσιακή ψυχολογία το κυρίαρχο ρεύμα στην ψυχολογία ήταν ο συμπεριφορισμός. Ο συμπεριφορισμός ιστορικά προέκυψε ως αντίδραση στην ενδοσκοπική ψυχολογία¹². Η χρήση της ενδοσκόπησης στην ψυχολογική έρευνα οδήγησε σε ανυπέρβλητα μεθοδολογικά προβλήματα. Η επίκληση ιδιωτικών οντοτήτων για τη στήριξη των ψυχολογικών θεωριών κλόνιζε την αντικειμενικότητα της ψυχολογίας. Επίσης, η υποψία ότι αυτές οι οντότητες είναι μη φυσικές επανέφερε το καρτεσιανό πρόβλημα για το πώς αλληλεπιδρά ο νους με το σώμα, για το πώς, δηλαδή, κάτι μη φυσικό προκαλεί ένα φυσικό φαινόμενο χωρίς ταυτόχρονα να παραβιάζονται οι νόμοι της φύσης για την διατήρηση της μάζας και της ενέργειας¹³.

Η αντίδραση στην ενδοσκοπική ψυχολογία άρχισε σιγά-σιγά να εκδηλώνεται ήδη από τα τέλη του προηγούμενου αιώνα. Πολλοί ερευνητές, ο ένας μετά τον άλλο, εγκατέλειπαν την μέθοδο της ενδοσκόπησης και στρέφονταν προς την παρατήρηση της συμπεριφοράς. Σε αυτό συνέβαλλε και η ανάπτυξη των διαφόρων οργάνων για την μέτρηση των χρόνων αντίδρασης και άλλων παραμέτρων που αφορούν την συμπεριφορά. Έτσι, η νέα μεθοδολογία που σταδιακά αναπτύσσεται βασίζεται αποκλειστικά στις παρατηρήσεις και τις μετρήσεις της συμπεριφοράς. Οι πειραματιστές μελετούν πλέον την ακρίβεια και την

¹² Βασικοί εκπρόσωποι της ενδοσκοπικής ψυχολογίας, με τη μια ή την άλλη μορφή της, είναι οι Wundt και Titchener.

¹³ Για μια παρουσίαση και κριτική αποτίμηση της μεθοδολογίας της ενδοσκοπικής ψυχολογίας βλέπε Lyons 1986.

ταχύτητα εκτέλεσης μιας πράξης μεταβάλλοντας κάθε φορά με προσεκτικό τρόπο τις συνθήκες κάτω από τις οποίες αυτή εκτελείται. Κάθε επίκληση ενδοσκοπικών δεδομένων έχει πλέον εγκαταλειφθεί. Στα 1913, ο Watson, ο πατέρας του συμπεριφορισμού, γράφει: "η ψυχολογία, όπως τη βλέπει ο συμπεριφοριστής, είναι ένας καθαρά αντικειμενικός κλάδος των φυσικών επιστημών. Ο θεωρητικός της στόχος είναι η πρόβλεψη και ο έλεγχος της συμπεριφοράς. Η ενδοσκόπηση δεν αποτελεί ουσιώδες τμήμα των μεθόδων της, ούτε η επιστημονική αξία των δεδομένων της εξαρτάται από την αμεσότητα με την οποία προσφέρονται για ερμηνεία με όρους συνείδησης" (Watson 1913a, σελ. 158). Ο Watson διακήρυξε σε όλους τους τόνους ότι η συμπεριφορά δεν έχει νοητικές αιτίες: η συμπεριφορά αποτελεί την αντίδραση του οργανισμού σε εξωτερικά ερεθίσματα.

Επίσημα, οι συμπεριφοριστές δεν έπαιρναν κάποια θέση ως προς το μεταφυσικό καθεστώς των νοητικών οντοτήτων διότι θεωρούσαν ότι τέτοια ζητήματα υπερβαίνουν το πλαίσιο της επιστήμης. Στην καλύτερη περίπτωση, πρέσβευαν ότι ακόμα και αν υπάρχουν οι νοητικές οντότητες δεν αποτελούν αντικείμενο έρευνας της ψυχολογίας γιατί δεν επηρεάζουν την συμπεριφορά. Μεθοδολογικά, η ψυχολογία θα πρέπει να βασισθεί αποκλειστικά σε δημόσια διαπιστώσιμα ερεθίσματα και συμπεριφορές. Αυτός είναι και ο λόγος που ο συμπεριφορισμός ως ρεύμα της ψυχολογίας ονομάστηκε "μεθοδολογικός συμπεριφορισμός". Στην ενότητα 1.4.1 θα αναφερθούμε σ' ένα άλλο ρεύμα που αναπτύχθηκε στο πλαίσιο της φιλοσοφίας και ονομάστηκε "αναλυτικός συμπεριφορισμός" ή "λογικός συμπεριφορισμός".

Ο μεθοδολογικός συμπεριφορισμός αναπτύχθηκε κυρίως στην Βόρειο Αμερική και έφτασε στην ακμή του κατά τις δεκαετίες του 1930 και του 1940. Ο μεθοδολογικός συμπεριφορισμός δομείται κυρίως γύρω από τις ακόλουθες θέσεις: (α) το αντικείμενο της ψυχολογίας είναι ο έλεγχος και η πρόβλεψη της ανθρώπινης συμπεριφοράς, (β) τα μόνα δεδομένα που πρέπει να λαμβάνονται υπόψη είναι σχηματισμοί (patterns) φυσικής αντίδρασης σε φυσικά ερεθίσματα (βασικός εξηγητικός όρος καθίσταται η έννοια του αντανακλαστικού) και (γ) αναφορές σε εσωτερικές καταστάσεις ή διαδικασίες του οργανισμού -φυσιολογικές ή νοητικές- δεν έχουν εξηγητική αξία.

Κυριότερος, αλλά και πιο ακραίος, υποστηρικτής του μεθοδολογικού συμπεριφορισμού ήταν ο Skinner. Ο Skinner έθεσε ως στόχο να εξηγήσει κάθε είδος συμπεριφοράς με όρους ερεθίσματος και απόκρισης (stimulus - response). Σύμφωνα με αυτή τη θεώρηση, ακόμα και οι σύνθετες συμπεριφορές δεν είναι τίποτε άλλο παρά σύνθετες αποκρίσεις που προκαλούνται από σύνθετα ερεθίσματα. Ο ίδιος ο οργανισμός, από την άποψη του συμπεριφορισμού, αντιμετωπιζόταν ως ένα "μαύρο κουτί" τού οποίου η εσωτερική δομή και λειτουργία δεν αφορούσε τις ψυχολογικές εξηγήσεις, αλλά συνιστούσε αντικείμενο άλλων επιστημών, όπως η φυσιολογία και η βιολογία.

Εφόσον, σύμφωνα με τον Skinner, η συμπεριφορά δεν οφείλεται σε εσωτερικές αιτίες (νευροφυσιολογικές ή νοητικές), κάθε μεταβολή της

πρέπει να προέρχεται από το εξωτερικό περιβάλλον. Έτσι, ο συμπεριφορισμός είναι εξ αρχής αντίθετος σε κάθε έννοια έμφυτης γνώσης, δεδομένου ότι αυτή βασίζεται αναγκαστικά σε δυνάμεις εντός του οργανισμού. Κάθε γνώση που αποκτά ένας οργανισμός απορρέει από την επαφή του με το εξωτερικό περιβάλλον. Δεν είναι, λοιπόν, τυχαίο το γεγονός ότι ο μεθοδολογικός συμπεριφορισμός εστίασε το ενδιαφέρον στην ανάλυση της μάθησης και όχι της ήδη κεκτημένης γνώσης.

Ο Skinner για να εξηγήσει τη διαδικασία της μάθησης επεξεργάστηκε την έννοια της συντελεστικής εξαρτημένης μάθησης (*operant conditioning*): ένα συγκεκριμένο είδος απόκρισης ενισχύεται όταν η εμφάνισή του "ανταμείβεται". Αν, δηλαδή, μια συμπεριφορά ακολουθείται από μια συνέπεια η οποία δρα ευεργετικά στον οργανισμό, τότε ενισχύεται η τάση του οργανισμού να επαναλάβει τη συμπεριφορά στο μέλλον. Για παράδειγμα, αν ένας ποντικός μέσα σε ένα κλουβί πατήσει τυχαία ένα κουμπί και λάβει τροφή, τότε θα συνεχίσει να πατά το κουμπί. Αν, πάλι, το πάτημα του κουμπιού μετά από κάποια στιγμή σταματήσει να δίνει φαγητό, τότε ο ποντικός θα σταματήσει τελικά να το πατά.

Ένα βασικό πρόβλημα που αντιμετώπισε η θεωρία του Skinner ήταν ότι, ενώ δούλεψε σχετικά ικανοποιητικά για στοιχειώδεις συμπεριφορές, δεν φαινόταν αρκετά ισχυρή ώστε να εξηγήσει τις σύνθετες ανθρώπινες συμπεριφορές. Χαρακτηριστικό παράδειγμα σύνθετης ανθρώπινης συμπεριφοράς είναι η γλωσσική συμπεριφορά. Ο Skinner προσπάθησε να δώσει μια συμπεριφοριστική εξήγηση της γλωσσικής συμπεριφοράς στο βιβλίο του *Verbal Behavior*, όπου επεσήμει κριτικά: "Έχει γενικά υποτεθεί ότι προκειμένου να εξηγηθεί η συμπεριφορά ... θα πρέπει να αποδοθεί σε συμβάντα που λαμβάνουν χώρα εντός του οργανισμού. Στο πεδίο της λεκτικής συμπεριφοράς αυτή η πρακτική κάποτε αντιπροσώπευε από το δόγμα περί εκφράσεως ιδεών ... Εάν ο ομιλητής είχε μια διαφορετική ιδέα, θα έλεγε διαφορετικές λέξεις ... Εάν τα λεγόμενά του ήταν ασυνήθιστα αυτό θα οφειλόταν στο καινοφανές και την πρωτοτυπία των ιδεών του.... Υπάρχει προφανώς κάτι ύποπτο στην ευκολία με την οποία ανακαλύπτουμε σε ένα σύνολο ιδεών ακριβώς εκείνες τις ιδιότητες που χρειάζονται για να εξηγήσουμε τη συμπεριφορά που τις εκφράζει. Είναι προφανές ότι κατασκευάζουμε τις ιδέες κατά βούληση από την προς εξήγηση συμπεριφορά. Δεν υπάρχει, φυσικά, κάποια πραγματική εξήγηση" (Skinner 1957, σελ. 5-6).

1.2.2 Η κριτική του συμπεριφορισμού και η εμφάνιση της γνωσιακής ψυχολογίας

Το εγχείρημα του Skinner να εξηγήσει με συμπεριφοριστικό τρόπο τη γλωσσική συμπεριφορά δεν στέφτηκε από επιτυχία όπως έδειξε ο Chomsky (1959/1980) με μια καταλυτική κριτική του. Συγκεκριμένα, ο Chomsky επεσήμανε ότι κάθε άνθρωπος έχει την ικανότητα να παράγει ένα αμέτρητο αριθμό καινοφανών προτάσεων. Όμως, "εφόσον η εκφώνηση μιας καινοφανούς πρότασης είναι μια συμπεριφορά που δεν

έχει ξανασυμβεί ποτέ, πώς θα μπορούσε μια προγενέστερη ενίσχυση να εξηγήσει τον τρόπο με τον οποίο οι άνθρωποι μαθαίνουν να παράγουν τέτοιες νέες προτάσεις;¹⁴ Ο Chomsky αντιπρότεινε ότι χωρίς την ύπαρξη κάποιων σύνθετων εσωτερικών δομών είναι αδύνατον να εξηγηθούν οι γλωσσικές ικανότητες.

Ένα άλλο γενικότερο πρόβλημα που παρουσίαζε ο μεθοδολογικός συμπεριφορισμός σχετιζόταν με τον τρόπο εξατομίκευσης των ερεθισμάτων και των αποκρίσεων. Τα πειράματα συντελεστικής εξαρτημένης μάθησης που παρέθετε ο Skinner εκτελούνταν σε ελεγχόμενα και πολύ απλά περιβάλλοντα. Ωστόσο, οι οργανισμοί, όταν είναι ελεύθεροι, έχουν την ικανότητα να μαθαίνουν μέσα σε πολύ σύνθετα και διαρκώς μεταβαλλόμενα περιβάλλοντα. Σε τέτοια περιβάλλοντα, κάθε στιγμή, τα ερεθίσματα είναι πάρα πολλά, ενώ οι προκαλούμενες αποκρίσεις χαρακτηρίζονται από μεγάλη προσαρμοστικότητα. Ένα πρόβλημα, λοιπόν, που τίθεται είναι πώς μπορούν να ενταχθούν επιμέρους ερεθίσματα και συμπεριφορές σε συγκεκριμένους τύπους ερεθίσματος και συμπεριφοράς¹⁵.

Συχνά, το ίδιο ερέθισμα μπορεί να προκαλέσει δύο διαφορετικές συμπεριφορές. Πώς μπορεί σε αυτή την περίπτωση να εξηγηθεί η παρατηρούμενη διαφορά; Ο Skinner θα ανέτρεχε εδώ στις διαφορές της ιστορίας ενίσχυσης: δύο οργανισμοί διαφέρουν στην συμπεριφορά παρότι δέχονται το ίδιο ερέθισμα γιατί έχουν διαφορετικές ιστορίες ενίσχυσης. Το πρόβλημα με αυτή την απάντηση είναι ότι οι παρελθούσες ιστορίες των δύο οργανισμών είναι αδύνατον από μόνες τους να μεταβάλλουν τη συμπεριφορά γιατί ακριβώς είναι *παρελθούσες*. Εφόσον, λοιπόν, δεν δεχόμαστε δράση εκ χρονικής αποστάσεως, ο μόνος τρόπος με τον οποίο παρελθόντα γεγονότα μπορούν να επηρεάσουν την τρέχουσα συμπεριφορά είναι αν έχουν μεταβάλλει την εσωτερική κατάσταση των οργανισμών. Έτσι, η διαφορά στις παρατηρούμενες συμπεριφορές μπορεί να αποδοθεί στις διαφορές των τρεχουσών εσωτερικών καταστάσεων. Ωστόσο, όπως ήδη επισημόναμε, ο μεθοδολογικός συμπεριφορισμός δεν αποδέχεται μια τέτοιου τύπου εξήγηση γιατί κάνει αναφορά στην εσωτερική κατάσταση του οργανισμού η οποία δεν είναι παρατηρήσιμη.

Τα ερεθίσματα, λοιπόν, δεν επαρκούν για να καθορίσουν τη συμπεριφορά και η αναδρομή στην ιστορία ενίσχυσης δεν λύνει το πρόβλημα εφόσον δεν δεχόμαστε δράση εκ χρονικής αποστάσεως. Η "ένδεια" του ερεθίσματος (poverty of stimulus), όπως την ονομάζει ο Chomsky, αναδεικνύει τη συνεισφορά του οργανισμού στη διαδικασία της αντίληψης: στην αντιληπτική απόκριση υπάρχει περισσότερη πληροφορία από την πληροφορία που περιλαμβάνεται στο ερέθισμα. Με άλλα λόγια, η αντίληψη υποκαθορίζεται από την αίσθηση. Χαρακτηριστικό παράδειγμα

¹⁴ Machamer 1998, σελ. 492.

¹⁵ Βλ. Preston 1994.

Ελληνικής και Αγγλικής Γλώσσας". Ευχαριστίες οφείλω, επίσης, στον καθηγητή νευρωνικών δικτύων του Τμήματος Ηλεκτρολόγων Μηχανικών και Ηλεκτρονικών Υπολογιστών του Ε.Μ.Π. Ανδρέα Σταφυλοπάτη καθώς και στο Γιώργο Μαγουλά για τα σχόλιά τους στο έκτο κεφάλαιο της διατριβής. Η Δήμητρα Κατή με προσκάλεσε το 1997 να μιλήσω στη διεπιστημονική διημερίδα με τίτλο "Νόηση και Γλώσσα" και στη συνέχεια συμπεριέλαβε την ανακοίνωσή μου με τίτλο "Υπάρχει Γλώσσα της σκέψης;" στο τόμο που δημοσίευσε με την Μαριάννα Κονδύλη και την Κική Νικηφορίδου. Τμήματα αυτού του άρθρου έχουν συμπεριληφθεί στη διατριβή. Οι συζητήσεις που είχα με τη Δήμητρα και τα σχόλιά της στο άρθρο μου με βοήθησαν να ξεκαθαρίσω ζητήματα που αφορούν διάφορες προσεγγίσεις στη ψυχολογία περί γνωσιακών αρχιτεκτονικών.

Ευχαριστώ το Δρακούλη Νικολινάκο ο οποίος με προσκάλεσε να συμβάλλω στο σχεδιασμό και τη διδασκαλία του σεμιναρίου "Αναπαράσταση και Προθετικότητα" που έγινε το εαρινό εξάμηνο του 1999 στο πλαίσιο των δύο προαναφερθέντων μεταπτυχιακών προγραμμάτων. Μέσα από τις συζητήσεις μας και την επαφή με την ενημερωμένη βιβλιοθήκη του ωφελήθηκα σημαντικά. Ευχαριστίες οφείλω επίσης στους κ.κ. Μιχάλη Ασημακόπουλο, Κατερίνα Ιεροδιακόνου, Βασίλη Καρασμάνη, Βαγγέλη Καφετζόπουλο, Γιώργο Κολέτσο, Κώστα Κριμπά, Παντελή Μπασάκο, Πανταζή Τσελεμάνη και Παύλο Χριστοδουλίδη οι οποίοι με τα σεμινάρια τους και τις συζητήσεις μας με βοήθησαν να διευκρινίσω ζητήματα που αφορούν τη διατριβή μου. Ο Άρης Αραγιώργης διάβασε από το πρώτο κεφάλαιο την ενότητα για τη μαθηματική λογική. Τον ευχαριστώ θερμά για τις παρατηρήσεις του. Ευχαριστίες οφείλω και στη Φαίη Ζήκα η οποία με βοήθησε σημαντικά λίγο πριν το τέλος της διατριβής μου.

Με τον Άρη Κουτούγκο κάναμε αρκετές συζητήσεις τα τελευταία χρόνια στον ΑΚΕΔ και η συμφωνία μας σε πολλά ζητήματα μου έδινε θάρρος να συνεχίσω. Θα ήθελα να ευχαριστήσω ιδιαίτερα το Γιώργο Ξηροπαίδη ο οποίος με τα συστηματικά και υψηλότατου επιπέδου σεμινάρια που έκανε από το 1993 έως το 2000 έπαιξε καθοριστικό ρόλο στη φιλοσοφική μου συγκρότηση και με επηρέασε βαθιά. Ο Γιώργος μού συμπαραστάθηκε και με στήριξε όλα αυτά τα χρόνια όχι απλά ως ένας δάσκαλος. Η Βάσω Κινή με έφερε σε επαφή με το έργο του Wittgenstein και με βοήθησε άμεσα με πολλούς τρόπους στη συγγραφή της διατριβής. Η γενικότερη στάση της απέναντί μου από το 1988 που τη γνώρισα μέχρι σήμερα είναι πολύ σημαντική για μένα.

Ευχαριστίες οφείλω στην επταμελή μου επιτροπή. Θα ξεκινήσω από τα τέσσερα μέλη που συμπληρώνουν την τριμελή συμβουλευτική επιτροπή. Ευχαριστώ θερμά το Στέλιο Βιρβιδάκη ο οποίος ήδη από το 1988 με κατέδυνε στην ανάγνωση του Kant και λίγο αργότερα με τα σεμινάρια του στο Deree με εισήγαγε στην αγγλοσαξονική γνωσιολογία. Ο Στέλιος συνέχισε σταθερά μέχρι σήμερα να με βοηθά με τις ανεξάντλητες γνώσεις και την κρίση του σε όλες τις φάσεις του διδακτορικού. Η Βούλα Τσινόρεμα διάβασε διάφορα κείμενά μου και με βοήθησε πολύ με τα σχόλιά της τα τελευταία χρόνια. Η καθηγήτρια Στέλλα Βοσνιάδου οργάνωσε το

μεταπτυχιακό πρόγραμμα Βασικής και Εφαρμοσμένης Γνωστικής Επιστήμης στο πλαίσιο του οποίου έγιναν πολλές διαλέξεις και σεμινάρια της ίδιας και άλλων που με βοήθησαν σημαντικά. Την ευχαριστώ θερμά για τις παρατηρήσεις που έκανε στη διατριβή καθώς και για την εμπιστοσύνη και την ευθύτητά της απέναντί μου. Τέλος, θα ήθελα να ευχαριστήσω τον καθηγητή Διονύσιο Αναπολιτάνο για τα διεισδυτικά σχόλιά του στη διατριβή μου. Η αδιάκοπη προσπάθειά του όλα αυτά τα χρόνια να βελτιώσει το διαπανεπιστημιακό μεταπτυχιακό πρόγραμμα Φιλοσοφίας και Ιστορίας των Επιστημών και της Τεχνολογίας σίγουρα επηρέασε και μένα ως μέλος του προγράμματος.

Η τριμελής συμβουλευτική μου επιτροπή απαρτίζεται από τον Αριστείδη Μπαλτά, επιβλέποντα της διατριβής, τον Παντελή Νικολακόπουλο και τον Κώστα Γαβρόγλου. Ο Παντελής Νικολακόπουλος, που τόσο αναπάντεχα έφυγε από κοντά μας, με βοήθησε πάρα πολύ από τα πρώτα χρόνια που εμφανίστηκε στο μεταπτυχιακό. Τα σεμινάρια του στη νεότερη φιλοσοφία ήταν πηγή σκέψης και έμπνευσης. Ειδικότερα, τα τελευταία χρόνια που άρχισα να του δίνω κείμενα, οι επαφές μας εντατικοποιήθηκαν. Η καθαρή του σκέψη και οι καίριες παρατηρήσεις του μου έδιναν πάντα πολύ μεγάλη βοήθεια και ικανοποίηση. Ο Παντελής μερμυνούσε μέχρι το τέλος για την πορεία της διατριβής μου παρά την κατάσταση της υγείας του. Η στάση του αυτή θα μείνει για πάντα στη μνήμη μου. Το δεύτερο μέλος της συμβουλευτικής μου επιτροπής είναι ο Κώστας Γαβρόγλου. Θα ήθελα να τον ευχαριστήσω για τη σταθερή υποστήριξή του και όλα όσα μου προσέφερε.

Τέλος θα ήθελα να αναφερθώ στον επιβλέποντα καθηγητή μου Αριστείδη Μπαλτά. Δεν είναι εύκολο να βρω λόγια να ευχαριστήσω τον Αριστείδη ο οποίος έπαιξε καθοριστικό ρόλο στη ζωή μου από την πρώτη στιγμή που τον γνώρισα το 1988. Χωρίς την παρουσία του σίγουρα δεν θα είχα διανοηθεί να ξεκινήσω διδακτορικό φιλοσοφίας στην Ελλάδα. Οι αμέτρητες συζητήσεις μας όλα αυτά τα χρόνια με διαμόρφωσαν και με επηρέασαν σε πολλά διαφορετικά επίπεδα που δεν αφορούν μόνο τη φιλοσοφία αλλά και τη γενικότερη στάση ζωής. Για όλα αυτά και πολλά άλλα θέλω να τον ευχαριστήσω θερμά.

Ευχαριστώ τους φίλους μου Θρασύβουλο Παπαδόπουλο, Ολυμπία Λαμπούση και Βάλια Δημοπούλου για τη στήριξη που μου παρείχαν όλα αυτά τα χρόνια. Ο Φραγκίσκος Σπανός με βοήθησε υπομονετικά σε πολλά τεχνικά ζητήματα και τον ευχαριστώ θερμά.

Η Ελευθερία Ασρινάκη με έκανε να πιστέψω ότι η διατριβή μου έχει ένα τέλος. Την ευχαριστώ για όλα.

Τέλος θα ήθελα να ευχαριστήσω ολόψυχα τους γονείς μου Κυριάκο και Μαργαρίτα Παγωνδιώτου καθώς και την αδελφή μου Ευαγγελία Παγωνδιώτου για την ακαταπόνητη στήριξη που μου παρείχαν όλα αυτά τα χρόνια, μια στήριξη που ήταν τόσο βαθιά ώστε να εμφανίζεται ως αστονόητη.

είναι ο τρόπος με τον οποίο οι άνθρωποι διακρίνουν τις παύσεις που οριοθετούν τις λέξεις σε μια φράση: "όταν ακούμε, π.χ. το όνομα 'Bob Lees' ακούμε μια διακοπή ή παύση μεταξύ του 'Bob' και του 'Lees'. Ωστόσο, αν εξετάσουμε τη γραφική παράσταση του ηχητικού κύματος που προκύπτει από μια φυσιολογική εκφώνηση αυτού του ονόματος, παρατηρούμε ότι η διακοπή στην ηχητική καμπύλη εμφανίζεται στην πραγματικότητα μεταξύ του 'BobL' και του 'ees'. Αυτό σημαίνει ότι ο ακροατής αντιλαμβάνεται μια παύση που δεν αντιστοιχεί στη φυσική διακοπή της μορφής του ηχητικού κύματος. Η συμπεριφορά που σχετίζεται με την αντίληψη της παύσης δεν διαμορφώνεται από τα φυσικά χαρακτηριστικά των ηχητικών κυμάτων."¹⁶

Μία κατεύθυνση προς την οποία χρησιμοποιήθηκε το επιχείρημα της "ένδειας" του ερεθίσματος οδήγησε στο συμπέρασμα ότι η αντίληψη δεν αποτελεί κάποιον είδους αντανakλαστικό αλλά είναι θεωρητικά φορτισμένη, δηλαδή, καθορίζεται (ως ένα βαθμό) από την θεωρία μας για τον κόσμο¹⁷. Επίσης, πολύ σημαντικό ρόλο προς την ίδια κατεύθυνση έπαιξε η θέση του Bruner για τη σημασία των αναγκών και των αξιών στη διαμόρφωση της αντίληψης. Έτσι, η αντίληψη θεωρήθηκε ότι βασίζεται, εκτός από τα αισθητηριακά ερεθίσματα, σε κάποια ήδη δεδομένη γνώση του οργανισμού και, κατά συνέπεια, ερμηνεύτηκε ως μια διαδικασία συναγωγών στην οποία συμμετέχουν ως προκειμένες πεποιθήσεις που απορρέουν απ' αυτή τη γνώση.

Η θεωρία του Bruner (1957) αποτέλεσε τη βάση για μια νέα προσέγγιση στο εσωτερικό της ψυχολογίας που ονομάστηκε "Νέα ματιά στην αντίληψη" (New Look in Perception)¹⁸. Σύμφωνα με αυτή, δεν υπάρχουν "γυμνά" δεδομένα, ούτε "αθά μάτια" –καθετί το βλέπουμε πάντα ως κάτι, δηλαδή κατηγοριοποιημένο. Ο Bruner θεωρεί ότι η αντίληψη είναι μια διαδικασία κατηγοριοποίησης που πραγματοποιείται μέσω συναγωγών, με προκειμένες που προέρχονται είτε από τις αισθήσεις είτε από τις πεποιθήσεις, τις ανάγκες και τις προσδοκίες. Το τελικό "πρόϊόν" (η έξοδος) αυτής της διαδικασίας είναι η κατηγορία του αντικειμένου που δίδεται στην αντίληψη. Σύμφωνα, λοιπόν, με τον Bruner η αντίληψη είναι κατά βάση μια διαδικασία επίλυσης προβλημάτων και ως τέτοια δεν διαφέρει ποιοτικά από τη σκέψη.

Οι παραπάνω απόψεις και οι κριτικές κατά του συμπεριφορισμού συνέβαλλαν στην ανάδυση ενός νέου δόγματος για την ψυχολογία που

¹⁶ Machamer 1998, σελ. 493. Το επιχείρημα χρησιμοποιείται από τον Fodor 1968.

¹⁷ Το εάν αυτή η θεωρία είναι επίκτητη ή εγγενής αποτελεί ένα περαιτέρω ζήτημα προς συζήτηση που δεν θα μας απασχολήσει γιατί αμφότερες οι θέσεις μπορούν να φιλοξενηθούν μέσα στο πλαίσιο μιας θεωρίας που δίνει προτεραιότητα στις εσωτερικές διαδικασίες του υποκειμένου.

¹⁸ Μια παρόμοια προσέγγιση είχε αναπτυχθεί τον 19ο αιώνα από τον μεγάλο φυσιολόγο H. von Helmholtz

ονομάστηκε "γνωσιοκρατία" (cognitivism). Η κεντρική ιδέα γύρω από την οποία αρθρώνεται το δόγμα είναι ότι ο οργανισμός συνεισφέρει δραστικά στην *οργάνωση των δεδομένων*. Οι ψυχολόγοι, προκειμένου να εξηγήσουν την ανθρώπινη συμπεριφορά, πρέπει να στραφούν στις εσωτερικές καταστάσεις του υποκειμένου. Ωστόσο, δεν πρόκειται πλέον για τις εσωτερικές καταστάσεις της ενδοσκοπικής ψυχολογίας, αλλά για φυσικές καταστάσεις του οργανισμού που εξετάζονται από την άποψη της λειτουργίας τους. Κάτω από την επίδραση της θεωρίας πληροφοριών και την ανάπτυξη των υπολογιστικών μηχανών¹⁹, ο οργανισμός από ψυχολογική άποψη αρχίζει πλέον να νοείται ως ένα σύστημα επεξεργασίας πληροφοριών. Τα κύρια ερωτήματα που απασχολούν όσους ασπάζονται τη γνωσιοκρατία, δηλαδή τους γνωσιακούς ψυχολόγους, είναι με ποιο τρόπο συλλέγονται οι πληροφορίες από το αντιληπτικό σύστημα, πώς γίνεται η επεξεργασία τους, πώς αποθηκεύονται και με ποιο τρόπο ενεργοποιούν το κινητικό σύστημα; Η γνωσιακή ψυχολογία, τελικά, θα συγκροτηθεί ως κλάδος και θα αποκτήσει το όνομά της μέσα στη δεκαετία του 1960 αντλώντας ιδέες μεταξύ άλλων από την τεχνητή νοημοσύνη και από τον λειτουργισμό όπως θα διαφανεί στις επόμενες ενότητες.

1.3 Εξελίξεις στη μαθηματική λογική που οδήγησαν στην ανάπτυξη της τεχνητής νοημοσύνης

Μια από τις σημαντικότερες επιστημονικές εξελίξεις στις αρχές του 20^{ου} αιώνα είναι η ανάπτυξη της μαθηματικής λογικής. Αυτή η εξέλιξη οδήγησε στην δημιουργία των τυποποιημένων γλωσσών, στην επινόηση των ηλεκτρονικών υπολογιστών και τελικά στην ανάπτυξη της τεχνητής νοημοσύνης.

1.3.1 Η συμβολή του Frege

Οι βάσεις για την ανάπτυξη της μαθηματικής λογικής τέθηκαν από τον Frege στο τέλος του 19^{ου} αιώνα. Ο Frege είχε θέσει ως στόχο του να αναγάγει τα μαθηματικά στη λογική. Με δεδομένο ότι η μαθηματική γνώση αποτελείται από ένα σύνολο θεωρημάτων που παράγονται με αποδεικτικές διαδικασίες από έναν αριθμό αξιωμάτων και ένα σύνολο κανόνων, ο Frege επεχείρησε να δείξει ότι οι μαθηματικές αποδείξεις μπορούν να αναχθούν στις λογικές αποδείξεις. Ένα τέτοιο εγχείρημα απαιτούσε να δειχθεί ότι οι μαθηματικές αποδείξεις βασίζονται μόνο στη λογική και όχι και σε κάποιο είδος περιεχομένου που είναι αποκλειστικά μαθηματικό και όχι λογικό. Ο Frege ερχόταν εδώ αντιμέτωπος με την καντιανή παράδοση, σύμφωνα με την οποία οι μαθηματικές προτάσεις είναι a priori συνθετικές. Η υποτιθέμενη συνθετικότητα των μαθηματικών

¹⁹ Βλ. Guttenplan 1994, σελ. 170.

προτάσεων σήμαινε ουσιαστικά ότι αυτές δεν βασίζονται μόνο στη λογική, αλλά βγαίνουν έξω από το χώρο της διάνοιας και συνδέονται με (καθαρές) εποπτείες. Ο Frege, προκειμένου να απορρίψει την παραπάνω καντιανή θέση, όφειλε να δείξει ότι οι μαθηματικές προτάσεις και αποδείξεις δεν είναι συνθετικές αλλά αναλυτικές. Πώς, όμως, θα μπορούσε να δείχθει κάτι τέτοιο, δεδομένου ότι ακόμα και για την πιο απλή μαθηματική συναγωγή δεν είναι σαφές αν το συμπέρασμα απορρέει καθαρά λογικά ή βασίζεται σε κρυφές προκειμένες που παρεισφρουν από το περιεχόμενο των εμπλεκόμενων μαθηματικών εννοιών; Για να ξεπεράσει αυτό το πρόβλημα ο Frege ανέπτυξε ένα είδος συμβολισμού με τον οποίο μπορούσε να παραστήσει ρητά τη δομή των προτάσεων και τις λογικές τους σχέσεις. Έτσι, μεταγράφοντας μια μαθηματική απόδειξη σε αυτόν τον συμβολισμό κρατούσε μόνο τη λογική της δομή και άφηνε εκτός κάθε τι που αφορούσε το σημασιολογικό της περιεχόμενο. Η μέθοδος του Frege είχε γενική εφαρμογή και δεν περιοριζόταν στον έλεγχο των μαθηματικών αποδείξεων. Οποιοδήποτε επιχείρημα στη φυσική γλώσσα θεωρήθηκε ότι μπορούσε πλέον να γραφτεί με το νέο συμβολισμό και να ελεγχθεί ως προς την ορθότητά του. Η ανάπτυξη του συγκεκριμένου συμβολισμού και ειδικότερα ο τρόπος που ο Frege ανέλυσε τη λογική δομή της γλωσσικής πρότασης (sentence)²⁰ αποτελούν το θεμέλιο λίθο της μαθηματικής λογικής²¹.

Ο Frege ουσιαστικά ανέπτυξε μια τυπική γλώσσα (formal language), δηλαδή έναν ανερχόμενο συμβολισμό, όπου μπορούσε να ελέγξει με καθαρά συντακτικά κριτήρια αν μια μαθηματική συναγωγή (ή ένα επιχείρημα) αποτελεί όντως απόδειξη ή όχι. Συντακτικά κριτήρια, εν προκειμένω, είναι ρητοί κανόνες που καθορίζουν τους επιτρεπτούς τρόπους συνδυασμού και μετασχηματισμού των συμβόλων με βάση τη μορφή των συμβόλων και ανεξάρτητα από το σημασιολογικό τους περιεχόμενο. Χαρακτηριστικό παράδειγμα ενός τέτοιου κανόνα είναι ο *modus tollens*:

²⁰ Για τη δικαιολόγηση της μεταφραστικής επιλογής βλέπε υποσημείωση 57.

²¹ Πριν από τον Frege, μια γλωσσική πρόταση θεωρούταν ότι αναλύεται λογικά στα εξής συστατικά: υποκείμενο, copula και κατηγορημα. Ο Frege μετέβαλε ριζικά αυτή την εικόνα προτείνοντας τη λογική ανάλυση της γλωσσικής πρότασης σε αντικείμενο και συνάρτηση (βλ. Coffa 1991, σελ. 63). Για παράδειγμα, η πρόταση *ο πύργος είναι κόκκινος*, σύμφωνα με την ανάλυση του Frege, υπολογίζεται από δύο συστατικά: το αντικείμενο στο οποίο αναφέρεται η φράση "ο πύργος" και τη συνάρτηση που συνοπτικά θα μπορούσαμε να εκφράσουμε ως *Κόκκινο(x)*, δηλαδή, ο *x* είναι κόκκινος. Η συνάρτηση αυτή αντιστοιχεί αντικείμενα σε αληθισμές. Η έννοια της συνάρτησης που εισάγει εδώ ο Frege πρέπει να νοηθεί ευρύτερα από την μαθηματική έννοια της συνάρτησης, υπό την έννοια ότι τα ορίσματα που δέχεται δεν περιορίζονται αποκλειστικά σε μαθηματικά αντικείμενα.

$$p \rightarrow q$$

$$\neg q$$

$$\neg p$$

Το βασικό, λοιπόν, επίτευγμα του Frege σε σχέση με το ζήτημα που μας ενδιαφέρει (δηλαδή την ανάπτυξη της μαθηματικής λογικής και την ανακάλυψη των ηλεκτρονικών υπολογιστών) ήταν ότι έδειξε ότι η έννοια της απόδειξης μπορεί να μελετηθεί καθαρά συντακτικά. Με άλλα λόγια, έδειξε ότι η αποδειξιμότητα (provability) είναι *συντακτική* και όχι σημασιολογική ιδιότητα. Κατ' αυτόν τον τρόπο, η αποδειξιμότητα διακρίθηκε από τη λογική εγκυρότητα (logical validity), ήτοι τη διατήρηση της αλήθειας κατά τη μετάβαση από τις προκειμένες στο συμπέρασμα²². Η λογική εγκυρότητα, σε αντίθεση προς την αποδειξιμότητα, εφόσον ορίζεται με βάση την έννοια της αλήθειας, αποτελεί ένα *εγγενώς σημασιολογικό* χαρακτηριστικό των επιχειρημάτων. Όπως επισημαίνει ο Rynasiewicz (1990, σελ. 19), "πριν τον Frege... όλοι προσπαθούσαν να αναλύσουν την φύση των αποδείξεων με σημασιολογικούς όρους. Η μεγάλη καινοτομία του Frege ήταν να θεωρήσει τις αποδείξεις αποκλειστικά ως χειρισμό συμβόλων σύμφωνα με συντακτικούς κανόνες. Κατ' αυτόν τον τρόπο, η ορθότητα των αποδείξεων καθίσταται αποτελεσματικά αποκρίσιμη (effectively decidable) ανεξάρτητα από ζητήματα που αφορούν τη λογική εγκυρότητα. Έτσι, παρότι η λογική εγκυρότητα παραμένει ένα σημασιολογικό χαρακτηριστικό των επιχειρημάτων, η ...αποδειξιμότητα του συμπεράσματος καθίσταται ένα αποκλειστικά συντακτικό χαρακτηριστικό".

Αυτός ο διαχωρισμός της αποδειξιμότητας από τη λογική εγκυρότητα διάνοιξε το δρόμο για την ανάπτυξη δύο ανεξάρτητων θεωριών: της θεωρίας αποδείξεων (proof theory) και της τυπικής σημασιολογίας (formal semantics). Η πρώτη αφορά το συντακτικό επίπεδο, ενώ η δεύτερη το σημασιολογικό. Το νέο βασικό πρόβλημα που δημιουργεί η αποσύνδεση της αποδειξιμότητας από τη λογική εγκυρότητα είναι πλέον ο τρόπος με τον οποίο συνδέονται αυτές οι δύο ιδιότητες. Αν, για παράδειγμα, η αποδειξιμότητα δεν συνεπάγεται τη λογική εγκυρότητα, τότε η μελέτη του συντακτικού επιπέδου μιας γλώσσας δεν επαρκεί για να εγγυηθεί τη διατήρηση ή μη της αλήθειας σε ένα επιχείρημα. Προτού εξετάσουμε αυτό το νέο πρόβλημα θα αναφερθούμε εντελώς συνοπτικά

²² Αυστηρότερα η έννοια της λογικής εγκυρότητας ορίζεται ως εξής: Έστω φ μια γλωσσική πρόταση και Σ ένα σύνολο γλωσσικών προτάσεων. Ένα επιχείρημα είναι λογικά έγκυρο αν το συμπέρασμα φ είναι λογική συνέπεια των προκειμένων Σ , όπου το φ είναι μια λογική συνέπεια του Σ αν και μόνο αν είναι αδύνατον για το φ να είναι ψευδές όταν όλα τα μέλη του Σ είναι αληθή (βλ. Rynasiewicz 1990, σελ. 6).

στον τρόπο με τον οποίο αρθρώνονται η θεωρία αποδείξεων και η τυπική σημασιολογία.

1.3.2 Η Θεωρία Αποδείξεων

Στη θεωρία αποδείξεων αναπτύσσονται και μελετώνται διάφορα αποδεικτικά συστήματα με καθαρά συντακτικό τρόπο. Ένα αποδεικτικό σύστημα συνίσταται από ένα σύνολο (ανερμήνευτων) αξιωμάτων καθώς και ένα σύνολο κανόνων συναγωγής. Έχουν αναπτυχθεί πολλά διαφορετικά αποδεικτικά συστήματα. Τα πιο γνωστά, και αυτά που προσεγγίζουν περισσότερο τις καθημερινές μας αποδεικτικές πρακτικές, είναι τα συστήματα Φυσικής Παραγωγής (Natural Deduction). Τα συστήματα Φυσικής Παραγωγής δεν διαθέτουν αξιώματα παρά μόνο κανόνες συναγωγής. Για παράδειγμα, στην προτασιακή λογική τέτοιοι κανόνες είναι οι ακόλουθοι:

1. **Διπλή Άρνηση:** Για κάθε πρόταση ϕ , μπορούμε να συναγάγουμε την ϕ από την $\neg\neg\phi$ καθώς και την $\neg\neg\phi$ από την ϕ .
2. **Επανάληψη:** Για κάθε πρόταση ϕ , μπορούμε από το ϕ να συναγάγουμε το ϕ .
3. **Modus Ponens:** Για κάθε πρόταση ϕ και ψ , μπορούμε από το ϕ και $(\phi \rightarrow \psi)$ να συναγάγουμε το ψ .
4. **Modus Tollens:** Για κάθε πρόταση ϕ και ψ , μπορούμε από το $\neg\psi$ και $(\phi \rightarrow \psi)$ να συναγάγουμε το $\neg\phi$.

Χρησιμοποιώντας τους κανόνες συναγωγής σε ένα σύστημα φυσικής παραγωγής μπορούμε να ελέγξουμε με καθαρά συντακτικό τρόπο αν ένα συμπέρασμα προκύπτει από ένα σύνολο προκειμένων. Για παράδειγμα, έστω οι προκειμένες $\neg P \rightarrow Q$ και $\neg Q$ από τις οποίες θέλουμε να ελέγξουμε αν απορρέει το συμπέρασμα P . Δεν έχουμε παρά να κατασκευάσουμε μια απόδειξη, δηλαδή να ξεκινήσουμε από τις προκειμένες και χρησιμοποιώντας τους κανόνες συναγωγής μέσα από έναν πεπερασμένο αριθμό βημάτων να καταλήξουμε στο συμπέρασμα:

1. $\neg P \rightarrow Q$ [προκειμένη]
2. $\neg Q$ [προκειμένη]
3. $\neg\neg P$ [με χρήση Modus Tollens από 1 και 2]
4. P [με χρήση Διπλής Άρνησης από 3]

Το ενδιαφέρον, και αυτό που πρέπει εδώ να κρατήσουμε, είναι ότι μια απόδειξη είναι μια καθαρά συντακτική διαδικασία στην οποία δεν παίζει κανένα ρόλο η σημασία των εμπλεκόμενων γλωσσικών προτάσεων. Γι' αυτό το λόγο, μια απόδειξη είναι μια διαδικασία που μπορεί να εκτελεστεί καθαρά μηχανικά ανατρέχοντας κάθε φορά απλώς στη μορφή των εμπλεκόμενων συμβόλων. Φυσικά, δεν αναφερόμαστε εδώ στη διαδικασία εύρεσης μιας απόδειξης, αλλά στη δυνατότητα μηχανικής

εκτέλεσης μιας ήδη υπάρχουσας απόδειξης. Η δυνατότητα μηχανικής εκτέλεσης των αποδείξεων θα γίνει προφανέστερη στη συνέχεια, όταν φτάσουμε στην παρουσίαση των υπολογιστικών μηχανών.

1.3.3 Η Τυπική Σημασιολογία

Ας έρθουμε τώρα εν συντομία στο δεύτερο είδος θεωρίας που αρχίζει να αναπτύσσεται με αφετηρία το έργο του Frege. Πρόκειται για την τυπική σημασιολογία. Η τυπική σημασιολογία εξετάζει τις σχέσεις των τυπικών γλωσσών με τον κόσμο ή, ακριβέστερα, με τις ερμηνείες που μπορούν να αποδοθούν στις τυπικές γλώσσες. Αν, για ευκολία, παραμείνουμε στο επίπεδο του προτασιακού λογισμού, τότε η τυπική σημασιολογία παίρνει τη μορφή της σημασιολογίας τιμών αληθείας (Truth Value Semantics). Σε αυτή την περίπτωση, εφόσον η απλούστερη συντακτική μονάδα της τυπικής γλώσσας είναι η ατομική γλωσσική πρόταση (atomic sentence), εκείνο που μπορεί να αναπαρασταθεί σημασιολογικά είναι η τιμή αληθείας της πρότασης, δηλαδή το εάν είναι αληθής (Α) ή ψευδής (Ψ)²³. Έτσι, λοιπόν, η συγκεκριμένη σημασιολογία βασίζεται στην απόδοση μιας εκ των δύο τιμών αληθείας (Τ, Ψ) στις προτάσεις της τυπικής γλώσσας. Το τι ακριβώς είναι τα δύο αντικείμενα Τ και Ψ που αποδίδουμε στις προτάσεις της τυπικής γλώσσας δεν έχει καμιά σημασία μέσα στο πλαίσιο της Λογικής γι' αυτό και για λόγους ευκολίας χρησιμοποιούμε στη θέση τους τις τιμές 1 και 0. Στη συνέχεια, καθορίζουμε την τιμή αληθείας για κάθε λογικό σύνδεσμο ανάλογα με την τιμή αληθείας των συνδεόμενων προτάσεων. Για παράδειγμα, οι τιμές αληθείας ή οι αληθοπίνακες των λογικών συνδέσμων \neg , \rightarrow , \wedge είναι αντίστοιχα οι εξής:

Άρνηση

P	$\neg P$
1	0
0	1

Συμπεπαγωγή

P	Q	$P \rightarrow Q$
1	1	1
1	0	0
0	1	1
0	0	1

Σύζευξη

P	Q	$P \wedge Q$
1	1	1
1	0	0
0	1	0
0	0	0

Με βάση τους αληθοπίνακες, μπορούμε να ελέγξουμε αν ένα συμπέρασμα είναι λογικά έγκυρο. Εφόσον η λογική εγκυρότητα είναι η διατήρηση της αλήθειας κατά τη μετάβαση από τις προκειμένες στο συμπέρασμα, δεν έχουμε παρά να κατασκευάσουμε τον αληθοπίνακα του συμπεράσματος και να ελέγξουμε αν είναι αληθές το συμπέρασμα σε κάθε περίπτωση που είναι αληθείς όλες οι προκειμένες. Ας πάρουμε το παράδειγμα που χρησιμοποιήσαμε παραπάνω στη θεωρία αποδείξεων.

²³ Για περισσότερα επ' αυτού του θέματος βλέπε κεφάλαιο 3, ενότητα 3.3.1.2.

Προκειμένη 1			Προκειμένη 2		Συμπέρασμα
P	Q	$\neg P$	$\neg Q$	$\neg P \rightarrow Q$	P
1	1	0	0	1	1
1	0	0	1	0	0
0	1	1	0	1	0
0	0	1	1	0	0

Από τον παραπάνω πίνακα διαπιστώνουμε ότι κάθε φορά που και οι δύο προκειμένες είναι αληθείς τότε και το συμπέρασμα είναι αληθές. Με άλλα λόγια, διατηρείται η αλήθεια και κατά συνέπεια το επιχείρημα είναι λογικά έγκυρο.

Αν παραμείνουμε στο πλαίσιο της προτασιακής λογικής, η τυπική σημασιολογία είναι μια μηχανική διαδικασία. Ωστόσο, αν προχωρήσει κανείς στην ανάλυση της λογικής δομής της ατομικής γλωσσικής πρότασης, τότε η τυπική σημασιολογία παύει να είναι μια απλή υπόθεση. Σε αυτή την περίπτωση περνάμε στον κατηγορηματικό λογισμό όπου η τυπική σημασιολογία είναι η θεωρία μοντέλων (model theory)²⁴.

1.3.4 Η πληρότητα του κατηγορηματικού λογισμού

Το βασικό ερώτημα που δημιουργείται μετά από αυτή την πολύ σύντομη παρουσίαση της θεωρίας αποδείξεων και της τυπικής σημασιολογίας είναι πώς συνδέεται η αποδειξιμότητα με τη λογική εγκυρότητα. Πιο συγκεκριμένα, δημιουργούνται, τουλάχιστον, δύο ερωτήματα:

A. Είναι κάθε αποδείξιμο επιχείρημα λογικά έγκυρο;

και

B. Είναι κάθε λογικά έγκυρο επιχείρημα αποδείξιμο;

Αν ισχύει το A τότε λέμε ότι το αποδεικτικό σύστημα (δηλαδή το σύνολο των κανόνων συναγωγής και των αξιωμάτων) είναι βάσιμο (sound)²⁵, ενώ όταν ισχύει το B ότι είναι πλήρες (complete). Αν ένα σύστημα δεν είναι βάσιμο τότε παράγει θεωρήματα που δεν είναι όλα αληθή. Για παράδειγμα σε ένα μη-συνεπές σύστημα μπορεί να αποδειχθεί τόσο μια πρόταση φ όσο και η άρνησή της ($\neg\varphi$). Όμως, είτε η φ είναι αληθής είτε η $\neg\varphi$. Κατά συνέπεια, σε ένα μη συνεπές αποδεικτικό σύστημα ορισμένες από τις αποδεικνύομενες προτάσεις είναι ψευδείς.

Αν ένα αποδεικτικό σύστημα δεν είναι πλήρες τότε υπάρχουν αληθείς προτάσεις της τυπικής γλώσσας που δεν αποδεικνύονται μέσα στο πλαίσιο του αποδεικτικού συστήματος. Σε αυτή την περίπτωση, το

²⁴ Βλ., ενδεικτικά, Αναπολιτάνος 1985, Κολέτσος 1992 και Rynasiewicz 1990.

²⁵ Ο όρος "βασικότητα" (soundness) έχει και μια δεύτερη σημασία που δεν πρέπει να συγχέεται με την παρούσα χρήση. Σύμφωνα με αυτή τη δεύτερη σημασία, βάσιμο είναι ένα έγκυρο επιχείρημα αν είναι αληθείς οι προκειμένες του.

αποδεικτικό σύστημα δεν είναι αρκετά ισχυρό ώστε να παραγάγει όλες τις αληθείς προτάσεις της γλώσσας.

Το βασικό πρόβλημα που τέθηκε κατά την ανάπτυξη του κατηγορηματικού λογισμού είναι αν τα διάφορα αποδεικτικά συστήματα που προτεινόταν στο πλαίσιο αυτής της γλώσσας μπορούσαν να παραγάγουν όλες τις αληθείς προτάσεις της γλώσσας και *μόνον αυτές*. Η λύση του προβλήματος ήρθε από τον Gödel, ο οποίος το 1930 απέδειξε το θεώρημα της πληρότητας για τον κατηγορηματικό λογισμό, σύμφωνα με το οποίο μια γλωσσική πρόταση φ είναι αποδείξιμη σε μια πρωτοβάθμια²⁶ θεωρία αν και μόνο αν η φ είναι αληθής σε όλα τα μοντέλα της θεωρίας²⁷.

Το θεώρημα αυτό εξασφαλίζει τον "παραλληλισμό" της σημασιολογίας με τη σύνταξη. Πιο συγκεκριμένα, εξασφαλίζει ότι οι σημασιολογικές σχέσεις μπορούν να "κατοπιριστούν" κατά κάποιο τρόπο στις συντακτικές σχέσεις: η απόδειξη – μια καθαρά μηχανική διαδικασία, αφού βασίζεται αποκλειστικά στη μορφή των συμβόλων, δηλαδή στη σύνταξη τους– διατηρεί την αλήθεια. Αυτό είναι το πρώτο σημαντικό αποτέλεσμα στο πλαίσιο της μαθηματικής λογικής το οποίο, όπως θα δούμε στο δεύτερο κεφάλαιο, αποτέλεσε μια από τις βάσεις για την ανάπτυξη της τεχνητής νοημοσύνης και της γνωσιακής επιστήμης γενικότερα.

1.3.5 Η ανάπτυξη της θεωρίας του υπολογισμού

Το επόμενο μεγάλο βήμα προς την ανάπτυξη των υπολογιστικών μηχανών έγινε αναμφίβολα το 1931 όταν ο Gödel απέδειξε το θεώρημα της μη πληρότητας για την αριθμητική²⁸. Σύμφωνα με αυτό το θεώρημα "οποιοδήποτε ... σύστημα με το οποίο μπορεί να αναπτυχθεί η αριθμητική είναι ουσιαστικά *μη πλήρες*. Με άλλα λόγια, αν δοθεί *οποιοδήποτε* συνεπές σύνολο αριθμητικών αξιωμάτων, υπάρχουν αληθείς αριθμητικές προτάσεις που δεν μπορούν να αποδειχθούν από το σύνολο αυτό"²⁹. Το θεώρημα του Gödel αποτέλεσε το καθοριστικό πλήγμα στο εγχείρημα αναγωγής των

²⁶ Μια θεωρία είναι πρωτοβάθμια αν "οι μεταβλητές έχουν σαν πεδίο τιμών τους μαθηματικά αντικείμενα και όχι κλάσεις τέτοιων αντικειμένων και ... οι ποσοδείκτες \forall και \exists μπορούν να δρουν μόνον πάνω σε τέτοιες μεταβλητές" (Αναπολιτάνος 1985, σελ. 179).

²⁷ Για μια αυστηρή διατύπωση του θεωρήματος του Gödel βλ. Αναπολιτάνος 1985 και Crossley et.al.1972.

²⁸ Για μια εκλαϊκευτική παρουσίαση του θεωρήματος βλ. Nagel & Newman 1991.

²⁹ Nagel & Newman 1991, σελ. 59. Ο Gödel έφτασε και σε ένα δεύτερο πολύ σημαντικό συμπέρασμα που δεν θα μας απασχολήσει εδώ: η συνέπεια της αριθμητικής δεν μπορεί να αποδειχθεί μέσα στο πλαίσιο της ίδιας της αριθμητικής (βλ. Αναπολιτάνος 1985, σελ. 257, Nagel & Newman 1991 και Crossley et.al.1972).

μαθηματικών στη λογική³⁰. Ωστόσο, ο λόγος για τον οποίον θεωρούμε ότι το εν λόγω θεώρημα συνέβαλε στην ανάπτυξη των υπολογιστικών μηχανών δεν είναι αυτός. Το θεώρημα του Gödel δεν είχε μόνο "αρνητικό" χαρακτήρα, δηλαδή το να επισημαίνει τα εγγενή όρια της λογικής, αλλά έδειξε ότι είναι δυνατόν να αποδείξουμε το τι δεν είναι αποδείξιμο ή, καλύτερα, υπολογίσιμο. Ωστόσο, όπως επισημαίνουν οι Crossley et al (1972, σελ. 31) "το να δείξει κανείς ότι καμιά μηχανική μέθοδος δεν επιλύει ένα δεδομένο πρόβλημα προϋποθέτει έναν ορισμό που συμπεριλαμβάνει όλους τους τύπους υπολογισμού". Με άλλα λόγια, για να δείχθει ότι ένα πρόβλημα δεν είναι υπολογίσιμο πρέπει με κάποιο τρόπο να γνωρίζουμε τι σημαίνει υπολογισμός. Οι προσπάθειες να απαντηθεί αυτό το ερώτημα οδήγησαν στην ανάπτυξη της θεωρίας του υπολογισμού.

Σε μια πρώτη προσέγγιση, η έννοια της υπολογισιμότητας συνεπάγεται την ύπαρξη ενός αλγορίθμου. Η έννοια του αλγορίθμου παρότι είναι μια από τις παλαιότερες έννοιες των μαθηματικών εξακολουθεί μέχρι σήμερα να μην είναι πλήρως κατανοητή³¹. Σε γενικές γραμμές, ένας αλγόριθμος είναι "μια μηχανική ή αυτόματη διαδικασία η οποία, τουλάχιστον ιδεατά, θα υπολογίζει κάτι για οποιαδήποτε περίπτωση από μια απειρία διαφορετικών δυνατών περιπτώσεων"³². Κάτω πιο συγκεκριμένα, ο Haugeland (1992, σελ. 93) ορίζει την έννοια του αλγορίθμου ως εξής:

"ένας αλγόριθμος είναι μια αλάθητη βήμα προς βήμα συνταγή για να επιτευχθεί ένας προκαθορισμένος στόχος. 'Αλάθητη' σημαίνει ότι είναι εγγυημένο πως η διαδικασία θα επιτύχει οπωσδήποτε έπειτα από πεπερασμένο αριθμό βημάτων (αν υποθέσουμε πως κάθε βήμα εκτελείται σωστά... 'βήμα προς βήμα' σημαίνει τρία πράγματα: (1) η συνταγή ορίζει ένα βήμα κάθε φορά, το ένα μετά το άλλο, (2) έπειτα από κάθε βήμα, το επόμενο είναι πλήρως καθορισμένο (δεν υπάρχουν επιλογές ή αβεβαιότητες), και (3) έπειτα από κάθε βήμα, το επόμενο είναι φανερό (δεν απαιτείται ούτε έμπνευση ούτε ευφυΐα για να ανακαλυφθεί ποιο πρέπει να είναι)."³³

Για παράδειγμα, ένα χαρακτηριστικό πρόβλημα που επιδέχεται αλγοριθμικής επίλυσης είναι το πώς εντοπίζουμε από μια αρμαθιά κλειδιά

³⁰ Το άρθρο του Gödel στρέφεται ρητά κατά των *Principia Mathematica* δηλαδή κατά του λογικού συστήματος που, συνεχίζοντας και διορθώνοντας το έργο του Frege, ανέπτυξαν ο Russell και ο Whitehead προκειμένου να αναγάγουν τα μαθηματικά στη λογική. Επίσης, το θεώρημα του Gödel έθεσε τέρμα και στο φορμαλιστικό πρόγραμμα του Hilbert ο οποίος ήθελε να αναγάγει τα μαθηματικά στην αριθμητική (βλ. Αναπολιτάνος 1985).

³¹ Βλ. Glymour 1998, σελ. 510.

³² Glymour 1998, σελ. 511.

³³ Για μια εις βάθος ανάλυση της έννοιας του υπολογισμού βλ. Copeland 1996.

εκείνα τα κλειδιά που ανοίγουν μια συγκεκριμένη κλειδαριά: απλώς δοκιμάζουμε με τη σειρά όλα τα κλειδιά.

Από τις αρχές της δεκαετίας του 1930, το πρόβλημα της υπολογισιμότητας οδήγησε στην ανάπτυξη πολλών διαφορετικών θεωριών. Ωστόσο, καμία από αυτές τις προσπάθειες δεν απέφερε έναν γενικό προσδιορισμό της έννοιας του αλγορίθμου και, κατ' επέκταση, της έννοιας των υπολογισμών συναρτήσεων, δηλαδή των συναρτήσεων για τις οποίες υπάρχει κάποιος αλγόριθμος. Όπως επισημαίνει ο Glymour (1998, σελ. 529): "Αντί να απαντήσει στα ερωτήματα αυτά, η θεωρία του υπολογισμού αναπτύχθηκε προτείνοντας συγκεκριμένα υπολογιστικά συστήματα, και χαρακτηρίζοντας κατόπιν ως υπολογισμό αυτό που μπορεί να υπολογιστεί σε κάποιο από αυτά τα συστήματα". Ένα τέτοιο σύστημα αναπτύχθηκε από τον Turing, ο οποίος επεχείρησε να προσδιορίσει την έννοια της υπολογίσιμης συνάρτησης με τη βοήθεια μιας μηχανής –της καθολικής μηχανής Turing. Το ερώτημα που δημιουργείται είναι αν η καθολική μηχανή Turing έχει τη δυνατότητα να υπολογίζει όλες τις δυνατές υπολογίσιμες συναρτήσεις. Σε αυτό το ερώτημα δεν υπάρχει μια μαθηματική απόδειξη αλλά πολύ ισχυρές ενδείξεις οι οποίες οδήγησαν τον Church στην διατύπωση της ακόλουθης θέσης:

Οι υπολογίσιμες συναρτήσεις αριθμών είναι οι συναρτήσεις που είναι υπολογίσιμες κατά Turing³⁴

Το ενδιαφέρον με τον προσδιορισμό των υπολογίσιμων συναρτήσεων μέσω της καθολικής μηχανής Turing είναι ότι με αυτόν τον τρόπο η έννοια του υπολογισμού συνδέεται με τις διαδικασίες μιας μηχανής. Όπως θα δούμε στο δεύτερο κεφάλαιο, η συγκεκριμένη προσέγγιση επιτυγχάνει ουσιαστικά να συνδέσει το συντακτικό επίπεδο με το επίπεδο της φυσικής υλοποίησης. Δεν μιλούμε πλέον αφηρημένα για κάποιους υπολογισμούς αλλά για διαδικασίες που μπορούν να πραγματοποιηθούν αυτόματα από μια μηχανή³⁵.

1.3.6 Πώς λειτουργεί μια μηχανή Turing;

Η καθολική μηχανή Turing είναι μια μηχανή που μπορεί να προσομοιώσει τη λειτουργία κάθε απλής μηχανής Turing. Προκειμένου να κατανοήσουμε τι ακριβώς σημαίνει αυτό πρέπει πρώτα να εξετάσουμε πώς λειτουργεί μια απλή μηχανή Turing³⁶.

³⁴ Βλ. Glymour 1998, σελ. 526.

³⁵ Βέβαια, ο Turing δεν κατασκεύασε αυτές τις μηχανές, απλώς προσδιόρισε θεωρητικά τον τρόπο λειτουργίας τους. Δεδομένου, μάλιστα, ότι ενέχουν άπειρα στοιχεία, όπως θα δούμε στην επόμενη ενότητα, θα ήταν αδύνατη η πιστή κατασκευή τους.

³⁶ Για μια αναλυτική περιγραφή της λειτουργίας των μηχανών Turing βλέπε Crossley et.al.1972, Weizenbaum 1976 και Penrose 1989.

Κάθε απλή μηχανή Turing υπολογίζει μια συγκεκριμένη συνάρτηση επί των φυσικών αριθμών. Έτσι, για κάθε αριθμό που εισάγουμε η μηχανή εξάγει την τιμή της συνάρτησης. Η μηχανή αποτελείται από μια απεριόριστη ταινία η οποία είναι χωρισμένη σε τετράγωνα. Σε κάθε τετράγωνο μπορεί να γραφτεί ένα σύμβολο από ένα πεπερασμένο αλφάβητο συμβόλων. Η μηχανή διαθέτει επίσης μια κεφαλή ανάγνωσης – εγγραφής η οποία κινείται κατά μήκος της ταινίας από τετράγωνο σε τετράγωνο. Από φυσική άποψη, η μηχανή μπορεί να κάνει τρία πράγματα: (1) να σβήσει ένα σύμβολο και να γράψει ένα άλλο στη θέση του, (2) να μετακινήσει την κεφαλή ανάγνωσης – εγγραφής προς τα δεξιά, και (3) να μετακινήσει την κεφαλή ανάγνωσης – εγγραφής προς τα αριστερά.

Ένα πεπερασμένο σύνολο οδηγιών προσδιορίζει πλήρως τη συμπεριφορά της μηχανής με βάσει δύο παράγοντες:

1. το περιεχόμενο του τετραγώνου στο οποίο βρίσκεται κάθε φορά η κεφαλή ανάγνωσης - εγγραφής και
2. την εσωτερική κατάσταση της κεφαλής.

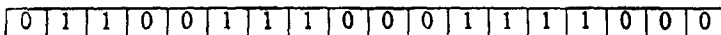
Οι οδηγίες της εκάστοτε μηχανής Turing μπορούν να συνοψιστούν σε έναν πίνακα διπλής εισόδου που ονομάζεται πίνακας μηχανής (machine table). Η μία είσοδος του πίνακα αφορά το περιεχόμενο των τετραγώνων και η άλλη την εσωτερική κατάσταση της κεφαλής. Κάθε οδηγία προσδιορίζει τρία πράγματα:

1. ποιο σύμβολο θα γραφτεί στο τετράγωνο που διαβάζει η κεφαλή
2. προς ποια κατεύθυνση θα κινηθεί η κεφαλή (ένα τετράγωνο προς τα αριστερά ή ένα τετράγωνο προς τα δεξιά) ή αν θα σταματήσει και
3. ποια θα είναι η νέα εσωτερική κατάσταση στην οποία θα βρεθεί η κεφαλή πηγαίνοντας στο επόμενο τετράγωνο

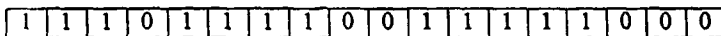
Ας δούμε ένα στοιχειώδες παράδειγμα³⁷. Έστω ότι τα σύμβολα που μπορούν να γραφτούν στη ταινία μιας μηχανής Turing είναι το 0 και το 1, ενώ οι δυνατές εσωτερικές καταστάσεις της κεφαλής ανάγνωσης – εγγραφής είναι η A και η B. Θα ορίσουμε μια μηχανή Turing που υπολογίζει για κάθε αριθμό τον επόμενο του ή, με άλλα λόγια, υπολογίζει τη συνάρτηση $f(x)=x+1$. Καταρχάς, τίθεται το ερώτημα πώς μπορούν να αναπαρασταθούν οι διάφοροι αριθμοί; Δεδομένου ότι έχουμε στη διάθεσή μας δύο σύμβολα, το 0 και το 1, θα πρέπει όλοι οι αριθμοί να εκφραστούν συναρτήσει των συγκεκριμένων συμβόλων. Ο πιο οικονομικός τρόπος είναι να χρησιμοποιήσουμε το δυαδικό σύστημα αρίθμησης. Ωστόσο εδώ για λόγους εποπτικότητας μπορούμε να χρησιμοποιήσουμε μια άλλη

³⁷ Βλ. Crane 1995, σελ. 94-95.

σύμβαση: κάθε αριθμός θα αναπαριστάται με ισάριθμο πλήθος από 1, ενώ το 0 δεν θα αναπαριστά κάποιον αριθμό αλλά το κενό που διαχωρίζει διαφορετικούς αριθμούς. Για παράδειγμα, το τμήμα της ταινίας που απεικονίζεται παρακάτω αναπαριστά κατά σειρά τους αριθμούς 2, 3 και 4.



Το ζητούμενο, λοιπόν, είναι να οριστούν οι κατάλληλες οδηγίες που θα επιτρέπουν στη μηχανή Turing να υπολογίζει τη συνάρτηση $f(x)=x+1$. Αν, για παράδειγμα, δοθεί η παραπάνω είσοδος θα πρέπει η μηχανή μετά την επεξεργασία³⁸ να παράγει στη θέση της, κατά σειρά, τους αριθμούς 3, 4 και 5, δηλαδή το ακόλουθο τμήμα ταινίας:



Μια μηχανή Turing με αυτή τη δυνατότητα ορίζεται από τον ακόλουθο πίνακα μηχανής:

		Περιεχόμενο του τετράγωνα που διαβάζει η κεφαλή ανάγνωσης - εγγραφής	
		1	0
Κατάσταση της κεφαλής ανάγνωσης - εγγραφής	A	<ul style="list-style-type: none"> • Μετάβαλλε την εσωτερική κατάσταση σε B, • Γράψε στο τετράγωνο το 1, • Μετακίνησε την κεφαλή προς τα αριστερά 	<ul style="list-style-type: none"> • Παράμεινε στην εσωτερική κατάσταση A, • Γράψε στο τετράγωνο το 0, • Μετακίνησε την κεφαλή προς τα αριστερά
	B	<ul style="list-style-type: none"> • Παράμεινε στην εσωτερική κατάσταση B, • Γράψε στο τετράγωνο το 1, • Μετακίνησε την κεφαλή προς τα αριστερά 	<ul style="list-style-type: none"> • Μετάβαλλε την εσωτερική κατάσταση σε A, • Γράψε στο τετράγωνο το 1, • Σταμάτα

Ο πίνακας αυτός καθορίζει πλήρως την λειτουργία της μηχανής Turing για κάθε διαφορετικό συνδυασμό της κατάστασης της κεφαλής ανάγνωσης - εγγραφής και του συμβόλου που διαβάζει η κεφαλή. Πρόκειται για οδηγίες της εξής μορφής: Αν η κεφαλή ανάγνωσης - εγγραφής βρίσκεται στην εσωτερική κατάσταση A και διαβάζει το σύμβολο 1, τότε μετάβαλλε την εσωτερική κατάσταση σε B, γράψε στο τετράγωνο το 1 και μετακίνησε την κεφαλή προς τα αριστερά.

Η μηχανή Turing του παραπάνω παραδείγματος θα τερματίσει τη λειτουργία της μόλις προσθέσει τη μονάδα στον αριθμό που της δίδεται ως είσοδος. Ας δούμε λίγο πιο αναλυτικά πώς ακριβώς συμβαίνει αυτό. Έστω ότι η μηχανή ξεκινάει από τα δεξιά με αρχική εσωτερική κατάσταση της κεφαλής την A και ότι η είσοδος είναι ο αριθμός 2 όπως φαίνεται στο παρακάτω τμήμα ταινίας:

³⁸ Για την ακρίβεια, η μηχανή πρέπει να τεθεί σε λειτουργία τρεις φορές.

0	0	0	0	0	1	1	0
---	---	---	---	---	---	---	---

Το συγκεκριμένο τμήμα της ταινίας θα τροποποιηθεί διαδοχικά ως εξής:

Μορφή της ταινίας ³⁹	Κατάσταση κεφαλής	Περιεχόμενο τετραγώνου	Νέα κατάσταση κεφαλής	Σύμβολο που γράφει η κεφαλή στο τετράγωνο	Κίνηση της κεφαλής
00000110	A	0	A	0	Προς τα αριστερά
00000110	A	1	B	1	Προς τα αριστερά
00000110	B	1	B	1	Προς τα αριστερά
00000110	B	0	A	1	Σταμάτημα
00001110					

Βλέπουμε, λοιπόν, ότι η μηχανή Turing που περιγράψαμε μπορεί να υπολογίσει τη συνάρτηση $f(x)=x+1$ χρησιμοποιώντας ένα πεπερασμένο αριθμό εντελώς στοιχειωδών λειτουργιών. Ο Turing υποστήριξε με διάφορα επιχειρήματα (αλλά δεν απέδειξε) ότι οποιαδήποτε υπολογίσιμη συνάρτηση⁴⁰, οσοδήποτε σύνθετη, μπορεί να υπολογιστεί από κάποια μηχανή Turing, δηλαδή μια μηχανή που χρησιμοποιεί τις στοιχειώδεις λειτουργίες που περιγράφηκαν παραπάνω⁴¹. Η σημαντικότερη, όμως, συνεισφορά του Turing είναι ότι απέδειξε ότι κάθε μηχανή Turing μπορεί να προσομοιωθεί από μία και μόνη μηχανή -την καθολική μηχανή Turing. Αυτή η μηχανή έχει τη δυνατότητα να μιμείται τη λειτουργία κάθε άλλης μηχανής Turing και έτσι έχει τη δυνατότητα να υπολογίζει κάθε

³⁹ Ο αριθμός που διαβάζει κάθε φορά η κεφαλή δηλώνεται με την υπογράμμιση.

⁴⁰ Σύμφωνα με τον Copeland (1997), ο όρος "οποιαδήποτε υπολογίσιμη συνάρτηση" επιδέχεται δύο τουλάχιστον αναγνώσεις. Η κυρίαρχη ανάγνωση είναι ότι σημαίνει οποιαδήποτε συνάρτηση που είναι υπολογίσιμη από οποιαδήποτε άλλη μηχανή. Ο Copeland ισχυρίζεται ότι ο Turing δεν υποστήριξε αυτή τη θέση αλλά μια ασθενέστερη, ότι δηλαδή οποιαδήποτε συνάρτηση που είναι υπολογίσιμη από ένα ανθρώπινο ον χωρίς τη βοήθεια μηχανισμών είναι υπολογίσιμη από μια μηχανή Turing. Σε ένα μεταγενέστερο άρθρο του ο Copeland (Copeland & Sylvan 1999) υποστηρίζει ότι "το σύνολο των καλώς ορισμένων υπολογισμών δεν εξαντλείται από τους υπολογισμούς που μπορούν να εκτελεστούν από μια μηχανή Turing".

⁴¹ Πρόκειται, ουσιαστικά, για την θέση του Church, στην οποία όμως ο Turing έφτασε ανεξάρτητα. Γι' αυτό στη βιβλιογραφία η θέση αυτή συχνά ονομάζεται θέση Church-Turing.

υπολογίσιμη συνάρτηση. Η καθολική μηχανή Turing εισάγει, ουσιαστικά, την έννοια του ψηφιακού υπολογιστή για κάθε σκοπό (*general purpose digital computer*). Αυτή η ιδιότητα αποτελεί ίσως τη χαρακτηριστικότερη ιδιότητα των ηλεκτρονικών υπολογιστών: οι σύγχρονοι υπολογιστές δεν κατασκευάζονται για να επιτελούν κάποιο συγκεκριμένο υπολογιστικό έργο, αλλά μπορούν να χρησιμοποιηθούν για την εκτέλεση οποιοδήποτε αλγορίθμου. Πρόκειται, ουσιαστικά, για μηχανές που μπορούν να αλλάζουν τη λειτουργία τους χωρίς να τροποποιούν τα μηχανικά τους μέρη.

Η μεγάλη συνεισφορά του Turing ήταν ότι συνέλαβε αυτό το είδος μηχανών και περιέγραψε με ακρίβεια τον τρόπο λειτουργίας τους. Η βασική ιδέα, όπως είπαμε, είναι ότι η καθολική μηχανή Turing έχει τη δυνατότητα να μιμείται τη λειτουργία κάθε άλλης μηχανής Turing. Αυτή η δυνατότητα μίμησης επιτυγχάνεται, ουσιαστικά, μέσω του προγραμματισμού της καθολικής μηχανής Turing. Χωρίς να εισέλθουμε σε λεπτομέρειες μπορούμε να δεχθούμε πως προγραμματισμός, σε αυτή την πρώτη του μορφή, σημαίνει κωδικοποίηση του πίνακα μηχανής της προς απομίμηση μηχανής και χρήση αυτής της κωδικοποίησης ως τμήμα της εισόδου της καθολικής μηχανής Turing.

1.3.7 Η μηχανή von Neumann

Από τις αρχές της δεκαετίας του 1940 άρχισαν να κατασκευάζονται οι πρώτοι ψηφιακοί υπολογιστές. Από τότε έχουν γίνει πολλές τεχνικές βελτιώσεις στα υλικά κατασκευής, στην αρχιτεκτονική των υπολογιστών και στον τρόπο προγραμματισμού. Μια εξέλιξη στην αρχιτεκτονική των υπολογιστών έγινε από τον von Neumann. Η εξέλιξη ήταν τόσο σημαντική που αναφερόμαστε πλέον σε αυτή τη νέα αρχιτεκτονική με τον όρο "μηχανή von Neumann". Η αρχιτεκτονική von Neumann χαρακτηρίζει στις γενικές της γραμμές ακόμα και τους σημερινούς υπολογιστές.

Ένα πρώτο βασικό χαρακτηριστικό της αρχιτεκτονικής που ανέπτυξε ο von Neumann είναι ότι προτείνει ένα νέο τρόπο προσπέλασης της μνήμης. Η μηχανή von Neumann μπορεί να προσπελάσει τη μνήμη όχι μόνο κατά σχετικό τρόπο, όπως συνέβαινε στη μηχανή Turing, αλλά και κατ' απόλυτο τρόπο. Στον σχετικό τρόπο προσπέλασης ο εντοπισμός της επόμενης θέσης γίνεται με βάση την τρέχουσα θέση. Για παράδειγμα, στη μηχανή Turing η επόμενη θέση στην οποία κάθε φορά κινείται η κεφαλή ορίζεται με αφετηρία την τρέχουσα θέση της. Αντίθετα, η απόλυτη προσπέλαση καθορίζει τις θέσεις της μνήμης μέσω κάποιου ονόματος (ή αριθμού) που χαρακτηρίζει την εκάστοτε θέση με μονοσήμαντο τρόπο. Κατ' αυτόν τον τρόπο, η κάθε θέση μνήμης ορίζεται ανεξάρτητα από όλες τις υπόλοιπες. Για παράδειγμα, στη μηχανή Turing θα υπήρχε η δυνατότητα απόλυτης προσπέλασης αν σε κάθε τετράγωνο της ταινίας αποδιδόταν ένας διαφορετικός αριθμός. Σε αυτή την περίπτωση, η κεφαλή θα μπορούσε να προσπελάσει κάποιο συγκεκριμένο τετράγωνο από

οποιαδήποτε άλλη θέση χρησιμοποιώντας την απόλυτη διεύθυνση του τετραγώνου, δηλαδή τον αριθμό που του έχει αποδοθεί.

Η δεύτερη καινοτομία της μηχανής von Neumann είναι ότι επιτρέπει την αποθήκευση όχι μόνο των δεδομένων αλλά και των προγραμμάτων. Η πραγματική καινοτομία εδώ είναι ότι τα προγράμματα μπορούν να χρησιμοποιήσουν τη μνήμη απόλυτης προσπέλασης. Αυτή η δυνατότητα δίνει την ευχέρεια στον υπολογιστή να μην ακολουθεί βήμα-βήμα ένα πρόγραμμα, αλλά να μπορεί να μετακινείται μπρος και πίσω στις εντολές του προγράμματος κάνοντας χρήση της απόλυτης διεύθυνσης κάθε εντολής του προγράμματος. Έτσι, ένα πρόγραμμα μπορεί να συσταθεί από ένα σύνολο υπορουτινών, δηλαδή ένα σύνολο μικρών αυτόνομων προγραμμάτων. Κάθε υπορουτίνα επιτελεί συνήθως κάποια σχετικά απλή λειτουργία που χρησιμοποιείται πολλές φορές από το βασικό πρόγραμμα. Προκειμένου, λοιπόν, να μην επαναλαμβάνονται μέσα στον κορμό του βασικού προγράμματος οι εντολές που περιέχει η υπορουτίνα κάθε φορά που αυτό απαιτείται, μπορεί να γίνει χρήση της απόλυτης διεύθυνσης, δηλαδή να καλείται η υπορουτίνα μέσω του ονόματός της και αφού εκτελεστεί να επιστρέφει ο έλεγχος στο βασικό πρόγραμμα.

Αυτή η ιδέα αποτέλεσε τη βάση για τη δημιουργία των διαφόρων γλωσσών προγραμματισμού. Κάθε γλώσσα προγραμματισμού συνίσταται από μια "βιβλιοθήκη" υπορουτινών και ένα σύστημα παρακολούθησης και συντονισμού της χρήσης τους. Οι στοιχειώδεις ενέργειες μέσω σε μια γλώσσα προγραμματισμού δεν είναι πλέον οι εντελώς "ανόητες" κινήσεις της μηχανής Turing, αλλά σύνθετες ενέργειες που μια μηχανή Turing θα απαιτούσε πάρα πολλά βήματα για να τις επιτελέσει. Κατ' αυτόν τον τρόπο, έχουν δημιουργηθεί γλώσσες προγραμματισμού που είναι πολύ φιλικές προς τον χρήστη υπό την έννοια ότι πλησιάζουν αρκετά στη φυσική γλώσσα με εντολές όπως "αναζήτησε", "ταξινόμησε", "τόπωσε" κ.λπ. Οι γλώσσες προγραμματισμού αποτελούν ουσιαστικά μηχανές υψηλότερου επιπέδου. Με την ανάπτυξη των γλωσσών προγραμματισμού αρχίζει πλέον να διακρίνεται με σαφήνεια το επίπεδο του λογισμικού (software) από την αρχιτεκτονική του υπολογιστή, δηλαδή το επίπεδο του υλισμικού (hardware).

1.3.8 Η ανάδυση της Τεχνητής Νοημοσύνης

Στις αρχές της δεκαετίας του 1950 οι αλματώδεις εξελίξεις στην τεχνολογία των ψηφιακών υπολογιστών είχαν αρχίσει να πείθουν όλο και περισσότερους ερευνητές ότι αυτές οι μηχανές μπορούν να χρησιμοποιηθούν ως πρότυπο για την εξήγηση του ανθρώπινου νου. Οι φωνές υποστήριξης έρχονταν από πολλά διαφορετικά γνωστικά πεδία και μια ρητορική είχε αρχίσει να αναπτύσσεται για την ομοιότητα του ηλεκτρονικού υπολογιστή/εγκέφαλου με τον ανθρώπινο εγκέφαλο και νου.

Σημαντικό ρόλο σε αυτή την εξέλιξη έπαιξε η σταδιακή συνειδητοποίηση ότι οι ψηφιακοί υπολογιστές δεν ήταν απλώς αριθμητικοί υπολογιστές αλλά γενικοί χειριστές συμβόλων. Βέβαια, η ιδέα ότι οι υπολογιστές είναι γενικοί χειριστές συμβόλων υπάρχει ήδη στην

αρχική σύλληψη αυτών των μηχανών από τον Turing. Ωστόσο οι πρώτοι υπολογιστές που κατασκευάστηκαν τη δεκαετία του 1940 ήταν αποκλειστικά στραμμένοι στην εκτέλεση αριθμητικών πράξεων. Η αντιμετώπιση των υπολογιστών ως γενικών χειριστών συμβόλων κατέστησε φανερό ότι μπορούν να χρησιμοποιηθούν για την αναπαράσταση και την επεξεργασία λέξεων, προτάσεων και επιχειρημάτων. Οι τυπικές γλώσσες της μαθηματικής λογικής παρείχαν τα βασικά εφόδια για την ανάπτυξη γλωσσών προγραμματισμού που εκτός από τον χειρισμό αριθμητικών δεδομένων επέτρεπαν και τον χειρισμό γλωσσικών δεδομένων.

Σημαντικό ρόλο στην εκκλαίκευση της ιδέας ότι οι υπολογιστές συνδέονται με κάποιο τρόπο με την ανθρώπινη νόηση έπαιξε το άρθρο του Turing "Computing Machinery and Intelligence". Ο Turing (1950) στο άρθρο αυτό έθεσε ρητά το ερώτημα αν οι υπολογιστικές μηχανές μπορούν να σκέφτονται και πρότεινε μια συμπεριφοριστική μέθοδο για την απάντηση του ερωτήματος. Πρόκειται για τον έλεγχο Turing (Turing test). Ο έλεγχος Turing βασίζεται σε ένα παιχνίδι προσποίησης με τρεις παίκτες. Οι δύο παίκτες είναι άνθρωποι και ο τρίτος υπολογιστής. Ο ένας από τους δύο ανθρώπους είναι ο ανακριτής και έχει ως αποστολή του να θέτει ερωτήματα μέσω ενός τερματικού στους δύο άλλους παίκτες προκειμένου να διακρίνει ποιος από τους δύο είναι ο υπολογιστής. Η αποστολή του υπολογιστή, από την άλλη μεριά, είναι να μιμηθεί όσο το δυνατόν καλύτερα την ανθρώπινη (γλωσσική) συμπεριφορά προκειμένου να μπερδέψει τον ανακριτή. Ο Turing ισχυρίζεται ότι αν σε αυτό το παιχνίδι ο ανακριτής φτάσει να κάνει λάθος τόσο συχνά όσο θα έκανε αν αποστολή του ήταν να διακρίνει έναν άνδρα από μια γυναίκα, τότε μπορούμε να δεχθούμε ότι ο υπολογιστής σκέφτεται.

Πολλά άλλα γεγονότα και εξελίξεις συνέβαλλαν στην ανάδυση της τεχνητής νοημοσύνης η οποία τελικά συστάθηκε ως κλάδος το 1956 στο συνέδριο του Dartmouth, όπου και απέκτησε το όνομά της από τον John McCarthy. Από τότε έχει ξεκινήσει μια πολυδάπανη και πολυδιαφημισμένη προσπάθεια για την επίτευξη των στόχων που τέθηκαν από τη σύσταση του κλάδου. Ένας από αυτούς τους στόχους είναι να κατασκευαστούν σκεπτόμενες μηχανές και, ειδικότερα, μηχανές που σκέφτονται όπως ο άνθρωπος. Το τμήμα της TN που υπηρετεί αυτόν το στόχο, όπως θα δούμε στο επόμενο κεφάλαιο, είναι ακριβώς εκείνο που αφορά τη γνωσιακή επιστήμη.

1.4 Εξελίξεις στην φιλοσοφία που συνέβαλαν στην ανάδυση της γνωσιακής επιστήμης

Στις αρχές του 20ου αιώνα με την επονομαζόμενη γλωσσική στροφή, η αναλυτική φιλοσοφία της γλώσσας καθίσταται "πρώτη φιλοσοφία" παίρνοντας τη θέση της φιλοσοφίας της συνείδησης. Μέσα στη νέα φιλοσοφική παράδοση που δημιουργείται, το πρόβλημα της συνείδησης αντιμετωπίζεται στο πεδίο της γλώσσας. Η έρευνα δεν αφορά πλέον το

πώς τα νοητικά φαινόμενα εμφανίζονται στη συνείδηση, αλλά τους τρόπους με τους οποίους μιλάμε για αυτά.

Η κριτική του ψυχολογισμού από τον Frege αναδεικνύει το αίτημα για τη διασφάλιση της διύποκειμενικότητας της γλωσσικής σημασίας. Αν οι γλωσσικοί όροι αναφέρονταν σε κάποιες εσωτερικές καταστάσεις του υποκειμένου όπως, για παράδειγμα, κάποιες ιδέες, τότε οι γλωσσικές σημασίες δεν θα ήταν διύποκειμενικά προσβάσιμες. Αυτό το αίτημα οδηγεί την αναλυτική φιλοσοφία και, ειδικότερα, το Λογικό Θετικισμό στην ανάπτυξη ενός εξειδικευμένου κριτηρίου για τον τρόπο που πρέπει να προσδιορίζεται η γλωσσική σημασία. Πρόκειται για το κριτήριο της επαληθευσσιμότητας (verification): η σημασία μιας γλωσσικής πρότασης εξαντλείται από τις συνθήκες που πρέπει να ισχύουν αν θεωρηθεί η πρόταση αληθής. Με άλλα λόγια, η σημασία μιας γλωσσικής πρότασης είναι οι συνθήκες επαληθευσσιμότητάς της⁴².

1.4.1 Ο Λογικός Συμπεριφορισμός

Η εφαρμογή του κριτηρίου της επαληθευσσιμότητας στις γλωσσικές προτάσεις που αναφέρονται σε νοητικά φαινόμενα οδηγεί σε αυτό που ονομάστηκε "λογικός" ή "αναλυτικός συμπεριφορισμός". Για παράδειγμα, η σημασία της πρότασης "Ο Πέτρος πονάει" εξαντλείται στο σύνολο των συμπεριφορών που θα ήταν δυνατό να παρουσιάσει ο Πέτρος αν όντως πονούσε. Γι' αυτό ο Hempel (1935/1980, σελ. 19), βασικός υποστηρικτής του λογικού συμπεριφορισμού, επισημαίνει: "Βλέπουμε καθαρά ότι η σημασία μιας ψυχολογικής δήλωσης συνίσταται αποκλειστικά στη λειτουργία της σύντηξης της περιγραφής συγκεκριμένων τρόπων φυσικής απόκρισης που χαρακτηρίζουν τα σώματα των ανθρώπων και των ζώων".

Ο λογικός συμπεριφορισμός, λοιπόν, είναι μια θεωρία που υποστηρίζει ότι οι ψυχολογικές προτάσεις είναι μεταφράσιμες χωρίς απώλεια περιεχομένου σε προτάσεις που αφορούν συμπεριφορές ή προδιαθέσεις για συμπεριφορά. Για παράδειγμα, σύμφωνα με τον Hempel (1935/1980, σελ. 17), η πρόταση "Ο Πέτρος έχει πονόδοντο" μπορεί να μεταφραστεί χωρίς απώλεια περιεχομένου σε ένα σύνολο προτάσεων ελέγχου, ορισμένες από τις οποίες είναι οι ακόλουθες:

α. Ο Πέτρος οδύρεται και κάνει χειρονομίες αυτών των ειδών.

β. Στην ερώτηση 'Τι συμβαίνει;' ο Πέτρος προφέρει τις λέξεις 'Έχω πονόδοντο'.

γ. Επί τόπου εξέταση αποκαλύπτει ένα χαλασμένο δόντι.

δ. Η πίεση, οι διαδικασίες χώνεψης και οι ταχύτητα των αντιδράσεων του Πέτρου δείχνουν αλλαγές αυτών των ειδών.

ε. Διαδικασίες αυτών των ειδών συμβαίνουν στο κεντρικό νευρικό σύστημα του Πέτρου."

⁴² Βλ., ενδεικτικά, Hempel 1935/1980, σελ. 17.

Οι παραπάνω προτάσεις συνιστούν προτάσεις ελέγχου γιατί μέσω αυτών μπορεί να ελεγχθεί η αλήθεια της πρότασης "Ο Πέτρος έχει πονόδοντο". Ο ισχυρισμός του Hempel, εν προκειμένω, είναι ότι οι συγκεκριμένες προτάσεις μαζί με κάποιες άλλες εξαντλούν τη σημασία της πρότασης "Ο Πέτρος έχει πονόδοντο". Ωστόσο, δεν χρειάζεται να προσπαθήσει κανείς πολύ για να διαπιστώσει την προβληματικότητα αυτού του ισχυρισμού. Καταρχάς, οι προτάσεις δ και ε μπορούν εύκολα να απορριφθούν γιατί δεν αποτελεί προϋπόθεση για τη γνώση της σημασίας της λέξης "πονόδοντος" η γνώση του τι συμβαίνει στο κεντρικό νευρικό σύστημα ή της πίεσης, των διαδικασιών χώνευσης και των ταχυτήτων αντίδρασης⁴³. Ούτε, όμως, η πρόταση γ περιγράφει ένα αναγκαίο κριτήριο για την ύπαρξη πονόδοντου, γιατί κάποιος μπορεί να νιώθει πόνο στα δόντια χωρίς να υπάρχει κάποιο πρόβλημα στα δόντια. Όσον αφορά την πρόταση β, αυτή συσχετίζει τον πονόδοντο με μια συγκεκριμένη γλωσσική συμπεριφορά. Ωστόσο, όπως επισημαίνει ο Kim (1996, σελ. 31), "προκειμένου η απόκριση του Πέτρου να είναι εδώ συναφής θα πρέπει να καταλαμβάνει την ερώτηση 'Τι συμβαίνει;' και να βεβαιώνει ότι η πρόταση 'Έχω πονόδοντο' σημαίνει ότι έχει πονόδοντο. Δηλαδή, ο Πέτρος πρέπει να είναι ένας επαρκής και να είναι επαρκής χρήστης αυτής της γλώσσας για γλώσσας και το να είναι κανείς επαρκής χρήστης αυτής της γλώσσας διαπροσωπική επικοινωνία είναι μια εκλεπτυσμένη και πολύ σύνθετη ψυχολογική ικανότητα, όχι κάτι που θα μπορούσε να υπαχθεί σε 'κινήσεις και ήχους'".

Τέλος, όσον αφορά την πρόταση α ούτε αυτή αποτελεί αναγκαία προϋπόθεση για την ύπαρξη πόνου. Μια πρώτη παρατήρηση προς αυτή την κατεύθυνση είναι ότι μπορούμε να φανταστούμε ανθρώπους που έχουν μάθει να καταπιézουν την έκφραση του πόνου σε τέτοιο βαθμό ώστε να μη διαφαίνεται καθόλου στην συμπεριφορά τους. Ωστόσο, η συγκεκριμένη ένσταση μπορεί να αντιμετωπιστεί από τους λογικούς συμπεριφοριστές με την επίκληση της έννοιας της προδιάθεσης για συμπεριφορά. Με άλλα λόγια, κάποιος που καταπιézει την έκφραση του πόνου διαφέρει από κάποιον άλλο που δεν πονά, κατά το ότι έχει την προδιάθεση, αν πληρούνται ορισμένες συνθήκες, να συμπεριφερθεί όπως προσδιορίζει η πρόταση α. Μια δεύτερη σημαντικότερη ένσταση κατά του ισχυρισμού ότι η ύπαρξη συγκεκριμένων συμπεριφορών αποτελεί μια αναγκαία προϋπόθεση για την ύπαρξη πόνου είναι ότι όσο γενικά κι αν προσδιορίσουμε αυτές τις συμπεριφορές μπορούμε πάντα να φανταστούμε κάποιο είδος οργανισμού που όταν πονά παράγει άλλου τύπου συμπεριφορές. Για παράδειγμα, αν θεωρήσουμε ότι αναγκαία προϋπόθεση για να πονά ένας οργανισμός είναι το να βογγά, τότε οδηγούμαστε στο συμπέρασμα ότι ζώα που δεν διαθέτουν φωνητικές χορδές δεν πονούν.

⁴³ Βλ. Kim 1996, σελ. 31.

Άρα, λοιπόν, φαίνεται να μην είναι δυνατό να οριστεί γενικά ο πόνος με όρους συμπεριφοράς.

Το πρόβλημα της δυνατότητας μετάφρασης ψυχολογικών προτάσεων σε προτάσεις που δεν περιλαμβάνουν ψυχολογικούς όρους γίνεται ακόμα μεγαλύτερο αν φύγουμε από προτάσεις που εκφράζουν πόνο και εξετάσουμε προτάσεις που αναφέρονται σε πιο εκλεπτυσμένα και διαμεσολαβημένα νοητικά φαινόμενα όπως, για παράδειγμα, οι πεποιθήσεις, οι προθέσεις, οι επιθυμίες κ.λπ. Ο λογικός συμπεριφορισμός στην πραγματικότητα ποτέ δεν προχώρησε, πέρα από στοιχειώδη παραδείγματα, στην ανάλυση των νοητικών φαινομένων με όρους συμπεριφοράς και προδιαθέσεων για συμπεριφορά. Υπάρχει, ίσως, ένας συστηματικός λόγος που αυτό δεν κατέστη δυνατό στην περίπτωση των πιο σύνθετων νοητικών φαινομένων. Ας πάρουμε μια επιθυμία, έστω την επιθυμία να πάω στη κουζίνα να πω νερό. Ποιες συμπεριφορές θα μπορούσαν να ισοδυναμούν με αυτήν την επιθυμία; Προφανώς, το να σηκωθεί από τη καρέκλα και να πάω στη κουζίνα να πω νερό. Όμως, τα πράγματα δεν είναι τόσο απλά, γιατί αν, για παράδειγμα, πιστεύω ότι στην κουζίνα είναι κλεισμένο ένα επικίνδυνο ζώο, τότε σίγουρα δεν θα επιδείξω αυτή τη συμπεριφορά. Προφανώς, μπορούμε, δυνάμει, να επινοήσουμε μια απειρία διαφορετικών περιπτώσεων εισάγοντας πρόσθετες πεποιθήσεις, επιθυμίες, προθέσεις κ.λπ. οι οποίες κάθε φορά παράγουν ένα διαφορετικό είδος συμπεριφοράς. Το γενικότερο συμπέρασμα που προκύπτει είναι ότι δεν είναι δυνατόν να αντιστοιχίσουμε συμπεριφορές σε μεμονωμένες πεποιθήσεις ή επιθυμίες. Με άλλα λόγια, κάθε συμπεριφορά εξαρτάται από το σύνολο των εκάστοτε πεποιθήσεων, επιθυμιών, προθέσεων κ.λπ. του υποκειμένου.

1.4.2 Η Θεωρία Ταυτότητας

Η αποτυχία του λογικού συμπεριφορισμού οδήγησε στην αναζήτηση μιας διαφορετικής θεώρησης για τα νοητικά φαινόμενα. Η αδυναμία μετάφρασης των ψυχολογικών προτάσεων σε προτάσεις που αναφέρονται σε συμπεριφορές έστρεψε και πάλι το ενδιαφέρον στο εσωτερικό του υποκειμένου. Ωστόσο, δεν πρόκειται για μια στροφή που υιοθετεί εκ νέου τη *res cogitans* του Καρτέσιου και τη δυστική οντολογία, αλλά για μια στροφή που επιτάσσει την αναζήτηση των νοητικών φαινομένων στο εσωτερικό του υποκειμένου με την κυριολεκτική έννοια. Η θεωρία που επικρατεί αυτή την περίοδο - η θεωρία ταυτότητας (*Identity Theory*)- είναι φυσικαλιστική (*physicalist*) και πρεσβεύει ότι οι νοητικές ιδιότητες ταυτίζονται με νευροφυσιολογικές ιδιότητες⁴⁴. Τα νοητικά φαινόμενα δεν είναι τίποτε άλλο παρά εγκεφαλικά φαινόμενα. Μια τέτοια πρόταση, εκ

⁴⁴ Οι βασικοί υποστηρικτές της θεωρίας ταυτότητας είναι ο Smart (1959/1990), ο Feigl (1958), και ο Place (1956/1980).

πρώτης όψεως, φαίνεται αρκετά εύλογη αν σκεφτούμε πόσο δραστικά οι εγκεφαλικές βλάβες επηρεάζουν τις νοητικές ικανότητες των ασθενών.

Με τη θεωρία ταυτότητας, εκτός από τη μετατόπιση από το πεδίο της συμπεριφοράς στο πεδίο των νευροφυσιολογικών φαινομένων, συμβαίνει και μια δεύτερη μετατόπιση από το χώρο της φιλοσοφίας της γλώσσας στον χώρο της μεταφυσικής. Η προτεινόμενη ταυτότητα αφορά πρωταρχικά τις νοητικές και τις νευροφυσιολογικές ιδιότητες και όχι τις σημασίες όρων που αναφέρονται σε νοητικές και νευροφυσιολογικές ιδιότητες. Πιο συγκεκριμένα, οι υποστηρικτές της θεωρίας ταυτότητας δεν ισχυρίζονται, κατ' αναλογία προς τους λογικούς συμπεριφοριστές, ότι προτάσεις που αναφέρονται σε νοητικές ιδιότητες μπορούν να οριστούν (εννοείται, a priori) χωρίς απώλεια περιεχομένου με τη βοήθεια προτάσεων που αναφέρονται σε νευροφυσιολογικές ιδιότητες. Η ταυτότητα μεταξύ νοητικών και νευροφυσιολογικών ιδιοτήτων δεν είναι a priori αλλά εμπειρική. Όπως χαρακτηριστικά επισημαίνει ο Place (1956/1990, σελ. 30) "... η πρόταση 'η συνείδηση είναι μια διαδικασία στον εγκέφαλο', παρότι δεν είναι αναγκαία αληθής, δεν είναι ούτε αναγκαία ψευδής. [Η πρόταση αυτή] κατά την άποψή μου, δεν είναι, από μόνη της, ούτε αντιφατική ούτε προφανής: είναι μια εύλογη επιστημονική υπόθεση, κατά τον τρόπο που η πρόταση 'Η αστραπή είναι μια κίνηση ηλεκτρικών φορτίων' είναι μια εύλογη επιστημονική υπόθεση". Οι υποστηρικτές της θεωρίας ταυτότητας δεν προτείνουν την ταυτότητα μεταξύ συγκεκριμένων νοητικών ιδιοτήτων με συγκεκριμένες φυσικές ιδιότητες του εγκεφάλου διότι κάτι τέτοιο αποτελεί έργο της επιστημονικής έρευνας. Αυτό που προτείνουν είναι ένα τρόπο ανάγνωσης των ευρημάτων της νευροφυσιολογικής έρευνας: η αναφορά σε νοητικές ιδιότητες συνιστά αναφορά σε νευροφυσιολογικές ιδιότητες⁴⁵.

Επειδή ακριβώς η προτεινόμενη ταυτότητα δεν είναι a priori αλλά εμπειρική, δεν ισχύει για την θεωρία ταυτότητας η ένσταση που συναντήσαμε στην περίπτωση του λογικού συμπεριφορισμού, ότι, δηλαδή, η γνώση της έννοιας του πόνου δεν προϋποθέτει τη γνώση νευροφυσιολογικών συμβάντων που λαμβάνουν χώρα όταν κανείς πονά. Όπως στην περίπτωση της αστραπής για πάρα πολλά χρόνια δεν γνωρίζαμε ότι ταυτίζεται ουσιαστικά με μια κίνηση ηλεκτρικών φορτίων, ενδέχεται κάποια στιγμή να ανακαλύψουμε ότι ο πόνος ταυτίζεται με την ενεργοποίηση συγκεκριμένων εγκεφαλικών ινών (έστω, για παράδειγμα, των ινών C⁴⁶). Οι έννοιες του πόνου και της ενεργοποίησης των ινών C, παρότι αναφέρονται στην ίδια ιδιότητα, είναι ανεξάρτητες και διακριτές μεταξύ τους. Η ταυτότητα εδώ ανακαλύπτεται εκ των υστέρων και όχι μέσω εννοιολογικής ανάλυσης γιατί οι όροι "πόνος" και "ενεργοποίηση των ινών C" δεν είναι καταρχάς συνώνυμοι.

⁴⁵ Βλ. Heil 1998, σελ. 78.

⁴⁶ Ισχυρισμός που, σύμφωνα με τους νευροφυσιολόγους, δεν ισχύει.

Η ταύτιση των νοητικών καταστάσεων με νευροφυσιολογικές καταστάσεις οδηγεί σε μια κατανόηση των πρώτων που είναι πολύ στενά δεσμευμένη με τη νευροφυσιολογική δομή. Για παράδειγμα, αν ο πόνος δεν είναι τίποτε άλλο από ενεργοποίηση των ινών C στον εγκέφαλο, τότε είμαστε ανεγκασμένοι να δεχθούμε ότι ζώα που δεν διαθέτουν τέτοιου είδους ίνες δεν πονούν. Ωστόσο, μια τέτοια συνέπεια είναι απαράδεκτη γιατί η έννοια του πόνου χρησιμοποιείται και για ζώα που δεν διαθέτουν ίνες C στον εγκέφαλό τους. Γενικότερα, αν δεχθούμε τον ανθρωποκεντρικό "σβινισμό"⁴⁷ της θεωρίας ταυτότητας, αποφασίζουμε εκ των προτέρων ότι δεν δικαιούμαστε να αποδίδουμε νοητικές καταστάσεις σε οργανισμούς με διαφορετική υλική σύσταση από εμάς, όπως, για παράδειγμα, φίδια, εξωγήινους και ρομποτό.

1.4.3 Ο Λειτουργισμός

Για την αποφυγή αυτής της δυσμενούς συνέπειας αλλά και, ταυτόχρονα, τη διατήρηση των πλεονεκτημάτων της θεωρίας ταυτότητας αναπτύσσεται στις αρχές της δεκαετίας του 1960 μια νέα προσέγγιση η οποία θα αποδειχθεί τόσο επιτυχής ώστε, με την μία ή την άλλη μορφή της, αποτελεί μέχρι σήμερα την κυρίαρχη προσέγγιση για τη σχέση νου - σώματος. Πρόκειται για το λειτουργισμό⁴⁸. Η γενική ιδέα γύρω από την οποία δομείται ο λειτουργισμός είναι ότι οι νοητικές καταστάσεις ταυτίζονται με λειτουργικές καταστάσεις και κατ' αυτόν τον τρόπο καθίστανται ανεξάρτητες από το μέσο υλοποίησης. Συγκεκριμένα, κάθε νοητική κατάσταση ορίζεται με βάση τον αιτιακό της ρόλο, δηλαδή με βάση τον τρόπο που σχετίζεται με τις εισόδους, τις εξόδους και τις άλλες νοητικές καταστάσεις του οργανισμού. Αυτό που είναι σημαντικό στις εγκεφαλικές καταστάσεις δεν είναι η ουσία τους, αλλά η λειτουργία που επιτελούν. Η ίδια λειτουργία μπορεί να επιτελεσθεί και από οντότητες με εντελώς διαφορετική χημική σύσταση. Έτσι, οι νοητικές καταστάσεις δεν ταυτίζονται πλέον ως τύπος (type-identified) με εγκεφαλικές καταστάσεις αλλά με λειτουργικές καταστάσεις. Αυτό δεν εμποδίζει φυσικά να δεχθούμε ότι σε επίπεδο εκτύπου (token) κάθε νοητική κατάσταση ενός συγκεκριμένου υποκειμένου ταυτίζεται με την εκάστοτε εγκεφαλική του κατάσταση. Για παράδειγμα, αν ένα υποκείμενο νιώθει πόνο κατά την χρονική στιγμή t, η συγκεκριμένη νοητική κατάσταση ταυτίζεται με τη συγκεκριμένη εγκεφαλική του κατάσταση την ίδια χρονική στιγμή. Με άλλα λόγια, ο λειτουργισμός, ενώ απορρίπτει την ταυτότητα τύπου μεταξύ

⁴⁷ Βλ. Block 1978/1980.

⁴⁸ Κυριότεροι εμπνευστές του λειτουργισμού είναι ο Putnam 1967/1980 και ο Lewis 1972/1980.

νοητικών και εγκεφαλικών καταστάσεων, είναι συμβατός με την ταυτότητα εκτύπου (token-token identity)⁴⁹ μεταξύ αυτών.

Ωστόσο, η έννοια της λειτουργικής κατάστασης είναι από μόνη της αρκετά ασαφής σε σημείο που επιτρέπει την ανεξέλεγκτη επέκταση των λειτουργιστικών εξηγήσεων ακόμα και σε φαινόμενα που δεν διαθέτουν κάποιο μηχανισμό για την υλοποίηση των υποτιθέμενων λειτουργιών. Οι λειτουργιστικές εξηγήσεις έτσι κινδυνεύουν να γίνουν ψευδοεξηγήσεις. Διόρθωση σε αυτό το πρόβλημα μπορεί να θεωρηθεί ότι προσφέρει μια ειδική μορφή λειτουργισμού - ο λειτουργισμός των μηχανών Turing (Turing-machine functionalism) στον οποίο η έννοια της λειτουργικής κατάστασης συγκεκριμενοποιείται ακόμα περισσότερο και ορίζεται με βάση την κατάσταση μιας καθολικής μηχανής Turing⁵⁰. Με αυτόν τον τρόπο εξασφαλίζεται η ύπαρξη ενός μηχανισμού για την υλοποίηση των λειτουργικών καταστάσεων, δεδομένου ότι οι μηχανές Turing είναι κατασκευάσιμες. Επίσης, για πρώτη φορά με τόση σαφήνεια οι νοητικές καταστάσεις ταυτίζονται με τις λειτουργικές καταστάσεις ενός ψηφιακού υπολογιστή.

1.5 Επίλογος

Η ανάπτυξη του λειτουργισμού παρείχε το ιδανικό πλαίσιο για την σύγκλιση της τεχνητής νοημοσύνης με την γνωσιακή ψυχολογία. Από τη μια μεριά, η θέση ότι οι νοητικές καταστάσεις πρέπει να κατανοηθούν ως μη συνειδητές λειτουργικές καταστάσεις που είναι αυτόνομες ως προς τις εγκεφαλικές καταστάσεις, περιχαράκωνει στο εσωτερικό του υποκειμένου ένα χώρο έρευνας για την γνωσιακή ψυχολογία που είναι σαφέστατα διακριτός τόσο από το χώρο έρευνας της νευροφυσιολογίας, όσο και από εκείνον της ενδοσκοπικής ψυχολογίας. Πρόκειται για το λειτουργικό επίπεδο περιγραφής που αντιδιαστέλλεται τόσο προς το υλικό όσο και προς το φαινομενολογικό επίπεδο. Από την άλλη μεριά, η κατασκευή των ψηφιακών υπολογιστών αποτελούσε την καλύτερη απόδειξη της ορθότητας του λειτουργισμού. Οι μηχανές Turing μπορούν να υλοποιηθούν σε οποιοδήποτε μέσο και οι ψηφιακοί υπολογιστές είναι εκείνες οι μηχανές που μπορούν να αλλάζουν την λειτουργία τους χωρίς να μεταβάλλουν τα μηχανικά τους μέρη.

Το αποτέλεσμα της σύγκλισης τεχνητής νοημοσύνης και γνωσιακής ψυχολογίας είναι η υιοθέτηση της θέσης ότι ο νους αποτελεί έναν ψηφιακό υπολογιστή. Γύρω από αυτή τη θέση αρχίζει να δομείται σιγά-σιγά η γνωσιακή επιστήμη. Στο επόμενο κεφάλαιο θα παρακολουθήσουμε από την σκοπιά της φιλοσοφίας την ανάπτυξη της θέσης ότι ο νους αποτελεί έναν ψηφιακό υπολογιστή. Παράλληλα, θα

⁴⁹ Θα αναφερθούμε αναλυτικά στην ταυτότητα εκτύπου στο όγδοο κεφάλαιο (ενότητα 8.3).

⁵⁰ Βλ. Putnam 1967/1980.

επιχειρήσουμε να αναδείξουμε άλλες βασικές παραδοχές που συνέβαλαν στη διαμόρφωση της παραδοσιακής γνωσιακής επιστήμης.

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ

Η Αναπαραστασιακή-Υπολογιστική Θεωρία του Νου (ΑΥΘΝ)

2.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο εξετάσαμε ορισμένα από τα ιστορικά νήματα που οδήγησαν στην ανάπτυξη της γνωσιακής επιστήμης. Η ανακάλυψη των ψηφιακών υπολογιστών και η ανάπτυξη της τεχνητής νοημοσύνης προσέφεραν ένα νέο μοντέλο για την κατανόηση των νοητικών διαδικασιών ως διαδικασιών επεξεργασίας πληροφοριών. Η γνωσιακή ψυχολογία, απ' τη μεριά της, ήρθε ως αντίδραση στον εξοβελισμό των εσωτερικών διαδικασιών από τον μεθοδολογικό συμπεριφορισμό. Ωστόσο, αυτή η επιστροφή στο τι συμβαίνει εντός του υποκειμένου δεν ήταν, όπως είδαμε, μια επιστροφή στην υποκειμενική σφαίρα που διερευνούσε η ενδοσκοπική ψυχολογία, διότι οι νοητικές διαδικασίες για τις οποίες κάνει λόγο η γνωσιακή ψυχολογία είναι εξ ορισμού μη-συνειδητές. Τέλος, οι εξελίξεις στην φιλοσοφία οδήγησαν σε μια νέα πρόταση για τη φύση του νου. Τα νοητικά φαινόμενα κατανοήθηκαν ως λειτουργικές καταστάσεις του εγκεφάλου. Σύμφωνα με τον λειτουργισμό, οι νοητικές καταστάσεις περιγράφονται στο λειτουργικό επίπεδο –ένα επίπεδο που διαφοροποιείται τόσο από το νευροφυσιολογικό επίπεδο όσο και από το φαινομενολογικό, δηλαδή το επίπεδο στο οποίο οριζόταν η έρευνα της ενδοσκοπικής ψυχολογίας. Η σύγκλιση TN και γνωσιακής ψυχολογίας μέσα και από τις επεξεργασίες της φιλοσοφίας του νου συνέβαλε ίσως περισσότερο από κάθε τι άλλο στην ανάπτυξη της γνωσιακής επιστήμης.

2.2 Αφετηριακές θέσεις της Γνωσιακής Επιστήμης

Ένα μεγάλο μέρος της σύγχρονης φιλοσοφίας του νου μπορεί να θεωρηθεί ως φιλοσοφία της γνωσιακής επιστήμης. Βασική αποστολή της, υπό αυτή την έννοια, είναι η διασαφήνιση του μεταφυσικού πλαισίου της γνωσιακής επιστήμης. Το πλαίσιο αυτό έχει ονομαστεί Αναπαραστασιακή - Υπολογιστική Θεωρία του Νου (AYΘN). Η AYΘN υιοθετεί τα πορίσματα του λειτουργισμού αλλά ταυτόχρονα προχωρεί πέρα από αυτόν. Όπως θα δούμε, μια βασική καινοτομία της είναι ότι εκτός από το μοντέλο του υπολογιστή επιστρατεύει και το μοντέλο της γλώσσας προκειμένου να εξηγήσει τη νόηση. Ωστόσο, θα δείξουμε ότι η αντίληψη περί γλώσσας που υιοθετεί η AYΘN έχει περισσότερο να κάνει με τις διάφορες τυπικές γλώσσες και λιγότερο με τη φυσική γλώσσα.

Ο συνδυασμός του μοντέλου του υπολογιστή με το μοντέλο της γλώσσας οδηγεί, σε μια πρώτη προσέγγιση, στην ακόλουθη κεντρική θέση της AYΘN: η νόηση συνίσταται στον υπολογιστικό μετασχηματισμό νοητικών αναπαραστάσεων γλωσσικής δομής. Έτσι, το ζητούμενο σύμφωνα με αυτή την προσέγγιση, είναι να δοθεί μια εξήγηση της νόησης με αφετηρία τις έννοιες της νοητικής αναπαράστασης και του υπολογισμού. Ωστόσο, οποιαδήποτε μορφή και αν πάρει τελικά αυτή η εξήγηση θα πρέπει να παραμείνει μέσα στο φυσιοκρατικό πλαίσιο. Η πίστη στην ύπαρξη μιας φυσιοκρατικής (naturalistic) εξήγησης του νου,

ό,τι και αν σημαίνει αυτό, αποτελεί μια αυτονόητη δέσμευση και αφετηρία όσων υποστηρίζουν την ΑΥΘΝ⁵¹.

Όσο εύκολο είναι να μιλήσει κανείς γενικά για την ΑΥΘΝ, τόσο δύσκολο είναι να την αναπτύξει ως μια συνεκτική θεωρία. Και αυτό γιατί υπάρχουν πολλές διαφορετικές εκδοχές της ΑΥΘΝ, αλλά και πολλές διαμάχες για επιμέρους ζητήματα. Η ανασυγκρότηση που ακολουθεί θα βασισθεί κυρίως στις θέσεις ενός φιλοσόφου, του Jerry Fodor, διότι, κατά γενική ομολογία, αποτελεί τον πιο συνεπή και επίμονο υποστηρικτή της ΑΥΘΝ. Έτσι, θα αποφευχθεί η κατασκευή ενός υβριδίου που δεν υποστηρίζεται από κανέναν. Βέβαια η επιλογή να ανασυγκροτήσουμε την ΑΥΘΝ μέσα από τις θέσεις του Fodor δημιουργεί ένα νέο πρόβλημα που έχει να κάνει με το αν και κατά πόσο θα πρέπει να παρουσιάσουμε τις, αρκετές είναι αλήθεια, μεταβολές που έχουν υποστεί οι απόψεις του τα τελευταία τριάντα χρόνια. Ωστόσο, δεν σκοπεύουμε να παρακολουθήσουμε στενά όλες αυτές τις μετατοπίσεις, ούτε να εμβαθύνουμε σε πτυχές του έργου του Fodor που δεν έτυχαν ευρύτερης αποδοχής. Ο στόχος του κεφαλαίου δεν είναι να αποτελέσει μια σχολαστική μελέτη του έργου του Fodor, αλλά να συστηματοποιήσει εκείνες από τις θέσεις του που άσκησαν μεγαλύτερη επιρροή για την συγκρότηση της γνωσιακής επιστήμης.

Ο Fodor είναι ουσιαστικά ένας από τους θεμελιωτές της γνωσιακής επιστήμης. Το βιβλίο του *Η Γλώσσα της Σκέψης* έπαιξε σημαντικό ρόλο στη συγκρότηση του μεταφυσικού πλαισίου αυτής της νέας επιστήμης. Η υπόθεση της γλώσσας της σκέψης αποτέλεσε τη βάση για την ανάπτυξη της ΑΥΘΝ.

Ο στόχος του Fodor άλλα και των υπολοίπων υποστηρικτών της ΑΥΘΝ είναι, όπως είπαμε, να δώσουν μια φυσιοκρατική εξήγηση της νόησης και γενικότερα των νοητικών φαινομένων. Για ποια νοητικά φαινόμενα, όμως, γίνεται λόγος; Ο Fodor ξεκινά από τη θέση ότι η ταξινόμηση των νοητικών φαινομένων όπως αυτή εκφράζεται μέσα στη

⁵¹ Στο έβδομο κεφάλαιο θα εξετάσουμε αναλυτικά τις κυριότερες προσπάθειες φυσικοποίησης του περιεχομένου των νοητικών αναπαραστάσεων. Έτσι, η έκφραση "φυσιοκρατική εξήγηση" θα αποκτήσει πιο συγκεκριμένο περιεχόμενο. Εντελώς γενικά θα μπορούσαμε να πούμε ότι η φυσιοκρατία "ως προς μια επικράτεια όντων (realm) είναι η άποψη ότι οτιδήποτε υπάρχει εντός αυτής της επικράτειας και όλα τα συμβάντα που λαμβάνουν χώρα εντός αυτής είναι εμπειρικά προσβάσιμα χαρακτηριστικά του κόσμου" (Guttenplan 1994, σελ. 449). Η φυσιοκρατία πολλές φορές συνδέεται με την αναγωγή των όντων και των συμβάντων μιας επικράτειας στα όντα και τα συμβάντα της επικράτειας της επιστήμης της φυσικής. Σε αυτή την περίπτωση είναι ακριβέστερο να μιλάμε για "φυσικαλισμό" (physicalism) ή υλισμό (materialism). Η φυσιοκρατία είναι συμβατή με τη θέση ότι το φυσικό (natural) δεν εξαντλείται από τα όντα και τα συμβάντα που μελετά η επιστήμη της φυσικής.

ψυχολογία του κοινού νου (folk psychology)⁵² είναι κατά βάση ορθή και γι' αυτό θα πρέπει να αποτελέσει το θεμέλιο για μια επιστημονική ψυχολογία. Είναι ορθό, με άλλα λόγια, το ότι για να εξηγήσουμε και να προβλέψουμε την ανθρώπινη συμπεριφορά ανατρέχουμε στις πεποιθήσεις, τις επιθυμίες, τους φόβους κλπ των δρώντων υποκειμένων. Αυτές οι "νοητικές καταστάσεις", σύμφωνα με τη ψυχολογία του κοινού νου, διαθέτουν σημασιολογικές αλλά και αιτιακές ιδιότητες. Για παράδειγμα, μια πεποίθηση μπορεί να είναι αληθής ή ψευδής, αλλά ταυτόχρονα μπορεί να αποτελεί και την αιτία κάποιας πράξης του υποκειμένου⁵³. Έτσι, τα νοητικά φαινόμενα για τα οποία ο Fodor ενδιαφέρεται να δώσει μια φυσιοκρατική εξήγηση είναι οι πεποιθήσεις, οι επιθυμίες, οι φόβοι και οι υπόλοιπες καταστάσεις του υποκειμένου που διαθέτουν σημασιολογικές και αιτιακές⁵⁴ ιδιότητες (σύμφωνα πάντα με τη ψυχολογία του κοινού νου). Τα φαινόμενα αυτά, μέσα στην αναλυτική φιλοσοφία, ονομάζονται, για λόγους που θα εξηγήσουμε στην επόμενη ενότητα, 'προτασιακές στάσεις' (propositional attitudes).

Ο Fodor πιστεύει ότι οι προτασιακές στάσεις παίζουν κεντρικό ρόλο στην εξήγηση της ανθρώπινης συμπεριφοράς και αυτό είναι κάτι που θέλει να το διατηρήσει στις εξηγήσεις που δίνει η γνωσιακή επιστήμη. Έτσι, ξεκινά τη θεωρία του προϋποθέτοντας την πραγματικότητα των προτασιακών στάσεων (με άλλα λόγια είναι αποβλεπτικός⁵⁵ ρεαλιστής) και το ερώτημα που θέτει αφορά το πώς είναι δυνατή η ύπαρξή τους μέσα στον φυσιοκρατικό κόσμο. Ωστόσο, αυτή η αποδοχή των προτασιακών στάσεων όπως εμφανίζονται στη ψυχολογία του κοινού νου, παρά τα φαινόμενα, δεν είναι καθολική: ο Fodor παραβλέπει εντελώς το ότι αυτές είναι, ως επί το πλείστον, *συνειδητές*. Το συγκεκριμένο χαρακτηριστικό των προτασιακών στάσεων δεν παίζει κανένα ρόλο στον τρόπο που αναπτύσσει την ΑΥΘΝ. Η συνειδηση εκλαμβάνεται ουσιαστικά ως ένα επαφαινόμενο που ούτε προσθέτει ούτε αφαιρεί τίποτα από τις σημασιολογικές και αιτιακές ιδιότητες των προτασιακών στάσεων. Ο

⁵² "Υποθέτω ότι αυτό που θέλουμε - και αυτό που μια επιτυχημένη γνωσιακή επιστήμη οφείλει να μας δώσει - είναι μια ψυχολογία των προτασιακών στάσεων [δηλαδή, των πεποιθήσεων, των επιθυμιών κ.λπ. (βλ. ενότητα 2.3.2)]. . . . Αυτό σημαίνει: οι προτασιακές στάσεις αποτελούν το κύριο αντικείμενο της θεωρίας: η θεωρία εμφανίζεται να μας λέει τι είναι οι προτασιακές στάσεις και τι κάνουν", Fodor 1985b

⁵³ Μάλιστα όπως έχει δείξει η ψυχανάλυση, με την υιοθέτηση ασυνειδητών πεποιθήσεων και επιθυμιών μπορούμε να επεκτείνουμε το εύρος των εξηγούμενων φαινομένων εντάσσοντας σε αυτά ένα μεγάλο αριθμό ανώμαλων συμπεριφορών (Loewer & Rey 1991, σελ. 277-78)

⁵⁴ Στη ψυχολογία του κοινού νου, οι πεποιθήσεις, οι επιθυμίες, οι φόβοι κλπ συνδέονται αιτιακά με δύο επιπλέον τρόπους: προκαλούν άλλες νοητικές καταστάσεις και προκαλούνται από αισθητηριακά ερεθίσματα.

⁵⁵ Βλ. επόμενη ενότητα.

Fodor ακολουθεί εδώ το γενικότερο κλίμα της φυσικοκρατικής παράδοσης στην οποία ανήκει και η οποία, προκειμένου να αποφυγεί το φάντασμα του καρτεσιανισμού και τα αδιέξοδα της ενδοσκοπικής ψυχολογίας, παραβλέπει εντελώς το φαινόμενο της συνείδησης (Searle 1992). Αυτή η στρατηγική οδηγεί σε σημαντικά αδιέξοδα όπως θα δείξουμε στο όγδοο κεφάλαιο, όπου θα ασχοληθούμε συστηματικά με το πρόβλημα της συνείδησης και το ρόλο που αυτή παίζει στην συγκρότηση των νοητικών καταστάσεων. Στο παρόν κεφάλαιο θα αρκεστούμε απλώς στην επισήμανση του φαινομένου της συνείδησης κάθε φορά που αυτό διαφαίνεται σε ορισμένα προβληματικά σημεία της ΑΥΘΝ.

Συνοψίζοντας, θα μπορούσαμε να πούμε ότι για τον Fodor το πρόβλημα της φυσικοποίησης των νοητικών φαινομένων ανάγεται στο πρόβλημα της φυσικοποίησης των προτασιακών στάσεων, δηλαδή εκείνων από τις νοητικές καταστάσεις που έχουν σημασιολογικές και αιτιακές ιδιότητες. Αυτό το δεύτερο πρόβλημα είναι που έρχεται να επιλύσει η ΑΥΘΝ με την υπόθεση της γλώσσας της σκέψης. Προτού αναφερθούμε στην θεωρία του Fodor θα πρέπει να δούμε υπό ποια μορφή παραλαμβάνει ο ίδιος το πρόβλημα, των προτασιακών στάσεων μέσα από την αναλυτική φιλοσοφία.

2.3 *Ιστορική τοποθέτηση του προβλήματος της φυσικοποίησης της νόησης στο πλαίσιο της φιλοσοφίας του 20ου αιώνα*

2.3.1 Το πρόβλημα της αποβλεπτικότητας

Το πρόβλημα των προτασιακών στάσεων αποτελεί ουσιαστικά μια επαναδιατύπωση και ερμηνεία του προβλήματος της αποβλεπτικότητας⁵⁶ (γερμανικά: Intentionalität, αγγλικά: intentionality), του λεγόμενου "προβλήματος του Brentano". Η αποβλεπτικότητα θεωρήθηκε ως το μη αναγώγιμο χαρακτηριστικό του νοητικού, εκείνη η ιδιότητα που διαφοροποιεί το νοητικό από το φυσικό. Ο Brentano, ο οποίος ήταν ο πρώτος που επανέφερε το πρόβλημα της αποβλεπτικότητας στην νεότερη φιλοσοφία, μίλησε για ένα περίεργο είδος σχέσης, μια "αποβλεπτική σχέση", η οποία συνίσταται στην ιδιότητα κάθε νοητικού φαινομένου να κατευθύνεται προς κάτι εκτός αυτού - υπάρχον ή μη. Κάθε συνείδηση είναι συνείδηση κάποιου πράγματος. Όταν φέρνω στον νου μου έναν κένταυρο, σκέφτομαι κάτι που δεν υπάρχει, ωστόσο αυτό δεν είναι το ίδιο με το να μη σκέφτομαι τίποτα. Ο Brentano, ακολουθώντας τους σχολαστικούς του μεσαίωνα, ονόμασε τον τρόπο ύπαρξης των αντικειμένων της σκέψης "αποβλεπτική ενύπαρξη". Επισημαίνει χαρακτηριστικά:

⁵⁶ Ακολουθώ την μεταφραστική επιλογή του Καθηγητή Ν. Σκουτερόπουλου (βλ. Χούσσερλ 1986)

"Αυτή η αποβλεπτική ενύπαρξη είναι αποκλειστικό χαρακτηριστικό των νοητικών φαινομένων. Κανένα φυσικό φαινόμενο δεν εκδηλώνει κάτι παρόμοιο. Συνεπώς, μπορούμε να ορίσουμε τα νοητικά φαινόμενα λέγοντας ότι είναι εκείνα τα φαινόμενα που περιλαμβάνουν εντός τους με αποβλεπτικό τρόπο ένα αντικείμενο" (Brentano 1874/1973)

Από τη συνοπτική αυτή παρουσίαση της έννοιας της αποβλεπτικότητας καθίσταται φανερό ότι δεν έχει κάποια ιδιαίτερη σχέση με τις ανθρώπινες προθέσεις, όπως εσφαλμένα υποδηλώνει η ρίζα του γερμανικού ή του αγγλικού όρου. Αποβλεπτικότητα δεν έχουν μόνο οι προθέσεις αλλά και οι επιθυμίες, οι πεποιθήσεις, καθώς και οι υπόλοιπες προτασιακές στάσεις. Κατά συνέπεια, ο όρος "προθετικότητα" είναι εξίσου παραπλανητικός με τον αγγλικό ή τον γερμανικό όρο. Επίσης, η έννοια της αποβλεπτικότητας δεν θα πρέπει να συγχέεται με την έννοια της εντασιακότητας (intensionality με s). Η εντασιακότητα ορίζεται κατ' αντιδιαστολή προς την εκτασιακότητα. Εντασιακή θεωρείται η αναφορά σε ιδιότητες και σχέσεις και εκτασιακή η αναφορά στις οντότητες που πληρούν κάποιες ιδιότητες ή σχέσεις. Για παράδειγμα, η ιδιότητα της φρόνησης θα πρέπει να διακριθεί από το σύνολο των φρόνιμων ανθρώπων. Ο λόγος που υπαγορεύει αυτή τη διάκριση είναι ότι δύο διαφορετικές ιδιότητες μπορεί να έχουν την ίδια έκταση.

2.3.2 Το πρόβλημα των προτασιακών στάσεων

Η αναλυτική φιλοσοφία παραλαμβάνει την προβληματική περί αποβλεπτικότητας και κάνοντας ορισμένες μεταποίσεις διατυπώνει το πρόβλημα των προτασιακών στάσεων. Μια βασική μετατόπιση είναι ότι τα νοητικά φαινόμενα δεν κατευθύνονται προς ένα μεμονωμένο αντικείμενο αλλά προς μια κατάσταση πραγμάτων, ή με την ορολογία του Russell, προς μια πρόταση (proposition).

Τα ρήματα που αναφέρονται σε νοητικές καταστάσεις συντάσσονται (ή μπορούν, τελικά, να συνταχθούν) με προτασιακά αντικείμενα: πιστεύω ότι ο Fodor είναι επιστήμονας, φοβάμαι ότι ο Fodor είναι επιστήμονας, ελπίζω ότι ο Fodor είναι επιστήμονας κ.λπ. Με άλλα λόγια, έχουν τη γενική δομή: πιστεύω, φοβάμαι, ελπίζω κ.λπ. ότι p, όπου p εκφράζει μια ενεργεία υπαρκτή (actual) ή δυνατή κατάσταση πραγμάτων, μια πρόταση. Οι προτάσεις (propositions) δεν θα πρέπει να συγχέονται με τις γλωσσικές προτάσεις (sentences). Για παράδειγμα, οι γλωσσικές προτάσεις "Ο Fodor είναι επιστήμονας" και "Fodor is a scientist" παρότι αποτελούνται από διαφορετικά σύμβολα εκφράζουν την ίδια πρόταση, την ίδια κατάσταση πραγμάτων⁵⁷. Όταν πιστεύω, ελπίζω, φοβάμαι κ.λπ. ότι p, υιοθετώ μια

⁵⁷ Στα επόμενα, για την αποφυγή σύγχυσης, θα μεταφράζω το proposition ως πρόταση και το sentence ως γλωσσική πρόταση. Επίσης, χάριν συντομίας, αντί του

στάση έναντι του *p*. Αυτός είναι ο λόγος για τον οποίο στην αναλυτική φιλοσοφία οι νοητικές καταστάσεις ονομάζονται, όπως αναφέραμε στη προηγούμενη ενότητα, προτασιακές στάσεις³⁸.

2.4 Η Αναπαραστασιακή Θεωρία του Νου (ΑΘΝ)

2.4.1 Η υπόθεση της γλώσσας της σκέψης

Η μορφή στην οποία παραλαμβάνει ο Fodor το πρόβλημα των προτασιακών στάσεων συνίσταται, κατά συνέπεια, στο ερώτημα πώς το υποκείμενο σχετίζεται με τις προτάσεις. Ο Fodor ευθύς εξαρχής δηλώνει ότι αδυνατεί να καταλάβει πώς ένας οργανισμός μπορεί να σχετίζεται άμεσα με καταστάσεις πραγμάτων. Για τον λόγο αυτό, κατηγορεί τον Frege ότι δεν έχει πει τίποτα για αυτή τη σχέση πέρα από κάποιους αρνητικούς προσδιορισμούς ότι, δηλαδή, οι προτάσεις δεν μπορούν να δοθούν ούτε στην αισθητηριακή αντίληψη (γιατί είναι αφηρημένες οντότητες) ούτε και στην ενδοσκοπήση (γιατί δεν είναι υποκειμενικές παραστάσεις [Vorstellungen]) (Fodor 1981a, σελ. 202). "Κάτι πρέπει να ειπωθεί για τη θέση του σημασιολογικού και του αποβλεπτικού μέσα στη φυσική τάξη δεν θα μας χρησιμεύσει το να έχουμε ανεξήγητες 'σχέσεις με προτάσεις' στα θεμέλια της φιλοσοφίας του νου" (Fodor 1985a, σελ 12). Για τον Fodor *κατανόηση* της σχέσης οργανισμού - κατάστασης πραγμάτων σημαίνει αναγωγή της σε μια φυσικοκρατική περιγραφή. Έτσι, στόχος του, ούτε λίγο ούτε πολύ, γίνεται η φυσικοποίηση των προτασιακών στάσεων, η φυσικοποίηση της αποβλεπτικότητας. Ο H. Field στο σημαντικό άρθρο του "Mental Representation" συνοψίζει ακριβώς αυτό το αίτημα:

"Το . . . πρόβλημα που τέθηκε από τον Brentano είναι το πρόβλημα της αποβλεπτικότητας. Πολλές νοητικές ιδιότητες - το πιστεύει, το επιθυμείν κ.ο.κ. - φαίνονται να είναι σχεσιακές ιδιότητες: ακριβέστερα, φαίνονται να σχετίζουν ανθρώπους με μη γλωσσικές οντότητες, τις ονομαζόμενες καταστάσεις πραγμάτων (propositions). Έτσι, κάθε υλιστής που παίρνει σοβαρά το πιστεύειν και το επιθυμείν -κάθε υλιστής που δέχεται ότι η πεποιθήσεις και οι επιθυμίες είναι σχέσεις μεταξύ ανθρώπων και καταστάσεων πραγμάτων - κάθε τέτοιος υλιστής

όρου "ενεργεία υπαρκτή ή δυνατή κατάσταση πραγμάτων" θα χρησιμοποίη τον όρο "κατάσταση πραγμάτων".

³⁸ Αξίζει να επισημάνουμε εδώ ότι η παραπάνω ανάλυση προϋποθέτει την παραδοχή ότι η δομή της γλώσσας εκφράζει τη δομή της σκέψης, πιο συγκεκριμένα, ότι ο τρόπος με τον οποίο μιλάμε για τις νοητικές καταστάσεις αντανακλά την οντολογία τους (βλ. ενότητα 2.4.2).

πρέπει να δείξει ότι οι εν λόγω σχέσεις δεν είναι μη αναγώγιμα νοητικές." (Field 1981, σελ. 78)

Διατυπωμένο σ' αυτή τη μορφή, το πρόβλημα της αποβλεπτικότητας, ως πρόβλημα των προτασιακών στάσεων, μοιάζει εξίσου απροσπέλαστο. Πώς μπορεί να εξηγηθεί με ένα φυσιοκρατικό τρόπο η σχέση οργανισμού - κατάστασης πραγμάτων, τη στιγμή μάλιστα που η έννοια της κατάστασης πραγμάτων / πρότασης φαίνεται να είναι περισσότερο ασαφής οντολογικά απ' ό,τι η έννοια του αντικειμένου;

Μια λύση θα ήταν να επιστρατεύσει κανείς το οπλοστάσιο του λειτουργισμού. Όπως είδαμε, οι λειτουργιστές μιλούσαν γενικώς για νοητικές καταστάσεις χωρίς να διευκρινίζουν εάν πρόκειται για προτασιακές στάσεις, αισθήματα, συναισθήματα, διαθέσεις κτλ. Αν λοιπόν η θεωρία τους ισχύει για τις νοητικές καταστάσεις εν γένει, θα πρέπει να ισχύει και για τις προτασιακές στάσεις. Ορισμένοι ερευνητές ακολούθησαν πράγματι αυτόν το δρόμο⁵⁹. Ωστόσο, ο Fodor θεώρησε ότι για την εξήγηση της αποβλεπτικότητας των προτασιακών στάσεων δεν επαρκεί ο λειτουργισμός⁶⁰. Συγκεκριμένα, η άποψή του ήταν ότι οι λειτουργιστικές θεωρήσεις επιτυγχάνουν να εξηγήσουν απλώς το πώς νοήμονες (intelligent) συμπεριφορές και διαδικασίες μπορούν να συμβαίνουν σε ένα φυσιοκρατικό σύμπαν. Το μοντέλο του υπολογιστή προσφέρει μια σχετικά απλή απάντηση. Μια "νοήμων" λειτουργία μπορεί να αναλυθεί σε λιγότερο νοήμονες λειτουργίες και αυτές με τη σειρά τους σε ακόμα απλούστερες κ.ο.κ μέχρι τελικά να καταλήξουμε σε στοιχειώδεις λειτουργίες που είναι εντελώς "χαζές" (όπως, για παράδειγμα, οι λειτουργίες μιας μηχανής Turing) και, κατά συνέπεια, άμεσα πραγματοποιήσιμες με μια φυσική διαδικασία. Ωστόσο, σύμφωνα με τον Fodor, ακόμα και αυτές οι στοιχειώδεις λειτουργίες στο βαθμό που παίζουν κάποιο ρόλο στις ψυχολογικές εξηγήσεις οφείλουν να είναι σημασιολογικά ερμηνεύσιμες, δηλαδή να μπορεί να τους αποδοθεί περιεχόμενο.

Ο Fodor προσπαθεί εδώ να κάνει μια διάκριση μεταξύ νοημοσύνης και αποβλεπτικότητας, για να υποστηρίξει ότι η δεύτερη δεν μπορεί να εξηγηθεί με βάση την πρώτη. Ωστόσο, δεν καταφέρνει να δώσει κάποιο πειστικό επιχειρήμα που να καθιστά σαφή την προτεινόμενη διάκριση. Ένας τρόπος που θα μπορούσαμε να καταλάβουμε το σε τι στοχεύει ο Fodor είναι ο ακόλουθος: οι λειτουργιστικές εξηγήσεις δεν αφορούν με κάποιο προνομιακό τρόπο τις νοητικές καταστάσεις. Λειτουργιστικές εξηγήσεις μπορούμε να δώσουμε για πράγματα που δεν έχουν τίποτα το νοητικό. Για παράδειγμα, μπορούμε να εξηγήσουμε το πώς κινείται ένα αυτοκίνητο περιγράφοντας τον τρόπο λειτουργίας του. Δεν υπάρχει τίποτα

⁵⁹ Dennett 1987, Loar 1981 κ.α.

⁶⁰ Βλ. Pylyshyn 1984, κεφ. 2

το αποβλεπτικό στην συμπεριφορά του αυτοκινήτου. Ενδεχομένως, θα μπορούσαμε να πούμε ότι υπάρχει κάτι νοήμον (intelligent), με την έννοια του ευφυούς, δηλαδή ότι επιτελείται μια σύνθετη λειτουργία που προσαρμόζεται σε κάποιους εξωτερικούς περιορισμούς. Αν είναι δύσκολο να αποκαλέσουμε "ευφύες", με αυτή την έννοια, ένα αυτοκίνητο, οι αναστολές μας είναι πολύ μικρότερες όταν πρόκειται για κάποιο σύστημα αυτόματου ελέγχου, όπως για παράδειγμα ο ρυθμιστής του Watt⁶¹. Το σίγουρο, ωστόσο, είναι ότι προθεωρητικά δεν θα αποδίδαμε *αποβλεπτικότητα* στον ρυθμιστή του Watt. Έτσι, λοιπόν, αν οι λειτουργιστικές εξηγήσεις πρόκειται να χρησιμοποιηθούν για την εξήγηση των νοητικών φαινομένων θα πρέπει να συμπληρωθούν με κάτι επιπλέον που να εξηγεί την αποβλεπτικότητά τους. Ο Fodor θεωρεί ότι αυτό το συμπλήρωμα βρίσκεται στην έννοια του συμβόλου, της αναπαράστασης. Οι υπολογιστές δεν είναι απλώς λειτουργικά περιγραφόμενες μηχανές, αλλά μηχανές χειρισμού *συμβόλων*.

Αλλά για ποιο λόγο ο Fodor θεωρεί ότι βασιζόμενος στην έννοια του συμβόλου θα μπορέσει να επιλύσει το πρόβλημα των προτασιακών στάσεων; Η απάντηση θα πρέπει να αναζητηθεί στην ομοιότητα που υπάρχει ανάμεσα στα σύμβολα και τις προτασιακές στάσεις⁶². Τα σύμβολα είναι τα μόνα, ίσως, από τα φυσικά αντικείμενα του καθημερινού μας βίου που διαθέτουν ταυτόχρονα αιτιακές και σημασιολογικές ιδιότητες, ό,τι, δηλαδή, διαθέτουν και οι προς εξήγηση νοητικές καταστάσεις. Για παράδειγμα, η γραπτή πρόταση "ο βράχος είναι μέσα στο ψυγείο", όντας ένα επιμέρους αντικείμενο έχει ένα πλήθος αιτιακών ιδιοτήτων: αντανακλά το φως, έλκει λόγω της μάζας της διάφορα σωματίδια κτλ. Ταυτόχρονα, έχει σημασιολογικές ιδιότητες: είναι αληθής αν και μόνο αν ο βράχος είναι μέσα στο ψυγείο και σε κάθε άλλη περίπτωση είναι ψευδής. Σε αντίθεση, λοιπόν, προς τους βράχους -οι οποίοι έχουν αιτιακές αλλά όχι σημασιολογικές ιδιότητες- και τις προτάσεις - οι οποίες έχουν σημασιολογικές αλλά όχι και αιτιακές ιδιότητες (καθότι αφηρημένα αντικείμενα) - τα σύμβολα είναι τα μόνα, ίσως, από τα γνωστά φυσικά αντικείμενα που έχουν ταυτόχρονα και τα δύο είδη ιδιοτήτων.

Ο Fodor ξεκινώντας από αυτή την ιδέα προσπαθεί να προσεγγίσει με ένα νέο τρόπο το πρόβλημα της φυσικοποίησης των προτασιακών στάσεων. Η λύση που προτείνει βασίζεται στην εισαγωγή μιας νέας οντότητας - μιας εσωτερικής αναπαράστασης, που μεσολαβεί ανάμεσα στον οργανισμό και την πρόταση. Αυτή η νέα οντότητα είναι μια νοητική αναπαράσταση S που έχει την πρόταση ως περιεχόμενό της.

⁶¹ Πρόκειται για ένα μηχανισμό που διατηρεί σταθερή την ταχύτητα περιστροφής ενός τροχού ρυθμίζοντας κατάλληλα την πίεση του ατμού που αποτελεί την κινητήρια δύναμη για την περιστροφή του τροχού (βλ. van Gelder 1995, σελ. 347). Βλ. παράρτημα 2.

⁶² Βλ. Fodor 1987, σελ. xi.

Κατά συνέπεια, η προτασιακή στάση [ο x πιστεύει ότι p] σπάει σε δύο επιμέρους σχέσεις:

1. ο x έχει μια σχέση R προς την S, όπου S είναι μια νοητική αναπαράσταση εντός του x, και
2. η νοητική αναπαράσταση S σημαίνει ότι p (Fodor 1987, σελ. 17, 135 και Fodor 1975, σελ. 198)

Η εσωτερική αναπαράσταση S πρέπει να νοηθεί εδώ κυριολεκτικά ως μια γλωσσική αναπαράσταση - μια γλωσσική πρόταση "γραμμένη" όχι στη φυσική γλώσσα αλλά στη γλώσσα της σκέψης. Για το λόγο αυτό ο Fodor ονόμασε την υπόθεσή του "Υπόθεση της Γλώσσας της Σκέψης" (ΥΓΣ). Όπως η γραφή προϋποθέτει πάντα κάποιο υλικό μέσο, έτσι ακριβώς και οι νοητικές αναπαραστάσεις, σύμφωνα με τον Fodor, δεν μπορούν παρά να έχουν κάποιο υλικό φορέα. Το γεγονός αυτό εξασφαλίζει ότι οι νοητικές αναπαραστάσεις, όντας επιμέρους (particular) αντικείμενα, εκτός από σημασιολογικές ιδιότητες διαθέτουν και αιτιακές ιδιότητες (καθότι μόνο οι φυσικές καταστάσεις διαθέτουν τέτοιες ιδιότητες). Το σύνθημα του Fodor είναι: "καμία αποβλεπτική αιτιότητα χωρίς ρητή αναπαράσταση" (Fodor 1987, σελ. 25). Έτσι, λοιπόν, οι νοητικές αναπαραστάσεις εκλαμβάνονται ως ένα ιδιαίτερο είδος φυσικών φαινομένων.

Η ΥΓΣ εκτός από μια νέα οντότητα θέτει και μια νέα σχέση -τη σχέση ανάμεσα στον οργανισμό και τη νοητική αναπαράσταση. Η σχέση αυτή, σύμφωνα με τον Fodor, ορίζεται υπολογιστικά και διαφοροποιείται ανάλογα με το είδος της προτασιακής στάσης. Για να δώσουμε ένα παράδειγμα, εάν ο Πέτρος πιστεύει ότι ο Fodor είναι φυσικός αυτό που συμβαίνει, σύμφωνα με τη ΥΓΣ, είναι ότι ο οργανισμός του Πέτρου σχετίζεται με ένα συγκεκριμένο υπολογιστικό τρόπο R_1 με κάποιο δείγμα-έκτυπο (token) νοητικής αναπαράστασης (γλωσσικής πρότασης του νου) το οποίο έχει ως περιεχόμενο ότι ο Fodor είναι φυσικός. Αντίστοιχα, αν ο Πέτρος ελπίζει ότι ο Fodor είναι φυσικός, πάλι ο οργανισμός του σχετίζεται με ένα έκτυπο της ίδιας γλωσσικής πρότασης του νου και η μόνη διαφορά που υπάρχει έγκειται στον υπολογιστικό τρόπο με τον οποίο σχετίζεται με αυτή την εσωτερική γλωσσική πρόταση. Το είδος της υπολογιστικής σύνδεσης καθορίζεται στο επίπεδο της αρχιτεκτονικής του νου από το είδος των μηχανισμών που έχουν πρόσβαση σε μια πρόταση της γλώσσας της σκέψης. Για να δώσουμε ένα υποθετικό παράδειγμα, μια νοητική (γλωσσική) πρόταση καθίσταται πεποίθηση αν αποτελεί έξοδο του αντιληπτικού συστήματος και είσοδο του συναγωγικού συστήματος⁶³. Έτσι, για τον Fodor, οι προτασιακές στάσεις δεν είναι τίποτα άλλο από

⁶³ Θα επανέλθουμε αναλυτικότερα σε αυτό το ζήτημα στην ενότητα 2.6.

υπολογιστικές σχέσεις του οργανισμού με εσωτερικές συμβολικές αναπαραστάσεις του.

Με την υπόθεση της γλώσσας της σκέψης, ο Fodor προσπαθεί να φυσικοποιήσει την αποβλεπτικότητα με βάση την έννοια της νοητικής αναπαράστασης. Δηλώνει χαρακτηριστικά:

"εάν η ΑΘΝ είναι αληθής, το πρόβλημα της αποβλεπτικότητας του νοητικού είναι σε μεγάλο μέρος - ίσως εξαντλητικά - το πρόβλημα της σημασιακότητας (semantivity) των νοητικών αναπαραστάσεων" (Fodor 1985a, σελ. 28)

Οι νοητικές αναπαραστάσεις είναι οι πρωτογενείς σημασιολογικές οντότητες από τις οποίες αποκτούν δευτερογενώς το περιεχόμενό τους τόσο οι προτασιακές στάσεις, όσο και τα ομιλιακά ενεργήματα (Fodor & Lepore 1992 σελ. 125). Έτσι, ο Fodor, προκειμένου να αποφύγει την κυκλικότητα, είναι υποχρεωμένος να εξηγήσει τις νοητικές αναπαραστάσεις, χωρίς να κάνει χρήση αποβλεπτικών όρων. Χρειάζεται, με άλλα λόγια, μια φυσιοκρατική εξήγηση της αναπαραστασιακής σχέσης - του πώς μπορεί, δηλαδή, κάποια υλική κατάσταση να αναπαριστά κάποια άλλη⁶⁴.

Σε τι έγκειται όμως η πρωτοτυπία της ιδέας του Fodor; Σίγουρα δεν έγκειται στην υιοθέτηση νοητικών αναπαραστάσεων για την εξήγηση της σχέσης του νου με τον κόσμο. Κάτι τέτοιο έχει γίνει από πολύ παλιά στη φιλοσοφία. Ήδη από την αρχαιότητα έχουμε συναφείς αναφορές, ενώ ένα μεγάλο μέρος της νεότερης φιλοσοφίας συγκροτείται γύρω από το παραστασιακό μοντέλο του νου: ο νους είναι μια οντότητα ριζικά διαφορετική από τον εκτατό κόσμο και επικοινωνεί με αυτόν μέσω παραστάσεων⁶⁵ (ή, κατά ορισμένους, ιδεών)⁶⁶. Οι νοητικές παραστάσεις, τουλάχιστον για τους βρετανούς εμπειριστές, αποτελούν μια εικόνα - ένα μοντέλο του κόσμου - είναι, δηλαδή, ένας αναδιπλασιασμός του κόσμου εντός του υποκειμένου⁶⁷.

Ο Fodor διαφοροποιείται σε τέσσερα τουλάχιστον σημεία από το παραδοσιακό αυτό μοντέλο για τη σχέση νου-κόσμου.

1. Απορρίπτει το δυϊσμό. Η οντολογία του νου δεν διαφέρει από την οντολογία του κόσμου. Οι νοητικές αναπαραστάσεις προκειμένου να

⁶⁴ Αυτό το πρόβλημα θα εξεταστεί αναλυτικά στο έβδομο κεφάλαιο.

⁶⁵ Vorstellungen - βλ. κεφάλαιο 3, ενότητα 3.1 για τη δικαιολόγηση της μεταφραστικής επιλογής.

⁶⁶ Για μια συστηματική παρουσίαση αυτού του μοντέλου στον Leibniz και τον Καρτέσιο βλ.επε Aparolitanos 1998, κεφάλαιο πρώτο.

⁶⁷ Βλ. κεφάλαιο 3, ενότητα 3.1. Βλ., επίσης, Haugeland 1985, κεφάλαιο πρώτο.

- μπορούν να επηρεάζουν την ανθρώπινη συμπεριφορά θα πρέπει να έχουν υλικό φορέα ώστε να επιδρούν αιτιακά⁶⁸.
2. Οι αναπαραστάσεις είναι ρητά γλωσσικού και όχι εικονιστικού ή άλλου τύπου. Ο Fodor υιοθετεί το μοντέλο της γλώσσας για να εξηγήσει τη νόηση.
 3. Η ταυτότητα των αναπαραστάσεων δεν εξαρτάται από τον αν είναι συνειδητές ή όχι. Οι αναπαραστάσεις, όντας αιροσπέλαστες από τη συνείδηση, διαθέτουν σημασιολογικές και αιτιακές ιδιότητες. Η σχέση που υπάρχει ανάμεσα στον οργανισμό και τις νοητικές του αναπαραστάσεις καθορίζεται υπολογιστικά.
 4. Ως συνέπεια της προηγούμενης θέσης, ο Fodor απορρίπτει την ενδοσκόπηση και την οπτική του πρώτου προσώπου ως μέθοδο προσέγγισης των νοητικών αναπαραστάσεων. Υιοθετεί αντιθέτως την οπτική του τρίτου προσώπου -την οπτική της επιστήμης σύμφωνα με την οποία, τα πράγματα περιγράφονται όπως δίδονται σε έναν εξωτερικό παρατηρητή⁶⁹.

2.4.2 Σε τι συνίσταται η γλωσσικότητα των προτάσεων του νου;

Τι χαρακτηριστικά όμως έχουν αυτές οι αναπαραστάσεις που τις καθιστούν γλωσσικές; Η γλώσσα της σκέψης, όντας διαφορετική από τη φυσική γλώσσα, είναι ουσιαστικά μια *έμφυτη* γλώσσα, γεγονός που αποκλείει το ενδεχόμενο τα σημαίνοντα αυτής να έχουν μια *συμβατική* σχέση με τα σημασιόμενά τους. Η γλώσσα της σκέψης αποτελεί ουσιαστικά μια *ιδιωτική* γλώσσα η οποία δεν έχει επικοινωνιακή λειτουργία (τουλάχιστον σε διύποκειμενικό επίπεδο). Για την ακρίβεια δεν πρόκειται ούτε καν για ιδιωτική γλώσσα γιατί, *εξ υποθέσεως*, δεν δίδεται στη συνείδηση. Η γλώσσα της σκέψης "χρησιμοποιείται" μόνο σε *υποπροσωπικό* επίπεδο. Ήδη, απ' αυτές τις πρώτες διαπιστώσεις γίνεται

⁶⁸ Για την ακρίβεια αυτό που επιδρά αιτιακά επηρεάζοντας την συμπεριφορά είναι οι προτασιακές στάσεις, αναπόσπαστο μέρος των οποίων είναι οι νοητικές αναπαραστάσεις. Ο Fodor έχει και ένα πρόσθετο λόγο να υποστηρίζει ότι οι νοητικές αναπαραστάσεις θα πρέπει να διαθέτουν αιτιακές ιδιότητες: προκειμένου να εξηγήσει το πώς μια ατομική νοητική αναπαράσταση αποκτά σημασιολογικές ιδιότητες προσφεύγει στην υποτιθέμενη αιτιακή σχέση της με το αναπαριστώμενο (βλ. Κεφάλαιο 7). Στη δεύτερη αυτή περίπτωση, οι νοητικές αναπαραστάσεις δεν έχουν τη θέση αιτίου αλλά αιτιατού.

⁶⁹ Με τα δύο τελευταία σημεία ο Fodor αποφεύγει - τουλάχιστον εκ πρώτης όψεως - την κριτική του Frege (1884/1959) και του Husserl (1900/1970) κατά του ψυχολογισμού, καθώς και το επιχείρημα του Wittgenstein (1953/1977) κατά της ιδιωτικής γλώσσας, γιατί οι νοητικές αναπαραστάσεις, όπως τις ορίζει, δεν αποτελούν ψυχολογικά αντικείμενα που δίδονται στην ενδοσκόπηση (ό,τι ο Frege ονόμαζε *Vorstellungen*, ήτοι παραστάσεις (βλ. κεφάλαιο 3, ενότητα 3.1).

φανερό ότι δεν πρόκειται για μια γλώσσα με τη συνήθη έννοια του όρου. Αυτό σημαίνει ότι η γλώσσα της σκέψης πρέπει να μοιράζεται μόνο κάποια πολύ γενικά χαρακτηριστικά με τις φυσικές γλώσσες. Όμως ποια θα μπορούσαν να είναι αυτά;

Δύο πολύ σημαντικές ιδιότητες που χαρακτηρίζουν τις φυσικές γλώσσες είναι, σύμφωνα με τον Fodor, η συστηματικότητα και η παραγωγικότητα.

- **Συστηματικότητα** είναι η ικανότητα που έχουμε, όταν παράγουμε και κατανοούμε μια γλωσσική πρόταση, να μπορούμε να παράγουμε και να κατανοούμε και ένα σύνολο άλλων προτάσεων που περιέχουν αναδιατεταγμένα (κάποια από) τα συστατικά της πρώτης γλωσσικής πρότασης. Για παράδειγμα, αν μπορούμε να παράγουμε και να κατανοούμε τη γλωσσική πρόταση "ο Fodor εκτιμά τον Dennett" τότε μπορούμε επίσης να παράγουμε και να κατανοούμε την γλωσσική πρόταση "ο Dennett εκτιμά τον Fodor", και
- **Παραγωγικότητα** είναι η ικανότητα να παράγουμε και να κατανοούμε ένα (δυνάμει) άπειρο αριθμό εντελώς νέων προτάσεων

Σύμφωνα με τον Fodor, τα μη-γλωσσικά αναπαραστασιακά συστήματα, όπως για παράδειγμα, τα εικονιστικά ή, γενικότερα, τα αναλογικά συστήματα, δεν παρουσιάζουν άμεσα αυτές τις ιδιότητες. Αντίθετα, η σκέψη, όπως ακριβώς και η φυσική γλώσσα, χαρακτηρίζεται από παραγωγικότητα και συστηματικότητα. Για το λόγο αυτό ο Fodor υποθέτει ότι το αναπαραστασιακό σύστημα της σκέψης πρέπει να είναι γλωσσικής μορφής. Το επιχείρημα αυτό, προφανώς, δεν απαντά στο ερώτημα: γιατί θα πρέπει η σκέψη να βασίζεται σε κάποιο αναπαραστασιακό σύστημα, αλλά στο ερώτημα: δεδομένου ότι η σκέψη βασίζεται σε κάποιο αναπαραστασιακό σύστημα, γιατί αυτό θα πρέπει να είναι γλωσσικό; Υπενθυμίζουμε ότι η απάντηση του Fodor στο πρώτο ερώτημα είναι ότι τα σύμβολα αποτελούν τα μόνα από τα γνωστά φυσικά αντικείμενα που έχουν ταυτόχρονα αιτιακές και σημασιολογικές ιδιότητες, όπως ακριβώς και οι προς εξήγηση προτασιακές στάσεις.

Το επόμενο βήμα στη συλλογιστική του Fodor είναι να υποστηρίξει ότι από τη παραγωγικότητα και τη συστηματικότητα της γλώσσας της σκέψης συνεπάγεται ότι η σημασιολογία αυτής της γλώσσας είναι *συνθετική* (compositional). Ένα αναπαραστασιακό σύστημα έχει συνθετική σημασιολογία αν η σημασία (ή, κατά ορισμένους, οι συνθήκες αλήθειας) μιας συντακτικά σύνθετης έκφρασης αυτού του συστήματος είναι συνάρτηση της συντακτικής δομής της και της σημασίας των συντακτικών συστατικών της. Θα πρέπει να επιστημονηθεί ότι ο Fodor δεν ισχυρίζεται ότι η αναγκαιότητα ύπαρξης μιας συνθετικής σημασιολογίας συνεπάγεται λογικά από την παραγωγικότητα και τη συστηματικότητα της γλώσσας της σκέψης. Η συναγωγή που προτείνει δεν είναι λογικού τύπου, αλλά συναγωγή προς την καλύτερη εξήγηση (inference to the best

explanation), δηλαδή, αποδοχή μιας υπόθεσης με βάση το γεγονός ότι παρέχει την καλύτερη εξήγηση των δεδομένων από όλες τις άλλες εναλλακτικές υποθέσεις⁷⁰. Χαρακτηριστικά επισημαίνει:

"Ο ισχυρισμός δεν ήταν ότι η παραγωγικότητα *συνεπάγεται* τη συνθετικότητα, αλλά απλά ότι υπάρχει μια παντελής έλλειψη κάποιας εναλλακτικής ιστορίας"⁷¹ (Loewer & Rey 1991, σελ.304-305)

Μια φυσική συνέπεια της αρχής της συνθετικότητας είναι ότι το περιεχόμενο των νοητικών αναπαραστάσεων δεν πρέπει να εξαρτάται από τα "συμφραζόμενα" προκειμένου οι αναπαραστάσεις να έχουν την ίδια σημασιολογική συνεισφορά σε κάθε "γλωσσική πρόταση" που χρησιμοποιούνται. Ο Fodor αποδεχόμενος την ύπαρξη συνθετικής σημασιολογίας αποδέχεται ουσιαστικά τον σημασιολογικό ατομισμό, δηλαδή τη θέση ότι υπάρχουν ατομικές νοητικές αναπαραστάσεις, το περιεχόμενο των οποίων δεν εξαρτάται από τον ρόλο τους μέσα στη γλώσσα της σκέψης, αλλά από τη σχέση των αναπαραστάσεων με τον κόσμο⁷². Υπ' αυτή την έννοια, το πρόγραμμα του Fodor έχει τα βασικά χαρακτηριστικά των θεμελιωκρατικών προγραμμάτων - σε αντίθεση προς τους εμπειριστές, ωστόσο, δεν επιδιώκει να θεμελιώσει το περιεχόμενο σε κάποιου είδους γυμνή εμπειρία. Το επίπεδο στο οποίο κάνει την ανάλυσή του δεν περιλαμβάνει εμπειρίες.

Εκτός, όμως, από τον σημασιολογικό ατομισμό, ο Fodor δέχεται και μια ισχυρότερη θέση, ότι, δηλαδή, υπάρχει ένας *ισομορφισμός* ανάμεσα στη δομή της γλωσσικής πρότασης του νου και τη δομή της κατάστασης πραγμάτων που αυτή εκφράζει. Με άλλα λόγια, δέχεται ότι τα συντακτικά συστατικά της γλωσσικής πρότασης του νου (υποκείμενα και κατηγορήματα) εκφράζουν συστατικά της κατάστασης πραγμάτων (αντικείμενα και ιδιότητες): "Η δομή των γλωσσικών προτάσεων είναι κατά τον ακόλουθο τρόπο *ισομορφική* προς τη δομή των προτάσεων που εκφράζουν: *εάν μια γλωσσική πρόταση S εκφράζει την πρόταση P, τότε συντακτικά συστατικά της S εκφράζουν τα συστατικά της P*" (Fodor &

⁷⁰ Πρόκειται ουσιαστικά γιαυτό που στην Λογική ονομάζεται απαγωγή (abduction).

⁷¹ Ωστόσο, θα πρέπει να επισημάνουμε ότι τα τελευταία χρόνια έχουν αναπτυχθεί μη-γλωσσικά συστήματα που επιδεικνύουν κάποιο βαθμό παραγωγικότητας και συστηματικότητας (van Gelder 1990). Επίσης, τουλάχιστον όσον αφορά τη φυσική γλώσσα, υπάρχουν πολύ σημαντικές αντιρρήσεις για το αν αυτή διαθέτει όντως μια συνθετική σημασιολογία (Schiffer 1987).

⁷² Ο Fodor έρχεται εδώ σε ρήξη με μια μεγάλη παράδοση στη σημασιολογία που ξεκινά από τη θέση ότι μονάδα νοήματος είναι η γλωσσική πρόταση:

"...οφείλουμε πάντοτε να έχουμε προ των ματιών μας μια πλήρη γλωσσική πρόταση. Μόνο μέσα σε μια γλωσσική πρόταση έχουν οι λέξεις σημασία.... Είναι αρκετό η γλωσσική πρόταση λαμβανόμενη ως όλον να έχει κάποιο νόημα· αυτό είναι που μεταφέρει επίσης στα μέρη της το περιεχόμενό τους" Frege 1959, § 60.

Lepore 1994, σελ. 147)⁷³. Αυτό, ουσιαστικά, σημαίνει ότι διακρίσεις που αφορούν το περιεχόμενο των νοητικών αναπαραστάσεων "κατοπτρίζονται" στο επίπεδο της μορφής των νοητικών αναπαραστάσεων. Όπως χαρακτηριστικά τονίζει ο Fodor: "η συντακτική δομή των νοητικών καταστάσεων κατοπτρίζει τις σημασιολογικές σχέσεις μεταξύ των αποβλεπτικών αντικειμένων τους" (Fodor 1987, σελ. 138)⁷⁴.

Προτού προχωρήσουμε όμως, ας κάνουμε μια μικρή ανακεφαλαίωση σε αυτό το σημείο. Είδαμε ότι, κατ' αρχάς, ο Fodor ξεκινά από το πρόβλημα της φυσικοποίησης της σχέσης υποκειμένου-κατάστασης πραγμάτων, το οποίο και ανάγει στο πρόβλημα της φυσικοποίησης των νοητικών αναπαραστάσεων. Στη συνέχεια, δείχνει ότι οι νοητικές αναπαραστάσεις πρέπει να έχουν συνθετική σημασιολογία και ότι η συνθετική σημασιολογία κατοπτρίζεται σε μια συνθετική σύνταξη. Ο Fodor επιτυγχάνει κατ' αυτόν τον τρόπο να φέρει το πρόβλημα στο επίπεδο περιγραφής που προσιδιάζει στην γνωσιακή επιστήμη, δηλαδή στο λειτουργικό επίπεδο, το επίπεδο που μεσολαβεί ανάμεσα στο φυσικό και το αποβλεπτικό. Το ερώτημα που εγείρεται πλέον είναι εάν μια εξήγηση με όρους του λειτουργικού επιπέδου μπορεί να θεωρηθεί *φυσιοκρατική* εξήγηση. Σε αυτό το σημείο είναι σκόπιμο να εξετάσουμε πως ακριβώς χρησιμοποιείται το μοντέλο του υπολογιστή γιατί, όπως είδαμε στο πρώτο κεφάλαιο, αυτό εξηγεί με ένα σαφή τρόπο τη σχέση του λειτουργικού με το φυσικό επίπεδο περιγραφής. Την επιμέρους θεωρία που προκύπτει από την χρήση του μοντέλου του υπολογιστή μπορούμε να την ονομάσουμε Υπολογιστική Θεωρία του Νου (ΥΘΝ).

2.5 Η Υπολογιστική Θεωρία του Νου (ΥΘΝ)

Η ΥΘΝ δεν έχει να κάνει τόσο με τις μεμονωμένες νοητικές καταστάσεις όσο με τις νοητικές διαδικασίες, δηλαδή με το πώς μεταβαίνουμε από τη μια νοητική κατάσταση στην άλλη. Συγκεκριμένα, αυτό που επιχειρεί ο Fodor είναι να δώσει μια φυσιοκρατική εξήγηση των αποβλεπτικών νόμων, να εξηγήσει δηλαδή με ποιο τρόπο οι νόμοι της ψυχολογίας μπορούν να υλοποιηθούν. Η συνήθης υλιστική απάντηση σε αυτό το πρόβλημα είναι ότι ο μηχανισμός που υλοποιεί τους αποβλεπτικούς νόμους θα πρέπει να περιγραφεί με νευροφυσιολογικούς όρους. Χαρακτηριστικό παράδειγμα μιας τέτοιας προσέγγισης αποτελεί, όπως είδαμε στο πρώτο κεφάλαιο, η θεωρία ταυτότητας τύπου. Ο Fodor,

⁷³ Βλ. επίσης Fodor & Lepore 1992, σελ. 176.

⁷⁴ Πρβλ. Frege 1979, σελ. 270: "Η γλωσσική πρόταση μπορεί να θεωρηθεί ως απεικόνιση της σκέψης, κατά την εξής έννοια: ότι οι σχέσεις των μερών προς το όλο στις σκέψεις και στα τμήματα των σκέψεων είναι, grosso modo, αντίστοιχες προς τις σχέσεις μερών προς όλο που υφίστανται στις γλωσσικές προτάσεις και στα τμήματα των γλωσσικών προτάσεων" (η μετάφραση με μικρές τροποποιήσεις ανήκει στον Μπασάκο 1999, σελ. 105)

πιστός εκφραστής της γνωσιακής επιστήμης, διαφοροποιείται ρητά από αυτή τη γραμμική σκέψης, υποστηρίζοντας ότι ο μηχανισμός που υλοποιεί τους αποβλεπτικούς νόμους είναι *υπολογιστικός* και, κατά συνέπεια, το κατάλληλο επίπεδο περιγραφής του είναι το λειτουργικό. Το κύριο επιχείρημα στο οποίο βασίζεται ο Fodor είναι εμπειρικού τύπου (συναγωγή προς την καλύτερη εξήγηση) και μπορούμε να το ανασυγκροτήσουμε ως εξής:

1. ένα κυρίαρχο χαρακτηριστικό των νοητικών διαδικασιών είναι ότι (συνήθως) διατηρούν σημασιολογικές ιδιότητες, όπως η αλήθεια.
2. ο υπολογιστής είναι ο μοναδικός μηχανισμός που γνωρίζουμε ο οποίος κατά τον συντακτικό μετασχηματισμό συμβόλων εισόδου σε σύμβολα εξόδου, διατηρεί την αλήθεια.
3. κατά συνέπεια, ο μηχανισμός που υλοποιεί τους αποβλεπτικούς νόμους είναι υπολογιστικός.

Η πρώτη προκείμενη, ουσιαστικά, αναδεικνύει ένα χαρακτηριστικό των νοητικών διαδικασιών που τις διαφοροποιεί από όλες τις άλλες φυσικές διαδικασίες- την ορθολογικότητα. Το παραπάνω επιχείρημα έχει την ίδια δομή με το επιχείρημα υπέρ της ύπαρξης νοητικών αναπαραστάσεων: Ο Fodor αρχικά εντοπίζει μια ιδιότητα που χαρακτηρίζει τις νοητικές διαδικασίες -την ορθολογικότητα- όπως, αντίστοιχα, για τις νοητικές καταστάσεις ήταν η ταυτόχρονη ύπαρξη αιτιακών και σημασιολογικών ιδιοτήτων. Κατόπιν βρίσκει κάποια γνωστά φυσικά φαινόμενα που παρουσιάζουν τις ίδιες ιδιότητες (οι υπολογιστικές διαδικασίες και τα σύμβολα αντίστοιχα). Και τέλος συμπεραίνει ότι οι νοητικές διαδικασίες και οι νοητικές καταστάσεις πρέπει να βασίζονται, αντίστοιχα, σε ανάλογα φυσικά φαινόμενα.

Η ΥΘΝ δεν αποτελεί κάποια ιδιαίτερα πρωτότυπη συνεισφορά του Fodor. Το μόνο που κάνει είναι να αναδειξει ορισμένες από τις βασικές ιδέες πάνω στις οποίες στηρίχθηκε η κατασκευή των ψηφιακών υπολογιστών. Οι αναφορές του δεν είναι ούτε εκτενείς ούτε συστηματικές. Στο πρώτο κεφάλαιο ανασυγκροτήσαμε τις εξελίξεις στη μαθηματική λογική που κατέστησαν δυνατή την επινόηση των ψηφιακών υπολογιστών. Έχοντας αυτό ως υπόβαθρο, μπορούμε να κατανοήσουμε άμεσα τις απόψεις του Fodor.

Είδαμε ότι με το θεώρημα της πληρότητας για τον κατηγορηματικό λογισμό ουσιαστικά εξασφαλίζεται μια σύνδεση του σημασιολογικού με το συντακτικό επίπεδο: οι σημασιολογικές σχέσεις παραλληλίζονται προς τις συντακτικές σχέσεις. Συγκεκριμένα, ο Gödel έδειξε ότι αν μια πρόταση είναι αποδείξιμη με βάση τα αξιώματα και τους κανόνες συναγωγής του κατηγορηματικού λογισμού, τότε είναι και αληθής. Κατέστη, έτσι, δυνατή η μελέτη της λογικής εγκυρότητας [logical validity] ενός επιχειρήματος ή της αλήθειας ενός θεωρήματος με βάση την αποδειξιμότητά τους [provability]. Με τους υπολογιστές αυτή η

διαδικασία αυτοματοποιήθηκε και ταυτόχρονα μηχανοποιήθηκε⁷⁵. Αυτό κατέστη δυνατό γιατί όπως επισημαίνει ο Fodor:

Οι λειτουργίες της μηχανής [δηλ. του H/Y] συνίστανται αποκλειστικά σε μετασχηματισμούς συμβόλων κατά τη διάρκεια εκτέλεσης αυτών των λειτουργιών, η μηχανή είναι ευαίσθητη μόνο στις συντακτικές ιδιότητες των συμβόλων και οι λειτουργίες που η μηχανή εκτελεί επί των συμβόλων περιορίζονται αποκλειστικά στην αλλαγή των σχημάτων τους. (Fodor 1987, σελ. 19)

Στους υπολογιστές, λοιπόν, οι συντακτικοί μετασχηματισμοί των συμβόλων, παρότι δεν βασίζονται στη σημασία των συμβόλων, έχουν κατασκευαστεί κατά τέτοιο τρόπο ώστε να διατηρούν τις σημασιολογικές σχέσεις που ισχύουν μεταξύ των συμβόλων. Με άλλα λόγια, οι υπολογιστές αποτελούν ένα από παράδειγμα για το πώς οι σημασιολογικές σχέσεις μπορούν να υλοποιηθούν σε ένα φυσικό μέσο, και αυτό γιατί οι συντακτικές σχέσεις μπορούν να περιγραφούν πλήρως μέσω αιτιοκρατικών σχέσεων, δηλαδή χωρίς καμιά αναφορά στη σημασία των συμβόλων. Έτσι, σύμφωνα με την ερμηνεία του Fodor, οι υπολογιστές μας δείχνουν "πώς συνδέονται οι σημασιολογικές ιδιότητες των συμβόλων με τις αιτιακές τους ιδιότητες μέσω της σύνταξης" (Fodor 1987, σελ. 18)

Ο Fodor υποστηρίζει ότι οι νοητικές διαδικασίες μπορούν, καθ' όμοιο τρόπο, να περιγραφούν ως υπολογιστικές διαδικασίες που εκτελούνται ακολουθώντας τους συντακτικούς κανόνες της γλώσσας της σκέψης (Βλέπε Πίνακα 1). Επισημαίνει χαρακτηριστικά:

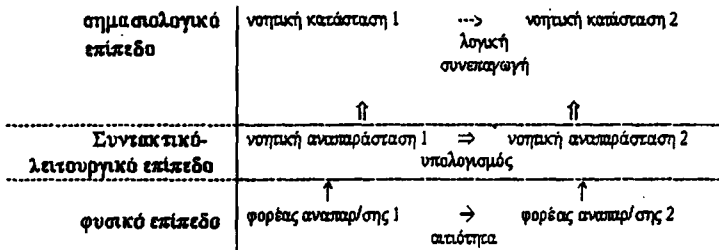
"...μια κεντρική ιδέα που υπόκειται στην κλασσική αντιμετώπιση των νοητικών διαδικασιών είναι ότι είναι δυνατό να επινοήσουμε μηχανές των οποίων η λειτουργία συνίσταται στον μετασχηματισμό συμβόλων, και των οποίων οι λειτουργίες είναι ευαίσθητες στη συντακτική δομή των συμβόλων επί των οποίων λειτουργούν..... Εάν, κατ' αρχήν, οι συντακτικές σχέσεις μπορούν να παραλληλιστούν προς τις σημασιολογικές σχέσεις, και εάν, κατ' αρχήν, μπορούμε να έχουμε ένα μηχανισμό του οποίου οι λειτουργίες επί των τύπων είναι ευαίσθητες στη σύνταξη τους, τότε είναι ίσως δυνατό να κατασκευαστεί μια συντακτικά καθοδηγούμενη μηχανή της οποίας οι μεταβάσεις από τη μια κατάσταση στην άλλη ικανοποιούν τα σημασιολογικά κριτήρια συνεκτικότητας (coherence). Μια τέτοια μηχανή θα ήταν, ό,τι ακριβώς

⁷⁵ Ο Fodor δεν παύει να επαινεί τον εμπνευστή αυτής της ιδέας, τον Turing. Επισημαίνει χαρακτηριστικά: "Στην πραγματικότητα ο Turing είπε: 'Όταν καταστάσεις ενός μηχανισμού έχουν τόσο συντακτικές όσο και σημασιολογικές ιδιότητες, τότε είναι συχνά δυνατό να ορισθούν αιτιακές διαδικασίες επί των συντακτικών ιδιοτήτων κατά τέτοιο τρόπο ώστε οι σημασιολογικές ιδιότητες να διατηρούνται'. Και εν συνεχεία εγένετω φως", Fodor 1991, σελ. 285.

απαιτείται για ένα μηχανικό μοντέλο της σημασιολογικής συνεκτικότητας της σκέψης: κατά συνέπεια, η ιδέα ότι ο εγκέφαλος είναι μια τέτοια μηχανή αποτελεί μια θεμελιακή υπόθεση της κλασικής γνωσιακής επιστήμης" (Fodor & Pylyshyn 1988, σελ. 30)⁷⁶

Όσον αφορά τη φύση των συντακτικών ιδιοτήτων, ο Fodor υποστηρίζει ότι πρόκειται για φυσικές ιδιότητες ανώτερης τάξης⁷⁷. Για παράδειγμα, αναφέρει ότι οι συντακτικές ιδιότητες θα μπορούσαν να είναι κάποιο "αφηρημένο χαρακτηριστικό του σχήματος" (Fodor 1987, σελ. 18), ενός συμβόλου. Γενικότερα, αποδέχεται τη βασική θέση του λειτουργισμού ότι οι συντακτικές σχέσεις μπορούν να υλοποιηθούν με *οποιοδήποτε νομοτελειακές (nomisic) ιδιότητες των συμβόλων* (Fodor 1987, σελ. 156). Κατά συνέπεια, το λειτουργικό - συντακτικό επίπεδο δεν εξαρτάται άμεσα από το φυσικό επίπεδο στο οποίο υλοποιείται.

Πίνακας 1



Έτσι, λοιπόν, η κεντρική θέση γύρω από την οποία συγκροτείται η ΥΘΝ είναι η εξής: οι νοητικές διαδικασίες είναι *υπολογιστικές διαδικασίες* οι οποίες υλοποιούν τους *αποβλεπτικούς νόμους* του σημασιολογικού επιπέδου (Loewer & Rey 1991, σελ 283)⁷⁸. Ο νους, σύμφωνα με την ΥΘΝ, είναι ένας ψηφιακός υπολογιστής ή, με άλλα λόγια, μια μηχανή διακριτών καταστάσεων που χειρίζεται συμβολικές αναπαραστάσεις μέσω συντακτικών κανόνων, δηλαδή κανόνων που είναι ευαίσθητοι στις συντακτικές-μορφικές και όχι στις σημασιολογικές ιδιότητες των

⁷⁶ Βλέπε επίσης Fodor 1987, σελ. 19: "αν ο νους είναι κάποιο είδος υπολογιστή, τότε αρχίζουμε να κατανοούμε πώς... θα μπορούσαν να υπάρχουν μη-αυθαίρετες νοηματικές σχέσεις μεταξύ αιτιακά συνδεδεμένων σκέψεων"

⁷⁷ Ένα πρόβλημα εδώ, το οποίο ωστόσο δεν θα εξετάσουμε περαιτέρω, είναι ότι ο Fodor χρησιμοποιεί την έννοια των συντακτικών ιδιοτήτων με έναν αμφίσημο τρόπο: άλλοτε τις έννοεί ως μορφικές και άλλοτε ως τυπικές ιδιότητες. (Βλ. Devitt 1989 και Devitt 1991, σελ. 99-102)

⁷⁸ Βλ. επίσης Fodor 1987, σελ. 166.

αναπαραστάσεων. Οι Fodor και Pylyshyn (1988, σελ. 294) επισημαίνουν χαρακτηριστικά:

"Στα κλασικά μοντέλα, οι αρχές βάσει των οποίων μετασχηματίζονται οι νοητικές καταστάσεις, ή βάσει των οποίων μία είσοδος επλέγει την αντίστοιχη έξοδο, ορίζονται επί δομικών ιδιοτήτων των νοητικών αναπαραστάσεων. Επειδή οι κλασικές νοητικές αναπαραστάσεις έχουν συνδυαστική δομή, είναι δυνατόν οι κλασικοί νοητικοί χειρισμοί να γίνονται σε αυτές με αναφορά στην μορφή τους. Το αποτέλεσμα είναι ότι μια παραδειγματική κλασική νοητική διαδικασία τελείται επί οποιασδήποτε νοητικής αναπαράστασης που ικανοποιεί μια δεδομένη δομική περιγραφή και τη μετασχηματίζει σε μια νοητική αναπαράσταση που ικανοποιεί μια άλλη δομική περιγραφή"

Ο νους δηλαδή είναι, όπως λέει και ο Dennett (1987, σελ. 61), μια συντακτική και όχι μια σημασιολογική μηχανή⁷⁹. Οι σημασιολογικές ιδιότητες των νοητικών αναπαραστάσεων έχουν αιτιακή επίδραση μόνο μέσω αντιπροσώπων, δηλαδή μέσω των συντακτικών τους ιδιοτήτων. Ο νους είναι μια συντακτική μηχανή που καθοδηγεί μια σημασιολογική μηχανή.

2.6 Η Αρχιτεκτονική του Νου

2.6.1 Τα τρία είδη γνωσιακών συστημάτων

Θα ολοκληρώσουμε αυτό το κεφάλαιο με την παρουσίαση των θέσεων του Fodor σε ό,τι αφορά την αρχιτεκτονική του νου. Έτσι, θα αποκτήσουμε μια πιο λεπτομερή εικόνα του τρόπου εφαρμογής της ΑΥΘΝ καθώς και του εύρους αυτής της εφαρμογής.

Στο βιβλίο *The Modularity of the Mind* ο Fodor προτείνει έναν τριμερή⁸⁰ χωρισμό των γνωσιακών συστημάτων του νου:

1. τους μορφομετατροπείς (transducers) οι οποίοι είναι αναλογικά συστήματα που μετατρέπουν τα ερεθίσματα του περιβάλλοντος σε συμβολικές περιγραφές τους (σε νευρωνικό κώδικα) ή, αντίστροφα, τα σύμβολα σε μυϊκές κινήσεις,
2. τα συστήματα εισόδου και εξόδου⁸¹ τα οποία αποτελούν αυτόνομους και εξειδικευμένους υπομηχανισμούς (modules) και

⁷⁹ Βλ. επίσης Fodor 1987, σελ. 20: "...ο νους ως μια συντακτικά καθοδηγούμενη μηχανή".

⁸⁰ Ο Fodor αφήνει ανοιχτό το ενδεχόμενο ορισμένα είδη γνωσιακών μηχανισμών να μην υποπίπτουν σε μια από τις τρεις παρακάτω κατηγορίες.

⁸¹ Θα πρέπει να επισημανθεί ότι ο Fodor δεν επιχειρηματολογεί υπέρ των δευτέρων. Αρκείται απλώς να επισημάνει ότι θα ήταν ιδιαίτερα ευχαριστημένος αν κάποιος έδειχνε ότι και στην περίπτωση των συστημάτων εξόδου ισχύουν τα ίδια

3. τα κεντρικά συστήματα

Σύμφωνα, με αυτόν τον χωρισμό ο νους δεν είναι ένα, εν πολλοίς, ομογενές σύστημα, ένας "γενικός επιλυτής προβλημάτων", όπως υπέθεταν οι ερευνητές στο ξεκίνημα της Τεχνητής Νοημοσύνης⁸². Ο Fodor υποστηρίζει ότι υπάρχει μια σαφής διάκριση ανάμεσα στην αντίληψη και τη σκέψη/νόηση. Προσπαθεί δε να θεμελιώσει αυτή την προθεωρητική διάκριση εισάγοντας τη διάκριση ανάμεσα στα συστήματα εισόδου και τα κεντρικά συστήματα. Ο Fodor κατατάσσει επίσης στα συστήματα εισόδου και τους διάφορους μηχανισμούς ανάλυσης και επεξεργασίας της φυσικής γλώσσας.

Σε ό,τι ακολουθεί θα εξετάσουμε λίγο πιο αναλυτικά τα χαρακτηριστικά γνωρίσματα των τριών προτεινόμενων γνωσιακών συστημάτων και τις μεταξύ τους διαφορές και στη συνέχεια θα αναφερθούμε στους λόγους που παραθέτει ο Fodor για να υποστηρίξει την αναγκαιότητα αυτής της τριμερούς διάκρισης.

Ας ξεκινήσουμε με τους μορφομετατροπείς. Με αυτή την κατηγορία ο Fodor αναφέρεται, μεταξύ άλλων, σε ό,τι προθεωρητικά ονομάζουμε "αισθητηριακά όργανα". Τα συστήματα αυτά αποτελούν τα σημεία επαφής του οργανισμού με το περιβάλλον και γι' αυτό βρίσκονται στην εξωτερική του επιφάνεια ή κοντά σε αυτή⁸³. Έργο τους, όπως είπαμε, είναι να μετατρέπουν τα προσπίπτοντα ερεθίσματα σε συμβολικές περιγραφές των ιδίων. Πιο συγκεκριμένα, οι μορφομετατροπείς παίρνουν ως εισόδο εγγύς ερεθίσματα (proximal stimuli) υπό φυσική περιγραφή και δίδουν ως έξοδο μια συμβολική (συντακτική) περιγραφή των εγγύς ερεθισμάτων (βλ. Σχήμα 1). Τι είναι όμως τα εγγύς ερεθίσματα; Τα ερεθίσματα αυτά ορίζονται κατ' αντιδιαστολή προς τα άπω ερεθίσματα (distal stimuli). Τα άπω ερεθίσματα είναι τα αντικείμενα⁸⁴ του περιβάλλοντος (ή οι ιδιότητές τους) που προκαλούν τον ερεθισμό των αισθητηρίων οργάνων του οργανισμού. Ακριβώς αυτός ο ερεθισμός είναι τα εγγύς ερεθίσματα. Για παράδειγμα, τα εγγύς ερεθίσματα που προκαλούνται από την παρουσία ενός σκύλου (άπω ερέθισμα) στο οπτικό πεδίο του οργανισμού δεν είναι τίποτα άλλο από την κατανομή της ενέργειας πάνω στον αμφιβληστροειδή εξαιτίας αυτής της παρουσίας. Οι μορφομετατροπείς, σύμφωνα με τον Fodor, είναι αναλογικοί μηχανισμοί και, κατά συνέπεια, η λειτουργία τους δεν βασίζεται σε εσωτερικές

επιχειρήματα με αυτά που ισχύουν για τα συστήματα εισόδου, βλ. Fodor 1983, σελ. 42. Σε ό,τι ακολουθεί, όταν αναφερόμαστε στους αυτόνομους υπομηχανισμούς θα εννοούμε τα συστήματα εισόδου.

⁸² Βλέπε επόμενη ενότητα 2.6.2.

⁸³ Ωστόσο, δύσκολα θα μπορούσε να ενταχθεί σε αυτό το μοντέλο και η κιναισθηση

⁸⁴ Ή, όσον αφορά την αίσθηση της αφής, τμήματα των αντικειμένων.

συμβολικές αναπαραστάσεις⁸⁵ και συναγωγές, ούτε μπορεί να περιγραφεί υπολογιστικά⁸⁶. Ως τέτοιοι, οι μορφομετατροπείς δεν εμπνέουν στο ερευνητικό πεδίο της κλασικής γνωσιακής επιστήμης αλλά της ψυχοφυσικής (psychophysics).

Οι αυτόνομοι υπομηχανισμοί, αντίθετα, αποτελούν το κατεξοχήν πεδίο έρευνας της γνωσιακής επιστήμης. Είσοδός τους είναι η έξοδος των μορφομετατροπέων, δηλαδή κάποια συμβολική/συντακτική περιγραφή των εγγύς ερεθισμάτων (βλ. Σχήμα 1). Η έξοδος των αυτόνομων υπομηχανισμών είναι κάποια συμβολική περιγραφή των άνω ερεθισμάτων και προκύπτει μέσα από μια υπολογιστική διαδικασία συναγωγών⁸⁷. Έτσι, σύμφωνα με τον Fodor, οι αυτόνομοι υπομηχανισμοί μπορούν να θεωρηθούν ότι 'προτείνουν' υποθέσεις σχετικά με τις άνω πηγές των αισθητηριακών (εγγύς) ερεθισμάτων. "Αυτές οι υποθέσεις μεταφέρονται με ένα περιορισμένο (δηλ. αντιληπτικό) λεξιλόγιο και στηρίζονται αντίστοιχα σε ένα περιορισμένο σύνολο πληροφοριών, ήτοι τρέχουσες αισθητηριακές πληροφορίες μαζί με οποιοδήποτε τμήμα της θεωρίας υποβάθρου στο οποίο έχει πρόσβαση ο υπομηχανισμός. Οι υποθέσεις που προτείνουν οι αυτόνομοι υπομηχανισμοί συγκρίνονται εν συνεχεία με το υπόλοιπο της θεωρίας υποβάθρου του οργανισμού και από αυτή τη σύγκριση απορρέει η αντιληπτική παγίωση μιας πεποιθήσης" (Fodor 1990, σελ. 249)

Σύμφωνα, λοιπόν, με τον Fodor, η λειτουργία της σύγκρισης των υποθέσεων με το σύνολο της θεωρίας υποβάθρου και, κατά συνέπεια, η παγίωση των πεποιθήσεων δεν επιτελείται από τους πληροφοριακά απομονωμένους υπομηχανισμούς αλλά από τα κεντρικά συστήματα τα οποία έχουν πρόσβαση στο σύνολο της θεωρίας υποβάθρου. Οι υπομηχανισμοί είναι υπεύθυνοι μόνο για την παγίωση των φαινομένων⁸⁸

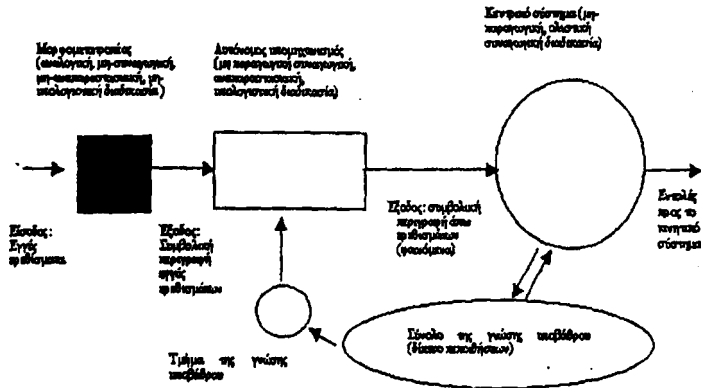
⁸⁵ Μόνο η έξοδος αυτών βρίσκεται σε συμβολική μορφή.

⁸⁶ Το ερώτημα που παραμένει ανοιχτό, ωστόσο, είναι: δεδομένου ότι όλες οι διαδικασίες είναι αιτιακές, τι είναι εκείνο που διαφοροποιεί τις υπολογιστικές/συναγωγικές αιτιακές διαδικασίες από τις μη-υπολογιστικές/συναγωγικές αιτιακές διαδικασίες;

⁸⁷ Για την ακρίβεια μεταξύ ενός μορφομετατροπέα και κάποιου κεντρικού συστήματος δεν μεσολαβεί μόνο ένας αυτόνομος υπομηχανισμός: η είσοδος του κεντρικού συστήματος μπορεί να προκύπτει μέσα από μια σειρά διαδοχικών αυτόνομων υπομηχανισμών

⁸⁸ Ένα ενδιαφέρον ερώτημα είναι αν μπορούμε να μιλάμε για φαινόμενα χωρίς να προϋποθέτουμε ότι αυτά δίδονται στη συνείδηση του υποκειμένου. Εν προκειμένω, φαίνεται σαν οι ίδιες οι έξοδοι των υπομηχανισμών να είναι από μόνες τους συνειδητές. Κάτι τέτοιο μοιάζει παραδοξολογικό. Αν, από την άλλη μεριά, αυτό που εννοεί ο Fodor με τον όρο "φαινόμενο" είναι το περιεχόμενο της εμπειρίας μας, τότε υπάρχει ένα άλμα στη θεωρία του: πώς από εκεί που μιλούσαμε αποκλειστικά για νοητικές αναπαραστάσεις που "διέθεται" περιεχόμενο το οποίο δεν δίδεται στη συνείδηση (αναφερόμαστε στις αναπαραστάσεις που χειρίζονται οι αυτόνομοι υπομηχανισμοί), φτάνουμε ξαφνικά

(ή, όπως λέει ο Fodor, για την παρατήρηση) και όχι για το τι πιστεύουμε για αυτά (δηλαδή, για την παγίωση των πεποιθήσεων).



Σχήμα 1: Σχηματική απεικόνιση της σχέσης μεταξύ μορφομετατροπέα, αυτόνομου υποσυστήματος, κεντρικού συστήματος και συνόλου γνώσης υποβάθρου.

Ο Fodor παραθέτει μια σειρά χαρακτηριστικών τα οποία διακρίνουν τους υπομηχανισμούς από τα κεντρικά συστήματα. Αναφέρουμε τα πιο σημαντικά:

1. **εξειδίκευση:** κάθε υπομηχανισμός είναι προσαρμοσμένος στην επεξεργασία συγκεκριμένου είδους ερεθισμάτων, και κατ' επέκταση, συγκεκριμένου είδους περιεχομένων.
2. **πληροφοριακή απομόνωση (informational encapsulation):** ένας υπομηχανισμός έχει πρόσβαση σε περιορισμένες πληροφορίες οι οποίες είτε δίδονται στην είσοδό του είτε υπάρχουν στους υπομηχανισμούς που τον αποτελούν. Υπό αυτή την έννοια, ένας υπομηχανισμός έχει μια αυτονομία γιατί δεν εξαρτάται από τις πληροφορίες άλλων υπομηχανισμών (εκτός φυσικά και αν αυτές αποτελούν είσοδό του)

να μιλάμε για αναπαραστάσεις που το περιεχόμενό τους εμφανίζεται στην συνείδηση του οργανισμού (κάτι σαν τα αντιληπτικά δεδομένα [percepts]). Ποιος είναι ο οργανισμός και τι είναι η συνείδηση εν προκειμένω; Όπως θα δούμε στο όγδοο κεφάλαιο έχουμε εδώ μια αδικαιολόγητη μετάβαση από την οπτική του τρίτου προσώπου στην οπτική του πρώτου προσώπου η οποία μάλιστα συνοδεύεται από μια λανθασμένη ανάλυση της φαινομενολογίας της αντίληψης, καθώς τα πράγματα που μας δίδονται στην αντίληψη δεν μας εμφανίζονται ως υποθέσεις αλλά ως ακλόνητες βεβαιότητες. Όπως θα προσπαθήσουμε να δείξουμε στα επόμενα, αυτό που βλέπουμε είναι ο ίδιος ο κόσμος και όχι κάποιες ενδιάμεσες οντότητες που χρειάζεται να ερμηνεύσουμε.

3. *υποχρεωτικότητα*: η λειτουργία των υπομηχανισμών δεν εξαρτάται από τη βούλησή μας (π.χ. αν ανοίξουμε τα μάτια μας δεν μπορούμε παρά να δούμε)

4. *Περιορισμένη δυνατότητα συνειδητής πρόσβασης στα διάφορα επίπεδα αναπαράστασης που προκύπτουν μεταξύ διαδοχικών αυτόνομων υπομηχανισμών*: όσο πιο απομακρυσμένη από τα κεντρικά συστήματα είναι η έξοδος (αναπαράσταση) ενός αυτόνομου υπομηχανισμού τόσο μικρότερη είναι η δυνατότητα συνειδητής πρόσβασης σε αυτήν.

5. *εντοπισμένοι στο νευρωνικό επίπεδο*: οι υπομηχανισμοί δεν είναι μόνο λειτουργικά διακριτοί αλλά είναι και χωρικά διακριτοί μέσα στον εγκέφαλο.

6. *Παρουσιάζουν επιλεκτικές βλάβες*: ακριβώς επειδή οι υπομηχανισμοί είναι εντοπισμένοι, τραύματα σε συγκεκριμένες περιοχές του εγκεφάλου προκαλούν συγκεκριμένες και αναμενόμενες βλάβες.

7. *φτωχή πληροφοριακά έξοδος*: η έξοδος των υπομηχανισμών δεν περιέχει πολλές πληροφορίες γιατί αυτό αντιβαίνει στο γεγονός ότι είναι πληροφοριακά απομονωμένοι.

8. *Μη δυνατότητα πρόσβασης στις εσωτερικές αναπαραστάσεις ενός υπομηχανισμού από τους υπόλοιπους*: το χαρακτηριστικό αυτό αποτελεί άμεση συνέπεια της πληροφοριακής απομόνωσης των υπομηχανισμών.

9. *ταχύτητα*: η λειτουργία των υπομηχανισμών είναι πάρα πολύ γρήγορη

Από τις παραπάνω ιδιότητες των αυτόνομων υπομηχανισμών οι πιο σημαντικές, είναι η πληροφοριακή απομόνωση και η εξειδίκευση.

Ας δούμε όμως τώρα τι είναι τα κεντρικά συστήματα. Κατ' αρχάς, θα πρέπει να επισημανθεί ότι η υποτιθέμενη λειτουργία τους δεν έχει καμία ομοιότητα με εκείνη του κεντρικού επεξεργαστή των ηλεκτρονικών υπολογιστών. Το κυρίαρχο χαρακτηριστικό τους είναι η έλλειψη εξειδίκευσης: κάθε κεντρικό σύστημα χειρίζεται ένα σύνολο διαφορετικού τύπου περιεχομένων και, υπ' αυτή την έννοια, δεν αφορά μόνο κάποια μεμονωμένη γνωσιακή επικράτεια. Ο Fodor ονομάζει τα κεντρικά συστήματα οριζόντιες γνωσιακές δυνάμεις (*horizontal faculties*) αντιδιαστέλλοντάς τα από τους αυτόνομους υπομηχανισμούς τους οποίους ονομάζει κάθετες γνωσιακές δυνάμεις (*vertical faculties*). Χαρακτηριστικά παραδείγματα οριζόντιων γνωσιακών δυνάμεων συναντάμε στην παραδοσιακή φιλοσοφία αλλά και στην ψυχολογία του κοινού νου. Ενδεικτικά αναφέρουμε την δύναμη της μνήμης στον Πλάτωνα (*Θεαίτητος*) και τις δυνάμεις της αισθητικότητας, της φαντασίας, της διάνοιας και της κρίσης στον Kant (1929). Σε αυτές τις περιπτώσεις, οι εν λόγω δυνάμεις δεν περιορίζονται σε ένα συγκεκριμένο τύπο περιεχομένου, αλλά υποτίθεται ότι παραμένουν αμετάβλητες από τη μια θεματική περιοχή στην άλλη. Υπ' αυτή την έννοια, οι οριζόντιες γνωσιακές δυνάμεις είναι ανεξάρτητες από το θεματικό τους αντικείμενο. Όπως ήδη αναφέραμε, η τεχνητή νοημοσύνη και η γνωσιακή ψυχολογία στα πρώτα βήματά τους αντλαμβάνονταν τις γνωσιακές δυνάμεις ως οριζόντιες.

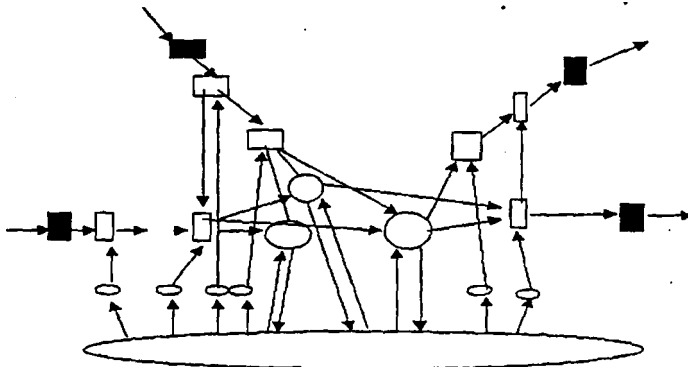
Ο Fodor αποδίδει δύο βασικές ιδιότητες στα κεντρικά συστήματα οι οποίες χαρακτηρίζουν τη λειτουργία τους. Είναι, όπως λέει, “κουαϊνικά” και ισότροπα.

Κουαϊνικός θα πει ότι “ο βαθμός επικύρωσης που απονέμεται σε μια υπόθεση εξαρτάται από τις ιδιότητες ολόκληρου του συστήματος πεποϊθήσεων” (Fodor 1983, σελ. 107), και

Ισότροπος θα πει ότι οι πεποϊθήσεις που είναι σημαντικές για την επικύρωση μιας νέας πεποϊθήσης μπορούν να βρίσκονται οπουδήποτε μέσα στο δίκτυο των πεποϊθήσεων.

Οι δύο αυτές ιδιότητες είναι αλληλένδετες και δείχνουν ότι η λειτουργία των κεντρικών συστημάτων έχει να κάνει με τον χειρισμό πεποϊθήσεων⁸⁹ και μάλιστα κατά ένα τρόπο ολιστικό (βλ. Σχήμα 2). Για τον λόγο αυτό ο Fodor είναι πολύ απαισιόδοξος για τη δυνατότητα εξήγησης της λειτουργίας των κεντρικών συστημάτων με υπολογιστικά μοντέλα. Όσο πλησιάζουμε, επισημαίνει, προς τις ανώτερες γνωσιακές λειτουργίες τόσο περισσότερο ισότροπο γίνεται το σύστημα. Χαρακτηριστικό παράδειγμα αποτελεί ο αναλογικός συλλογισμός για τον οποίο “κανείς δεν γνωρίζει πώς δουλεύει”. Έτσι, μεταξύ σοβαρού και αστείου, καταλήγει στη διατύπωση ενός νόμου που, όπως λέει, κάποια μέρα θα ήθελε να μείνει γνωστός ως “Ο πρώτος νόμος του Fodor για την ανυπαρξία της γνωσιακής επιστήμης”:

“όσο περισσότερο σφαιρική (δηλ. ισότροπη) γίνεται μια γνωσιακή διαδικασία τόσο λιγότερο την καταλαβαίνει κανείς” (Fodor 1983, σελ. 107)



Σχήμα 2: Μια σχηματική απεικόνιση ολόκληρου του γνωσιακού συστήματος η οποία ελπίζουμε ότι δεν αδικεί ιδιαίτερα τη θεωρία του Fodor. (Να επισημανθεί ότι τα κεντρικά συστήματα δεν θα πρέπει να

⁸⁹ Ενδεχομένως και των υπολοίπων προτασιακών στάσεων

θεωρηθούν ούτε καν λειτουργικά εντοπισμένα. Επίσης, ανοιχτό παραμένει το ζήτημα του εάν η γνώση υποβάθρου πρέπει να διαχωριστεί από τα συστήματα που την χρησιμοποιούν).

Είναι αξιοσημείωτο το ότι ένας από τους πιο συνεπείς και επίμονους υποστηρικτές της γνωσιακής επιστήμης φτάνει μέσα από την εσωτερική λογική του έργου του να μιλήσει για κάποιους εγγενείς περιορισμούς του αναπαραστασιακού-υπολογιστικού παραδείγματος. Με την θεωρία του για την αρχιτεκτονική του νου, ο ίδιος ουσιαστικά αποκλείει τα κεντρικά συστήματα από το πεδίο έρευνας της γνωσιακής επιστήμης. Οι συγκεκριμένες θέσεις του Fodor αποτελούν ίσως το μοναδικό σημείο όπου η θεωρία του όχι μόνο δεν υποστηρίζει το πρόγραμμα της γνωσιακής επιστήμης αλλά έρχεται σε ευθεία αντίθεση με αυτό.

2.6.2 Επιχειρήματα υπέρ της ύπαρξης αυτόνομων υπομηχανισμών

Ας εξετάσουμε τώρα τα επιχειρήματα με τα οποία ο Fodor προσπαθεί να στηρίξει τη θέση περί ύπαρξης αυτόνομων υπομηχανισμών. Μια πρώτη κατηγορία είναι εξελικτικού τύπου επιχειρήματα, δηλαδή επιχειρήματα που δείχνουν ότι οι αυτόνομοι υπομηχανισμοί συμβάλλουν στην επιβίωση ενός οργανισμού.

Η λειτουργία της αντίληψης έγκειται στο να καθορίζει, γρήγορα και αξιόπιστα, πώς είναι τα πράγματα στο άμεσο περιβάλλον του οργανισμού χωρίς να ασχολείται με το μακρινό παρελθόν ή το απώτερο μέλλον. "Μια συνθήκη για την αξιοπιστία της αντίληψης, τουλάχιστον για έναν οργανισμό που δεν είναι αλάνθαστος, είναι ότι γενικά βλέπει τι είναι εκεί, και όχι τι θέλει ή τι προσδοκά να είναι εκεί. Οργανισμοί που δεν το κάνουν αυτό πεθαίνουν" (Fodor 1983, σελ 68). Για τον ίδιο λόγο πεθαίνουν και οργανισμοί που καθυστερούν να δουν τι υπάρχει στο άμεσο περιβάλλον τους⁹⁰. Επίσης, ενδείξεις υπέρ των αυτόνομων υπομηχανισμών θεωρούνται οι διάφορες εξειδικευμένες γνωσιακές ανικανότητες που προκαλούνται

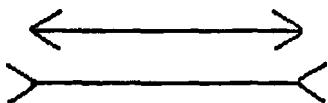
⁹⁰ Με τέτοιου είδους επιχειρήματα ο Fodor δείχνει ότι τα χαρακτηριστικά που έχει αποδώσει στους αυτόνομους υπομηχανισμούς (εξειδίκευση, απομόνωση, ταχύτητα κ.λπ.) συμβάλλουν στην επιβίωση ενός οργανισμού. Ωστόσο, αυτό που οφείλει επιπλέον να δείξει είναι το γιατί αυτά τα φαινομενολογικά και συμπεριφοριστικά χαρακτηριστικά θα πρέπει να υλοποιούνται από μηχανισμούς με τη συγκεκριμένη αρχιτεκτονική. Θα δείξουμε ότι η υπόθεσή του αντιεί την αληθοφάνειά της από την παραδοχή ότι η γνώση στο σύνολό της κωδικοποιείται προτασιακά. Μόνο με αυτή την παραδοχή δημιουργείται η συνδυαστική "έκρηξη" κατά την αναζήτηση μέσα στην βάση γνώσης του οργανισμού και κατά συνέπεια εμφανίζεται λογικό οι υπομηχανισμοί να μην έχουν πρόσβαση στο σύνολο της γνώσης. Θα επανέλθουμε σε αυτό το ζήτημα στο τέταρτο κεφάλαιο.

από εντοπισμένες εγκεφαλικές βλάβες (βλέπε, αφασίες, απραξίες, αγνωσίες κλπ).

Ο Fodor συγκρίνοντας την αντίληψη με τη σκέψη διαπιστώνει ότι ενώ η δεύτερη μπορεί να εξηγηθεί ως μια διαδικασία επίλυσης προβλημάτων, δεν ισχύει το ίδιο και για την πρώτη. Η επίλυση ενός προβλήματος έχει δύο στάδια:

1. Έρευνα για το πώς θα λυθεί και
2. Εύρεση της λύσης.

Ο Fodor πιστεύει ότι στην αντίληψη υπάρχει μόνο το δεύτερο στάδιο. Η αντίληψη είναι ουσιαστικά ένα λυμένο πρόβλημα. Γενικότερα, μπορούμε να πούμε ότι η έννοια των αυτόνομων και πληροφοριακά απομονωμένων υπομηχανισμών αναπτύχθηκε από τον Fodor για να συλλάβει ακριβώς όλες εκείνες τις διαδικασίες του οργανισμού που δεν λύνουν άγνωστα προβλήματα αλλά απλά βρίσκουν τις εκάστοτε λύσεις γνωστών προβλημάτων. Η ιδέα αυτή είναι εύλογη από εξελικτική άποψη, διότι δεν απαιτεί την απότομη εμφάνιση ενός γενικού επλυτή προβλημάτων με προσωμοσιτικότητα στις εκάστοτε περιβαλλοντικές συνθήκες. Αντίθετα, δείχνει πώς μπορούμε να φτάσουμε σιγά - σιγά σε αυτό το αποτέλεσμα με τη σταδιακή προσθήκη εξειδικευμένων ικανοτήτων⁹¹. Επίσης, αυτή η ιδέα είναι και από υπολογιστική άποψη περισσότερο εύλογη: ένα γενικού σκοπού σύστημα εκμάθησης θα βασίζονταν για τη λειτουργία του σε μια σωρεία πληροφοριών, ορισμένες μόνο από τις οποίες θα ήταν χρήσιμες για το εκάστοτε πρόβλημα. Η συλλογή ενός υπερβολικού αριθμού πληροφοριών καθιστά εξαιρετικά δαπανηρή από υπολογιστική άποψη την επεξεργασία τους.⁹²



Σχήμα 3

Ο Fodor, για να υποστηρίξει τη θέση του υπέρ της ύπαρξης αυτόνομων υπομηχανισμών, επικαλείται επίσης διάφορες αισθητηριακές απάτες οι οποίες καταδεικνύουν την απομόνωση του αντιληπτικού συστήματος από τη γνώση μας για τον κόσμο. Συγκεκριμένα, υπάρχουν ορισμένα φαινόμενα τα οποία είναι αδύνατο να τα δούμε όπως πραγματικά

⁹¹ Βλέπε, επίσης, Simon (1969) ο οποίος υποστηρίζει ότι τα βιολογικά συστήματα είναι ιεραρχικά γιατί έτσι εξελίσσονται ταχύτερα καθότι τα διάφορα υποσυστήματα σχηματίζουν πολλά ενδιάμεσα σταθερά στάδια σε μια διαδικασία.

⁹² Σε αυτές τις ιδέες έχει βασισθεί μια νέα προσέγγιση στο χώρο της ψυχολογίας, η λεγόμενη "εξελικτική ψυχολογία" (evolutionary psychology). Η εν λόγω προσέγγιση έχει ως στόχο την αναζήτηση εξειδικευμένων γνωσιακών ικανοτήτων που έπαιξαν ρόλο στην εξέλιξη ενός οργανισμού.

γνωρίζουμε ότι είναι. Τα φαινόμενα αυτά δείχνουν ότι ο ανθρώπινος εγκέφαλος δεν παρουσιάζει πλαστικότητα όσον αφορά *ορισμένες* αντιληπτικές λειτουργίες⁹³. Για παράδειγμα, στη ψευδαισθηση Mueller-Lyer (βλ. Σχήμα 3), παρά το γεγονός ότι *γνωρίζουμε* ότι οι δύο γραμμές είναι ίσες, δεν μπορούμε παρά να τις *βλέπουμε* ως άνισες. Όση θεωρία και αν αποκτήσουμε γύρω από αυτό το φαινόμενο, δεν μπορούμε να σταματήσουμε να έχουμε την ψευδαισθηση. Αυτό συμβαίνει, σύμφωνα με τον Fodor, ακριβώς επειδή το αντιληπτικό σύστημα είναι πληροφοριακά απομονωμένο. Έτσι, *το πώς μας φαίνονται τα πράγματα*⁹⁴ δεν εξαρτάται από το σύνολο των πεποιθήσεων της θεωρίας μας. Η αντίληψη, σε αντίθεση με τη σκέψη, είναι πληροφοριακά απομονωμένη και γι' αυτό βασίζεται σε ένα περιορισμένο μόνο αριθμό πληροφοριών υποβάθρου. Ακριβώς αυτή η διαφοροποίηση είναι εκείνη που, σύμφωνα με τον Fodor, *απαγορεύει τον σχετικισμό*⁹⁵:

"Το επιχείρημα υπέρ της σχετικότητας της παρατήρησης απαιτεί ... όχι απλά τον συναγωγικό χαρακτήρα της αντίληψης, αλλά την ιδέα ότι όλη σας η γνώση υποβάθρου, συμπεριλαμβανομένων ιδιαίτερα των επιστημονικών σας θεωριών, είναι προσβάσιμη ως προκείμενη για την αντιληπτική ολοκλήρωση. Αντίθετα, εάν σκεφτείτε ότι η αντίληψη, παρότι συναγωγική, είναι παρόλα αυτά απομονωμένη από πολλά από αυτά που πιστεύει ο αντιλαμβανόμενος οργανισμός, τότε αρχίζει να διαφαίνεται η κοινή γνωσιολογική κατάσταση του επιστήμονα και του απλού ανθρώπου. Υπάρχει, ενδεχομένως, μόνο ένας αντιληπτικός κόσμος, αν και οι ειδικοί μερικές φορές γνωρίζουν περισσότερα για αυτόν από τους ερασιτέχνες." (Fodor 1990c, σελ 244)

2.7 Επίλογος

Σε αυτό το κεφάλαιο εξετάσαμε τα βασικά στοιχεία γύρω από τα οποία αρθρώνεται η ΑΥΘΝ του Fodor. Μια θεωρία η οποία άρχισε σιγά-σιγά να αναπτύσσεται από τα μέσα της δεκαετίας του '60 και με τα χρόνια έχει υποστεί αρκετές μεταβολές και προσθήκες.

Για το λόγο αυτό, σταθήκαμε στις πιο χαρακτηριστικές θέσεις του Fodor δεδομένου ότι σκοπός μας δεν είναι να εξαντλήσουμε το θέμα -αλλά να συγκροτήσουμε το πλαίσιο πάνω στο οποίο δομείται η κλασική γνωσιακή επιστήμη και οι βασικές διαμάχες εντός αυτής. Έτσι,

⁹³ Για επιχειρήματα υπέρ της πλαστικότητας του εγκέφαλου βλ. Churchland 1979.

⁹⁴ Στο σημείο αυτό γίνεται φανερή η επίκληση εκ μέρους του Fodor της οπτικής του πρώτου προσώπου για να αντλήσει επιχειρήματα υπέρ της θέσεως του.

⁹⁵ Η βασική διάκριση εδώ είναι ανάμεσα σε δύο διαφορετικά είδη γνώσης υποβάθρου -ένα προσβάσιμο και ένα μη-προσβάσιμο. Ένα ζητούμενο είναι το γιατί θα πρέπει (και τα δύο) να έχουν προτασιακή δομή; Στο τέταρτο κεφάλαιο θα εξετάσουμε αναλυτικά αυτό το ζήτημα.

προσπαθήσαμε όσο είναι δυνατόν να παρουσιάσουμε την εσωτερική ενότητα που διέπει τις διάφορες επιμέρους θέσεις του Fodor, μια ενότητα που δεν είναι τόσο προφανής όταν κανείς διαβάζει το έργο του, δεδομένου του εύρους του εγχειρηματός του, του αποσπασματικού τρόπου που παρουσιάζει τις απόψεις του (κυρίως μέσω άρθρων), αλλά και των μετατοπίσεων των θέσεων του με το πέρασμα του χρόνου. Για τον ίδιο λόγο αναβάλαμε για τα επόμενα κεφάλαια κάθε είδους συστηματική κριτική.

Το πρόγραμμα του Fodor εκφράζει στις γενικές του κατευθύνσεις το σύνολο, σχεδόν, της σύγχρονης φιλοσοφίας του νου. Συγκεκριμένα, η πλειονότητα των σύγχρονων φιλοσόφων ασπάζονται με τον ένα ή τον άλλο τρόπο κάποια αναπαραστασιακή-υπολογιστική θεωρία του νου. Κεντρικό ρόλο σε όλες αυτές τις θεωρίες παίζει η έννοια της νοητικής αναπαράστασης. Γύρω από αυτή την έννοια αρθρώνονται οι βασικές διαφοροποιήσεις μεταξύ των θεωρητικών της γνωσιακής επιστήμης.

Κατά συνέπεια, υπάρχουν δύο κυρίως τρόποι με τους οποίους κανείς θα μπορούσε να ασκήσει κριτική στο θεωρητικό πλαίσιο της παραδοσιακής γνωσιακής επιστήμης, δηλαδή στην ΑΥΘΝ: ο πρώτος είναι να προτείνει μια διαφορετική έννοια νοητικής αναπαράστασης και ο δεύτερος να απορρίψει συνολικά την έννοια της νοητικής αναπαράστασης και να αναπτύξει έναν τύπο εξήγησης των νοητικών φαινομένων ο οποίος δεν βασίζεται σε αυτήν την έννοια. Στο πλαίσιο της γνωσιακής επιστήμης αλλά και της φιλοσοφίας του νου μπορεί κανείς να συναντήσει και τους δύο αυτούς εναλλακτικούς τρόπους αντίδρασης. Στον πρώτο εντάσσονται, όπως θα δούμε, οι υποστηρικτές των νοητικών ειδώλων, των νοητικών μοντέλων και των κατανομημένων αναπαραστάσεων. Στο δεύτερο εντάσσονται, όπως έχουμε ήδη αναφέρει, οι συμπεριφοριστές. Επίσης, κατά της έννοιας της νοητικής αναπαράστασης μπορούν να θεωρηθούν ότι τάσσονται και οι υποστηρικτές της οικολογικής ψυχολογίας του Gibson και, τα τελευταία χρόνια, οι υπέρμαχοι των δυναμικών συστημάτων.

Προτού προχωρήσουμε, ωστόσο, θα ήταν σκόπιμο να εξετάσουμε κάπως πιο συστηματικά αυτή καθαυτή την έννοια της αναπαράστασης γιατί μέχρι στιγμής την εκλάβαμε ως αυτονόητα κατανοητή. Στο επόμενο, λοιπόν, κεφάλαιο θα προβούμε σε μια συστηματική ανάλυση αυτής της έννοιας.

ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ

Μια πρώτη προσέγγιση της έννοιας της αναπαράστασης

3.1 Εισαγωγή

Στο δεύτερο κεφάλαιο είδαμε ότι οι νοητικές αναπαραστάσεις που υιοθετεί η ΑΥΘΝ δεν μπορούν να προσεγγιστούν από την οπτική του πρώτου προσώπου, δηλαδή δεν είναι συνειδητές. Είδαμε, επίσης, ότι οι εν λόγω αναπαραστάσεις είναι γλωσσικής μορφής. Ξεκινάμε το παρόν κεφάλαιο με αυτές τις δύο διαπιστώσεις προκειμένου να κάνουμε μια σημαντική διευκρίνιση: Όπως αναφέραμε και προηγουμένως, οι νοητικές αναπαραστάσεις που υιοθετεί η ΑΥΘΝ δεν πρέπει να συγχέονται με τις (υποκειμενικές) παραστάσεις (Vorstellungen) που συναντάμε στη νεότερη φιλοσοφία⁹⁶. Μέσα σε αυτή την παράδοση, όροι που ονομάζουν είδη παραστάσεων είναι, μεταξύ άλλων, οι ακόλουθοι: "ιδέα", "εντύπωση" (impression), εποπτεία (αγγλικά: "intuition", γερμανικά: "Anschauung"), "έννοια" (αγγλικά: "concept", γερμανικά: "Begriff"), "νοητικό είδωλο"⁹⁷ (mental image). Ειδικότερα, στον Βρετανικό Εμπειρισμό κάθε είδους νοητικό φαινόμενο συναρτάται με την κατοχή ή τον χειρισμό ιδεών. Για παράδειγμα, ο Locke (1984, βιβλίο II, κεφάλαιο VIII, σελ. 8) δηλώνει: "Οπδήποτε αντιλαμβάνεται ο νους ή οπδήποτε αποτελεί άμεσα αντικείμενο αισθητηριακής αντίληψης, σκέψης ή νόησης, το αποκαλώ ιδέα"⁹⁸. Έτσι, η επαφή του υποκειμένου με τον κόσμο κατανοείται ως μια έμμεση πρόσβαση στον κόσμο που πραγματοποιείται πάντα με τη διαμεσολάβηση παραστάσεων: Αυτή η θέση συνέχισε να υποστηρίζεται και μέσα στο πλαίσιο της αναλυτικής φιλοσοφίας από αρκετούς φιλοσόφους μέχρι περίπου τα μέσα του εικοστού αιώνα. Για παράδειγμα, ο Ράσσελ (1963, σελ. 11) επισημαίνει: "Έτσι, καθίσταται φανερό ότι το πραγματικό τραπέζι, αν υπάρχει κάτι τέτοιο, δεν είναι το ίδιο με αυτό του οποίου έχουμε άμεσα την εμπειρία μέσω της όρασης ή της αφής ή της ακοής. Το πραγματικό τραπέζι, αν υπάρχει τέτοιο, δεν μας γίνεται με κανέναν τρόπο άμεσα γνωστό, αλλά πρέπει να το συναγάγουμε από ό,τι μας είναι άμεσα γνωστό ... Ας δώσουμε το όνομα 'αισθητηριακά δεδομένα' στα πράγματα εκείνα που μας είναι άμεσα γνωστά στις αισθήσεις: στα χρώματα, τους ήχους, τις οσμές, τη σκληρότητα, την τραχύτητα κ.ο.κ."⁹⁹. Οι παραστάσεις, δηλαδή οι οντότητες που μεσολαβούν μεταξύ του υποκειμένου και του κόσμου, αποκτούν, λοιπόν,

⁹⁶ Οι διαφορές μεταξύ νοητικών αναπαραστάσεων και (υποκειμενικών) παραστάσεων που θα αναφερθούν στη συνέχεια δικαιολογούν την υιοθέτηση διαφορετικών όρων για την αναφορά σε αυτές τις δύο έννοιες.

⁹⁷ Για τη μεταφραστική επιλογή βλ. υποσημείωση 222.

⁹⁸ Βλ., επίσης, Tye 1991, σελ. 5: "... αντίθετα προς τον Καρτέσιο και τον Hobbes, οι βρετανοί εμπειριστές υποστήριξαν ότι κάθε είδους σκέψη συνίσταται στον χειρισμό είτε απλών νοητικών ειδώλων που πηγάζουν από την αισθητηριακή εμπειρία είτε σύνθετων νοητικών ειδώλων που κατασκευάζονται από αυτά τα απλά νοητικά είδωλα".

⁹⁹ Βλ., επίσης, Moore 1913, Ayer 1956 κ.α.

στον Εμπειρισμό του 20^{ου} αιώνα ένα νέο όνομα: "αισθητηριακά δεδομένα"¹⁰⁰ (sense data)¹⁰¹. Συμπερασματικά, οι παραστάσεις δίδονται στο επίπεδο της συνείδησης, δηλαδή αφορούν το φαινομενολογικό επίπεδο.

Οι νοητικές αναπαραστάσεις από την άλλη μεριά είναι ασυνείδητες: δεν μπορούν να δοθούν στη συνείδηση του υποκειμένου που τις φέρει. Μάλιστα, δεν είναι ασυνείδητες κατά τον τρόπο που μια φροϋδική επιθυμία μπορεί να είναι ασυνείδητη, διότι μια ασυνείδητη επιθυμία μπορεί δύναμι να γίνει συνειδητή. Οι φοντοριανές νοητικές αναπαραστάσεις είναι *ουσιωδώς* ασυνείδητες. Πρόσβαση στις νοητικές αναπαραστάσεις δεν έχει το υποκείμενο αλλά αυτόνομο γνωστικά υποσυστήματα του υποκειμένου. Υπό αυτή την έννοια, οι νοητικές αναπαραστάσεις δεν είναι απλώς ασυνείδητες αλλά, όπως επισημάναμε στο δεύτερο κεφάλαιο, υποπροσωπικές (subpersonal)¹⁰².

Επίσης, σε αντίθεση προς τις νοητικές αναπαραστάσεις, οι παραστάσεις δεν εξηγούνται με βάση το μοντέλο της γλώσσας. Ο λόγος είναι ότι η φιλοσοφία της συνείδησης προηγείται της φιλοσοφίας της γλώσσας στο Βρετανικό Εμπειρισμό και γι' αυτό οι έννοιες που χρησιμοποιούνται για την ανάλυση της συνείδησης δεν μπορούν να προέρχονται από μια οντολογικά και εξηγητικά παράγωγη περιοχή όπως η γλώσσα.

Στη σύγχρονη φιλοσοφία, ωστόσο, η κριτική του ψυχολογισμού επέτρεψε την ανάπτυξη μιας θεωρίας της γλωσσικής σημασίας αποσυνδεδεμένη από τη συνείδηση. Η ανάπτυξη της μαθηματικής λογικής και η συνακόλουθη πρόοδος στις μεθόδους τυποποίησης της φυσικής

¹⁰⁰ Βέβαια τα αισθητηριακά δεδομένα είναι κάτι πολύ διαφορετικό από μια εσωτερική εικόνα που δίδεται στο "μάτι" του νου, ωστόσο αυτό δεν εμπόδισε τους υποστηρικτές των αισθητηριακών δεδομένων να τα εκλάβουν ως παραστάσεις στο βαθμό που λειτουργούν ως υποκατάστατα των εξωτερικών αντικειμένων και των ιδιοτήτων τους.

¹⁰¹ Μια σημαντική διευκρίνιση που πρέπει να γίνει είναι ότι στην *παρούσα εργασία* περιοριζόμαστε στην αντίληψη περί παραστάσεων-ιδεών που κυριαρχούσε στον Βρετανικό Εμπειρισμό και στον λογικό Εμπειρισμό του 20^{ου} αιώνα, σύμφωνα με την οποία οι ιδέες νοούνται ως εσωτερικά αντικείμενα. Σε αυτή την έννοια της παράστασης θα ασκήσουμε κριτική στο όγδοο κεφάλαιο (ενότητα 8.2.2.2). Εκτός από την προαναφερθείσα αντίληψη, σύμφωνα με τον McRae (1965, σελ. 175), υπάρχουν δύο ακόμα εναλλακτικές αντιλήψεις στους επιγόνους του Descartes κατά τον 17^ο αιώνα: (α) "Οι τρεις αντιλήψεις είναι (α) ότι μια ιδέα είναι ένα αντικείμενο (Malebranche, Locke, Berkeley): (β) ότι μια ιδέα είναι ένα ενέργημα (Spinoza, Arnauld): (γ) ότι μια ιδέα είναι μια προδιάθεση (Leibniz)". Επίσης, για την έννοια της παράστασης στον Descartes και τον Leibniz βλέπε Anapolitanos 1998 (πρώτο κεφάλαιο). Το πώς ακριβώς κατανοούνται οι παραστάσεις στο σύνολο της νεώτερης φιλοσοφίας αποτελεί ένα ανοικτό ερευνητικό ζήτημα το οποίο υπερβαίνει το πλαίσιο της διατριβής.

¹⁰² Περισσότερα επ' αυτού στο όγδοο κεφάλαιο.

γλώσσας θεωρήθηκε από πολλούς ως η καλύτερη απόδειξη για το ότι η γλώσσα μπορεί να μελετηθεί ανεξάρτητα από τη συνείδηση. Η ΑΥΘΝ εμπνέεται ακριβώς από αυτό το φιλοσοφικό υπόδειγμα και υπό μια έννοια αποτελεί την κορύφωση της "γλωσσικής στροφής": πρόκειται για το *εγχείρημα που επιδιώκει να εξηγηθούν τα ίδια τα νοητικά φαινόμενα με βάση τη γλώσσα και τις γλωσσικές αναπαραστάσεις.*

3.2 Προβλήματα με τη δυνατότητα πρόσβασης στις νοητικές αναπαραστάσεις

Οι νοητικές αναπαραστάσεις δεν αποτελούν το πιο τυπικό παράδειγμα αναπαραστάσεων. Δεν είναι καν συνήθη φυσικά αντικείμενα. Πρόκειται για οντότητες που, όπως επισημάναμε, δεν μπορούν να δοθούν στη συνείδηση του υποκειμένου που τις φέρει. Ούτε, όμως, όντας κρυμμένες κάπου μέσα στο κεφάλι, είναι προσβάσιμες από τους (συνήθεις) εξωτερικούς παρατηρητές. Μάλιστα, απομένει να αποδειχθεί το εάν, έστω και κατ' αρχήν, οι νοητικές αναπαραστάσεις μπορούν να δοθούν σε πιο εξειδικευμένους παρατηρητές, π.χ. νευροφυσιολόγους. Φυσικά, θα μπορούσε κανείς να ισχυρισθεί ότι οι νοητικές αναπαραστάσεις αποτελούν κάποιο είδος θεωρητικών οντοτήτων όπως τα ηλεκτρόνια και τα quarks. Ωστόσο, αντίθετα από τα ηλεκτρόνια και τα quarks, αυτές δεν γίνονται αντιληπτές, όχι εξαιτίας του μεγέθους τους, αλλά εξαιτίας του γεγονότος ότι δεν γνωρίζουμε καν το τι ψάχνουμε. Εάν οι νοητικές αναπαραστάσεις έχουν πράγματι μια γλωσσική μορφή, όπως διατείνεται ο Fodor, η άγνοιά μας για το τι να ψάξουμε μοιάζει, εκ πρώτης όψεως, να οφείλεται στο ότι δεν γνωρίζουμε τη γλώσσα του νου: Μόλις αποκωδικοποιήσουμε αυτή τη γλώσσα, όπως οι αρχαιολόγοι αποκωδικοποιούν τα ιερογλυφικά, το πρόβλημα θα έχει ξεπεραστεί¹⁰³. Πόσο εύλογη είναι αυτή η εικόνα; Το πρώτο πράγμα που πρέπει να επισημανθεί είναι ότι στηρίζεται σε μια μεταφορά: οι νοητικές αναπαραστάσεις συγκρίνονται με τα γραπτά σύμβολα κάποιας άγνωστης φυσικής γλώσσας. Πόσο κυριολεκτικά θα πρέπει να πάρουμε αυτή τη σύγκριση; Το ερώτημα εξαρτάται από το τι θεωρεί κανείς φυσική γλώσσα και γλωσσική αναπαράσταση. Αυτό είναι ένα τεράστιο και πολύ σύνθετο ζήτημα που μόνο σε μια πρώτη προσέγγιση θα μπορούσαμε να εξετάσουμε. Για την απλοποίηση του προβλήματος, είναι σκόπιμο να προσδιορίσουμε αρχικά το πότε μια οντότητα θεωρείται

¹⁰³ Μια ενδεχόμενη δυσαναλογία είναι ότι στην περίπτωση της αρχαίας γλώσσας έχουν ήδη ανακαλυφθεί τα ιερογλυφικά και το πρόβλημα είναι να αποκωδικοποιηθούν, ενώ στην περίπτωση της γλώσσας του νου δεν έχουν ανακαλυφθεί ούτε καν οι "λέξεις" και οι "εκφράσεις" αυτής της γλώσσας. Ωστόσο, ένας υποστηρικτής της υπόθεσης της γλώσσας της σκέψης θα μπορούσε να αντιτείνει ότι διαθέτουμε πλέον αρκετά εξελιγμένες τεχνικές απεικόνισης της λειτουργίας του εγκεφάλου οι οποίες μας παρέχουν αυτό το λεξιλόγιο.

αναπαράσταση και ποια είναι τα κυριότερα χαρακτηριστικά της, ανεξάρτητα από το αν πρόκειται για γλωσσική ή άλλου τύπου αναπαράσταση. Μεθοδολογικά, είναι σκόπιμο να βασιστούμε στις συνήθεις εξωτερικές αναπαραστάσεις μια και αυτές, σε αντίθεση με τις νοητικές αναπαραστάσεις, αποτελούν το πλέον οικείο είδος αναπαραστάσεων. Έτσι, θα εντοπίσουμε τις κυριότερες διακρίσεις που γίνονται όσον αφορά τις εξωτερικές αναπαραστάσεις. Οι διακρίσεις αυτές θα μας παράσχουν ταυτόχρονα τρόπους ταξινόμησης των αναπαραστάσεων. Έτσι, έχοντας προσδιορίσει τα βασικά χαρακτηριστικά και τις κυριότερες διακρίσεις που αποδίδονται στις εξωτερικές αναπαραστάσεις θα αποκτήσουμε τα εφόδια για να επιμερίσουμε το πολυσύνθετο πρόβλημα των νοητικών αναπαραστάσεων σε συγκεκριμένα υποπροβλήματα κάθε ένα από τα οποία θα αποτελέσει το αντικείμενο ξεχωριστού κεφαλαίου.

3.3 Το πιο οικείο είδος αναπαραστάσεων: εξωτερικές αναπαραστάσεις

Οι εξωτερικές αναπαραστάσεις είναι δημόσιες οντότητες οι οποίες εμπλέκονται σχεδόν σε κάθε ανθρώπινη δραστηριότητα. Μάλιστα, είμαστε τόσο εξοικειωμένοι με την κατανόηση και χρήση των εξωτερικών αναπαραστάσεων που, τις περισσότερες φορές, δεν τις προσέχουμε καν. Υπάρχουν πολλές διαφορετικές μορφές εξωτερικών αναπαραστάσεων: λέξεις, προτάσεις, σκίτσα, εικόνες, χάρτες, διαγράμματα, μοντέλα κλίμακας, οδικά σήματα κ.λπ. Υπάρχει άραγε κάτι κοινό ανάμεσα σε όλα αυτά τα διαφορετικά είδη εξωτερικών αναπαραστάσεων; Σε αυτό το κεφάλαιο θα επιχειρήσουμε να δώσουμε έναν προκαταρκτικό ορισμό του γένους αναπαράσταση, προϋποθέτοντας καταρχάς ότι τα διάφορα είδη αναπαραστάσεων συνιστούν μια κλασική κατηγορία και ότι, υπό αυτή την έννοια, υπάρχει ένα σύνολο χαρακτηριστικών που είναι κοινό σε όλα τα είδη αναπαραστάσεων. Η παραδοχή αυτή είναι απαραίτητη μεθοδολογικά, τουλάχιστον για να ξεκινήσουμε να συζητάμε το πρόβλημα των αναπαραστάσεων ως ενιαίο πρόβλημα. Φυσικά, αυτό κάθε άλλο παρά σημαίνει ότι δεχόμαστε a priori την ισχύ της συγκεκριμένης παραδοχής.

Τι είναι, λοιπόν, μια αναπαράσταση; Σε μια πρώτη προσέγγιση είναι κάτι που στέκεται στη θέση, παίρνει τη θέση κάποιου άλλου πράγματος, δηλαδή αντικαθιστά, αντιπροσωπεύει κάποιο άλλο πράγμα¹⁰⁴: η εικόνα του Σωκράτη παίρνει τη θέση του Σωκράτη, το ίδιο κάνει, αν και με διαφορετικό τρόπο, η λέξη "Σωκράτης". Ας δεχτούμε λοιπόν προσωρινά και ως αφετηρία τον ακόλουθο ορισμό:

¹⁰⁴ Αυτή η αντίληψη για το τι σημαίνει "αναπαράσταση" απαντά ευρύτατα στο πλαίσιο της σύγχρονης φιλοσοφίας του νου.

Ορισμός 1:	Ένα αντικείμενο Α συνιστά αναπαράσταση κάποιου πράγματος Ο αν μπορεί να πάρει τη θέση του Ο ή, με άλλα λόγια, να αντικαταστήσει το Ο.
------------	---

Αυτό τη θέση του οποίου παίρνει η αναπαράσταση ονομάζεται *περιεχόμενο* (content) της αναπαράστασης¹⁰⁵. Μια πρώτη, λυσιτελής, βασική διάκριση που πρέπει να αναλύσουμε είναι αυτή μεταξύ μορφής και περιεχομένου της αναπαράστασης.

3.3.1 Βασικές διακρίσεις

3.3.1.1 Η μορφή της αναπαράστασης και ο φορέας της αναπαράστασης

Η μορφή της αναπαράστασης σχετίζεται με τη φυσική της εμφάνιση. Ωστόσο, φαίνεται ότι δεν είναι όλα τα φυσικά χαρακτηριστικά μιας αναπαράστασης σημαντικά για τον προσδιορισμό της μορφής της. Για παράδειγμα, το χρώμα, το μέγεθος ή το υλικό της γλωσσικής αναπαράστασης "Σωκράτης", σε αντίθεση με το σχήμα της, δεν αφορούν τη μορφή της αναπαράστασης. Αυτή η διαπίστωση δεν φαίνεται, όμως, να ισχύει τόσο έντονα για άλλα είδη αναπαράστασεων. Έτσι, σε μια εικόνα εκτός από το σχήμα έχει σημασία, μεταξύ άλλων, το πάχος των γραμμών, το χρώμα και το σχετικό μέγεθος των εμφανιζόμενων σχημάτων. Αντίθετα, δεν παίζει, συνήθως, κανένα σημαίνοντα ρόλο η υφή του υλικού της εικόνας ή τα χαρακτηριστικά της μορφής της πίσω όψης της εικόνας. Αν τώρα φτάσουμε στα διάφορα μοντέλα κλίμακας¹⁰⁶, τότε διαπιστώνουμε ότι ακόμα περισσότερα χαρακτηριστικά της φυσικής εμφάνισής τους παίζουν σημαίνοντα ρόλο. Επίσης, πρέπει να παρατηρήσουμε ότι τα σημαίνοντα χαρακτηριστικά δεν περιορίζονται κατ' ανάγκη μόνο στα οπτικώς διαπιστώσιμα χαρακτηριστικά, δηλαδή στα χαρακτηριστικά της οπτικής εμφάνισης της αναπαράστασης - αρκεί να σκεφτούμε τον προφορικό λόγο και το σύστημα γραφής των τυφλών. Τέλος, όταν μιλάμε για σημαίνοντα φυσικά χαρακτηριστικά μιας αναπαράστασης δεν υπάρχει κανένας εγγενής λόγος που να υπαγορεύει ότι τα εν λόγω χαρακτηριστικά πρέπει να ανήκουν στην εξωτερική επιφάνεια της αναπαράστασης. Για παράδειγμα, σε ορισμένα μοντέλα μπορεί να παίζει σημαίνοντα ρόλο το υλικό από το οποίο είναι κατασκευασμένα καθώς και η εσωτερική τους δομή. Πάντως, όσο μεγάλο κι αν γίνει το ποσοστό των φυσικών χαρακτηριστικών που έχουν σημαίνοντα ρόλο σε μια αναπαράσταση ποτέ

¹⁰⁵ Πρόκειται για ό,τι προθεωρητικά στα προηγούμενα κεφάλαια ονομάζαμε "σημασία" (meaning).

¹⁰⁶ Ένα χαρακτηριστικό παράδειγμα είναι οι σμικρύνειες αεροπλάνων που χρησιμοποιούνται από τους αεροναυπηγούς για να μελετήσουν την πτητική συμπεριφορά των κανονικών αεροπλάνων.

δεν φτάνουμε σε αναπαραστάσεις όπου το σύνολο των φυσικών τους χαρακτηριστικών έχουν σημαίνοντα ρόλο. Για παράδειγμα, αν ο σημαίνων ρόλος βασίζεται σε μια σχέση ομοιότητας μεταξύ των φυσικών χαρακτηριστικών της αναπαράστασης και του αναπαριστώμενου, τότε σε περίπτωση ομοιότητας του συνόλου των φυσικών χαρακτηριστικών μιας οντότητας με εκείνα μιας άλλης, κάνουμε λόγο για *αντίγραφο* και όχι για αναπαράσταση και αναπαριστώμενο. Άρα λοιπόν, όταν μιλάμε για αναπαραστάσεις μόνο ένα υποσύνολο των φυσικών τους χαρακτηριστικών παίζουν σημαίνοντα ρόλο.

Για να διακρίνουμε τη (σημαίνουσα) μορφή από τη συνολική υλική υπόσταση της αναπαράστασης, θα αναφερόμαστε στην τελευταία με το όρο "φορέας της αναπαράστασης". Το τι ακριβώς είναι και από τι καθορίζεται η μορφή της αναπαράστασης αποτελεί ένα μεγάλο και ανοιχτό ερώτημα του οποίου, προς το παρόν, θα αναβάλουμε την εξέταση. Αυτό που πρέπει να κρατήσουμε εδώ είναι απλώς η διάκριση μεταξύ μορφής και φορέα, το ότι, δηλαδή, μια εξωτερική αναπαράσταση, στο βαθμό που είναι φυσικό αντικείμενο, αποτελεί κάποιον (φυσικό) φορέα του οποίου ορισμένα μόνο χαρακτηριστικά συνιστούν αυτό που ονομάζουμε *μορφή* της αναπαράστασης. Συγκεκριμένα, από όλα τα χαρακτηριστικά του φορέα μιας αναπαράστασης εκείνα που συνιστούν τη μορφή της είναι όσα σχετίζονται με κάποιο τρόπο με το περιεχόμενο της αναπαράστασης¹⁰⁷. Ο τρόπος σχετισμού, όπως θα δούμε, διαφέρει ανάλογα με το είδος της αναπαράστασης. Για παράδειγμα, στις εικόνες και τα μοντέλα κλίμακας το είδος της σχέσης θεωρείται ότι είναι σχέση ομοιότητας. Θα επανέλθουμε σε αυτό το ζήτημα αφού πρώτα αναφερθούμε αναλυτικότερα στην έννοια του περιεχομένου της αναπαράστασης.

Ποια είναι όμως η σκοπιμότητα εισαγωγής της διάκρισης μεταξύ μορφής και φορέα της αναπαράστασης; Γιατί αυτή η διάκριση μάς είναι χρήσιμη για τον προσδιορισμό της έννοιας της αναπαράστασης; Η απάντηση πιστεύουμε είναι προφανής: με αυτόν τον τρόπο κατορθώνουμε να απομονώσουμε τα χαρακτηριστικά ενός υλικού σώματος που εμπλέκονται στην αναπαραστασιακή σχέση από όλα τα υπόλοιπα χαρακτηριστικά του¹⁰⁸.

¹⁰⁷ Στο σημείο αυτό είναι σκόπιμο να επισημάνουμε ότι η έννοια της αναπαράστασης χρησιμοποιείται γενικώς με δύο, τουλάχιστον, τρόπους, γεγονός που δυσχεραίνει ακόμα περισσότερο το πρόβλημα γύρω από το τι είναι αναπαράσταση. Συγκεκριμένα, μπορεί να αναφέρεται είτε στον φορέα μεμονωμένα, είτε στην ενότητα φορέα και περιεχομένου. Αυτή η αμφισβημία του όρου "αναπαράσταση" είναι πολύ δύσκολο να αποφευχθεί στην πράξη.

¹⁰⁸ Χάρη στη μη ταύτιση αναπαραστασιακής μορφής και αναπαραστασιακού φορέα οι αναπαραστάσεις, όπως και οι υπολογιστές (βλ. ενότητες 1.3.6 και 2.5), παρουσιάζουν μια σχετική αυτονομία από το μέσο στο οποίο "υλοποιούνται". Αυτή η αυτονομία, ωστόσο, ποικίλει σημαντικά ανάλογα με το είδος της αναπαράστασης: όσο περισσότερα χαρακτηριστικά του αναπαραστασιακού φορέα

Ενσωματώνοντας τη διάκριση φορέα/μορφής στον προκαταρκτικό ορισμό της έννοιας της αναπαράστασης μπορούμε να πούμε ότι:

Ορισμός 2:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου πράγματος O , δηλαδή έχει ως περιεχόμενο το O , αν η μορφή M του A μπορεί να πάρει τη θέση του O .
------------	---

3.3.1.2 Το περιεχόμενο της αναπαράστασης και η διάκριση νοήματος/αναφοράς

Παραμένοντας πάντα σε ένα πολύ γενικό επίπεδο προσέγγισης, μπορούμε να παρατηρήσουμε ότι οι αναπαραστάσεις φαίνεται να είναι το μόνο είδος φυσικών αντικειμένων που διαθέτουν (αναπαραστασιακό) περιεχόμενο. Ωστόσο, αυτή η διαπίστωση κάθε άλλο παρά ξεκαθαρίζει την κατάσταση. Τα κρίσιμα ερωτήματα που θα πρέπει να διασαφηνισθούν είναι, εν προκειμένω, τι σημαίνει για κάτι να έχει περιεχόμενο και πώς κάτι τέτοιο καθίσταται δυνατό. Στην παρούσα ενότητα θα εξετάσουμε το πρώτο ερώτημα, ενώ το δεύτερο θα εξετασθεί στην επόμενη ενότητα. Η επίλυση αυτής της σειράς δεν είναι τυχαία. Για να απαντήσει κανείς στο δεύτερο ερώτημα πρέπει με κάποιο τρόπο να έχει προηγουμένως απαντήσει στο πρώτο: προκειμένου να εξετάσουμε πώς ένα πράγμα αποκτά περιεχόμενο πρέπει ήδη να γνωρίζουμε τι σημαίνει για ένα πράγμα να έχει περιεχόμενο.

Έχουμε ήδη δώσει μια πρώτη απάντηση στο ερώτημα τι σημαίνει για κάτι να έχει περιεχόμενο: κάτι έχει περιεχόμενο αν "παίρνει τη θέση" κάποιου άλλου πράγματος. Ωστόσο, μια προσεκτικότερη εξέταση αυτής της απάντησης δείχνει ότι είναι ανεπαρκής. Πολλά πράγματα παίρνουν τη θέση άλλων πραγμάτων χωρίς τα πρώτα να θεωρούνται αναπαραστάσεις των δευτέρων. Για παράδειγμα, ένα ανταλλακτικό μπορεί να πάρει τη θέση ενός χαλασμένου τμήματος κάποιου μηχανισμού, αλλά κανείς δεν θα ονομάζε το ανταλλακτικό αναπαράσταση του χαλασμένου τμήματος.

Τι παραπάνω, λοιπόν, κάνει μια αναπαράσταση από το να παίρνει τη θέση κάποιου άλλου πράγματος; Μια ιδέα είναι ότι η αναπαράσταση λειτουργεί ως υποκατάστατο αυτού που αναπαριστά, ότι δηλαδή υποκαθιστά το αναπαριστώμενο. Ωστόσο, αυτό δεν φαίνεται να είναι ακριβές. Για παράδειγμα, στον κινηματογράφο συνηθίζεται κατά το γύρισμα μιας επικίνδυνης σκηνής να υποκαθίσταται ο ηθοποιός από κάποιον σωσία του. Ωστόσο, το γεγονός ότι ο σωσίας παίρνει τη θέση του ηθοποιού επειδή ο τελευταίος δεν μπορεί να συμμετάσχει στη σκηνή, δεν καθιστά τον σωσία αναπαράσταση του ηθοποιού. Προς απόδειξη αυτού αρκεί να παρατηρήσουμε ότι στην περίπτωση μιας (επιτυχημένης)

συνιστούν την αναπαραστασιακή μορφή, τόσο περιορίζεται η αυτονομία από το μέσο "υλοποίησης".

κινηματογραφικής σκηνής οι θεατές δεν καταλαβαίνουν καν ότι σε αυτή δεν συμμετέχει ο ίδιος ο ηθοποιός, ενώ στην περίπτωση μιας αναπαράστασης δεν συμβαίνει κάτι ανάλογο: μια φωτογραφία του ηθοποιού δεν εκλαμβάνεται ως ο ίδιος ο ηθοποιός. Επίσης, για να αναφέρουμε ένα διαφορετικό παράδειγμα, η γλυκόζη πολλές φορές χρησιμοποιείται ως υποκατάστατο της ζάχαρης ωστόσο αυτό δεν καθιστά την πρώτη αναπαράσταση της δεύτερης. Γενικά, ένα υποκατάστατο του Ο έχει ως λειτουργία να μην διακρίνεται άμεσα από το Ο, ενώ αντίθετα μια αναπαράσταση του Ο πάντα διακρίνεται από το Ο.

Κατά συνέπεια, το ερώτημα παραμένει: τι παραπάνω κάνει μια αναπαράσταση από το να παίρνει τη θέση κάποιου άλλου πράγματος; Ένας πρόσθετος περιορισμός που συχνά αναφέρεται είναι ότι οι αναπαραστάσεις υποκαθιστούν κάτι που δεν είναι παρόν. Ωστόσο, ούτε αυτή η συνθήκη είναι επαρκής: όταν κάποιος δεν μπορεί να παρευρεθεί σε μια συνάντηση, στέλνει έναν αντιπρόσωπό του, κάποιον που παίρνει τη θέση του προκειμένου να ενεργήσει για λογαριασμό του. Παρόλα αυτά δεν λέμε, συνήθως, ότι ο αντιπρόσωπος αποτελεί μια αναπαράσταση του ατόμου που δεν είναι χωρικά παρόν. Βέβαια, ο Peirce, ο φιλόσοφος που ανέπτυξε μια από τις σημαντικότερες θεωρίες αναπαράστασης, έχει διαφορετική γνώμη ως προς αυτό το ζήτημα¹⁰⁹. Επιχειρώντας να ορίσει την αναπαραστασιακή σχέση μέσω της έννοιας της υποκατάστασης λέει χαρακτηριστικά:

"Το να υποκαθιστά (stand for), [είναι]... το να βρίσκεται κάτι σε μια τέτοια σχέση προς κάτι άλλο ώστε για συγκεκριμένους σκοπούς να αντιμετωπίζεται από κάποιον νου ως εάν να ήταν αυτό το άλλο. Έτσι, ένας εκπρόσωπος ομιλητής (spokesman), ένας βουλευτής, ένας πληρεξούσιος δικηγόρος, ένας αντιπρόσωπος (agent), ένας βικάριος [επισκοπικός επίτροπος], ένα διάγραμμα, ένα σύμπλωμα, ένας μετρητής, μια περιγραφή, μια έννοια, μια προκείμενη, μια μαρτυρία, όλα αυτά αναπαριστούν κάτι άλλο με τους ποικίλους τρόπους τους στον νου όσων τα θεωρούν κατ' αυτόν τον τρόπο." (Peirce 1967, 2.272).

Ωστόσο, ο ίδιος ο Peirce σε ένα γράμμα του προς τη Lady Welby, το οποίο κατά πάσα πιθανότητα δεν της εστάλη ποτέ¹¹⁰, κάνει αυτοκριτική στην επιλογή του να θεωρήσει τους αντιπροσώπους ως είδος αναπαραστάσεων -σημείων, όπως αυτός τις ονομάζει. Αναφερόμενος στην επιλογή του να χρησιμοποιεί τη λέξη "representamen" ως συνώνυμο της λέξης "σημείο", επισημαίνει:

¹⁰⁹ Ένας λόγος που ο Peirce ενέταξε τους αντιπροσώπους στην κατηγορία των αναπαραστάσεων είναι ότι στα αγγλικά ο αντιπρόσωπος αναφέρεται με τη λέξη "representative".

¹¹⁰ Hardwick 1977, σελ. 189.

"Ο λόγος που προτιμούσα τον όρο "representamen" ήταν ότι θα φαινόταν πιο φυσικό να τον εφάρμοζα στους βουλευτές του νομοθετικού σώματος (representatives in legislature), σε αντιπροσώπους διαφόρων ειδών (deputies of various kinds) κ.λπ. ... Εξακολουθώ να δέχομαι ότι βοηθά στην κατανόηση του ορισμού η προσεκτική σύγκρισή του με τέτοιες περιπτώσεις. Αλλά σίγουρα αυτές αποκλίνουν από τον ορισμό, κατά το ότι αυτός απαιτεί ότι η δράση του σημείου ως τέτοιου δεν θα επηρεάσει το αναπαριστώμενο αντικείμενο. Ένας βουλευτής του νομοθετικού σώματος, αντίθετα, αναμένεται με τις λειτουργίες του να βελτιώσει τις συνθήκες των εκλογέων του και οποιοσδήποτε πληρεξούσιος δικηγόρος, ακόμα και αν δεν έχει καμία σύνεση, αναμένεται να επηρεάσει τις συνθήκες του εντολέα του. ... Θεώρησα το representamen ως κάτι που παίρνει τη θέση του πράγματος: αλλά το σημείο δεν είναι ένα υποκατάστατο. Ο Ernst Mach έπεσε και αυτός στην ίδια παγίδα."¹¹¹

Βλέπουμε, λοιπόν, ότι δεν αρκεί για να καταστεί μια οντότητα αναπαράσταση το να παίρνει τη θέση κάποιου άλλου πράγματος επειδή αυτό δεν είναι φυσικά παρόν. Αποδεικνύεται ότι έχει σημασία ο λόγος για τον οποίο η οντότητα παίρνει τη θέση του πράγματος που δεν είναι παρόν. Στο παράδειγμα με τον αντιπρόσωπο είναι ακριβώς ο σκοπός για τον οποίο γίνεται η αντιπροσώπηση αυτός που αποτρέπει το να θεωρηθεί ως αναπαράσταση ο αντιπρόσωπος. Επειδή, με άλλα λόγια, ο αντιπρόσωπος έχει ως αποστολή να ενεργήσει για λογαριασμό του εντολέα του, έχει τη δυνατότητα να επηρεάσει τον εντολέα του -κάτι που, σύμφωνα με τον Peirce, δεν συμβαίνει με τις αναπαραστάσεις.

Όμως η χωρική ή χρονική απουσία του περιεχομένου δεν αποτελεί ούτε αναγκαία συνθήκη για να είναι κάτι αναπαράσταση. Η προσωπογραφία ενός κοριτσιού που μόλις ολοκληρώνει ένας πλανόδιος ζωγράφος δεν παύει να αποτελεί αναπαράσταση επειδή το περιεχόμενο του πίνακα, δηλαδή το κορίτσι, βρίσκεται δίπλα στον πίνακα.

Τι είναι, λοιπόν, αυτό το παραπάνω που κάνει μια αναπαράσταση εκτός από το να παίρνει τη θέση κάποιου άλλου πράγματος; Η απάντηση θα πρέπει να αναζητηθεί στη λειτουργία που εξυπηρετεί μια αναπαράσταση με το να παίρνει τη θέση ενός άλλου πράγματος. Μια αναπαράσταση δεν παίρνει τη θέση κάποιου άλλου πράγματος για να το αναπληρώσει επειδή δεν είναι φυσικά παρόν. Η κυρίαρχη λειτουργία της αναπαράστασης είναι να αναφέρεται (refer) σε αυτό του οποίου παίρνει τη θέση. Επειδή ο όρος "αναφορά" έχει μια αμφισημία, δηλαδή μπορεί να αναφέρεται είτε στη σχέση της αναπαράστασης (δηλαδή, του αναπαραστασιακού φορέα) με το αναπαριστώμενο είτε στο ίδιο το αναπαριστώμενο, στο παρόν κείμενο θα χρησιμοποιούμε τον όρο

¹¹¹ Hardwick 1977, σελ. 193.

αποκλειστικά για να αναφερόμαστε στη σχέση της αναπαράστασης με το αναπαριστώμενο, ενώ στο ίδιο το αναπαριστώμενο θα αναφερόμαστε με τον όρο "αναφερόμενο" (referent).

Ωστόσο, χρειάζεται κάτι επιπλέον για να συμπληρωθεί η δομή της αναπαραστασιακής λειτουργίας. Μια αναπαράσταση είναι ένα ειδικού τύπου εργαλείο: χρησιμοποιείται για να αναφέρεται σε αυτό του οποίου παίρνει τη θέση. Όμως, όπως όλα τα εργαλεία, χρειάζεται κάποιον χρήστη προκειμένου να επιτελέσει τη λειτουργία της. Αν ο χρήστης γνωρίζει πώς να χρησιμοποιεί την αναπαράσταση, δηλαδή είναι επαρκής, τότε αυτή εκτελώνει την αναπαραστασιακή της λειτουργία που είναι να αναφέρεται σε ένα άλλο αντικείμενο. Φυσικά, θα ήταν πολύ περιοριστικό το να ισχυριστεί κανείς ότι ένα αντικείμενο καθίσταται αναπαράσταση *μόνο τη στιγμή* που χρησιμοποιείται από κάποιον. Γι' αυτό, μιλώντας για χρήση θα εννοούμε απλώς τη δυνατότητα χρήσης και όχι την τρέχουσα χρήση¹¹².

Ενσωματώνοντας στον προκαταρκτικό ορισμό της έννοιας της αναπαράστασης τούς επιπλέον περιορισμούς που επισημάναμε, φτάνουμε στην ακόλουθη διατύπωση:

Ορισμός 3:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου πράγματος O, δηλαδή έχει ως περιεχόμενο το O, αν για κάποιον επαρκή χρήστη X η μορφή M του A μπορεί να αναφέρεται στο O.
------------	---

Ας εξετάσουμε όμως τώρα λίγο πιο συγκεκριμένα την έννοια της αναφοράς σε σχέση με τον χρήστη. Για την απλοποίηση του προβλήματος με τον όρο "χρήστης" θα αναφερόμαστε αποκλειστικά σε αυτόν που "δέχεται" την αναπαράσταση και όχι σε αυτόν που την παράγει. Τι ακριβώς λοιπόν σημαίνει ότι για κάποιον επαρκή χρήστη ένα πράγμα αναφέρεται σε κάποιο άλλο; Μια πιθανή απάντηση είναι ότι ο χρήστης αντιδρά στην παρουσία της αναπαράστασης του O, όπως θα αντιδρούσε στην παρουσία του ίδιου του O¹¹³. Ωστόσο, αυτός ο προσδιορισμός

¹¹² Δηλαδή, θα χρησιμοποιήσουμε υποθετικό λόγο αντιγεγονικού είδους ή, όπως από ορισμένους λέγεται, αντίθετο του πραγματικού (counterfactual conditional).

¹¹³ Αυτή είναι σύμφωνα με τον Tugendhat (1976, σελ. 289-290) η θέση της πραγματιστικής-συμπεριφοριστικής παράδοσης. Ο Tugendhat επικαλείται συγκεκριμένα τον Morris ο οποίος, μαζί με τον Peirce, έθεσε τις βάσεις για την ανάπτυξη μιας γενικής θεωρίας περί σημείων, της σημειωτικής. Ο Morris επισημαίνει συγκεκριμένα ότι "από την οπτική γωνία των συμπεριφοριστών, το να λαμβάνει κανείς υπόψη του το D μέσω της παρουσίας του S έχει να κάνει με το να αποκρίνεται στο D δυνάμει της απόκρισής του στο S" (Morris 1938, σελ. 84). Βλ., επίσης, Morris (1938, σελ. 109): "... το διερμηνεύον (interpretant) είναι η συνήθεια του οργανισμού να αποκρίνεται εξαιτίας του φορέα του σημείου, σε απόντα αντικείμενα που είναι σημαντικά σε μια παρούσα προβληματική κατάσταση ως εάν να ήταν παρόντα".

φαίνεται να ταιριάζει περισσότερο σε ό,τι η Judge (1983) ονομάζει σήματα (signals)¹¹⁴. Στο ζωικό βασίλειο είναι συχνό φαινόμενο ζώα ενός είδους να τρέπονται σε φυγή όχι μόνο όταν αντιλαμβάνονται κάποιον εχθρό τους, αλλά και όταν ακούσουν μια συγκεκριμένη κραυγή από ένα ζώο του είδους τους που τα προειδοποιεί για την παρουσία εχθρού. Η κραυγή αποτελεί για τα ζώα ένα σήμα κινδύνου για την ύπαρξη αρπακτικού στη γύρω περιοχή. Τα ζώα αντιδρούν στην κραυγή όπως θα αντιδρούσαν αν αντιλαμβάνονταν το ίδιο το αρπακτικό. Η αντίδραση αυτή φαίνεται να είναι αντανακλαστική και πάντα η ίδια¹¹⁵. Τα ζώα δεν έχουν τη δυνατότητα να αντιδράσουν διαφορετικά. Ωστόσο, οι αναπαραστάσεις με την πλήρη έννοια του όρου δεν προκαλούν πάντα την ίδια αντίδραση: αν ξαφνικά ακούσω την κραυγή "λιοντάρι!" δεν είναι προκαθορισμένος ο τρόπος που θα αντιδράσω. Ανάλογα με το πλαίσιο στο οποίο βρίσκομαι (στο σπίτι μου, στον ζωολογικό κήπο του Κονγκό, στο διαστημικό λεωφορείο κ.λπ.), ανάλογα με την προϊστορία μου (θηριοδαμαστής, θύμα επίθεσης λιονταριού, πρωταθλητής του κουνγκ-φου κ.λπ.), ανάλογα με την τρέχουσα κατάσταση μου (οπλισμένος κυνηγός λιονταριών, κυνηγημένος από λιοντάρι, τραυματισμένος κ.λπ.) κ.λπ. θα αντιδράσω διαφορετικά. Με άλλα λόγια, οι χρήστες αναπαραστάσεων με την πλήρη έννοια του όρου δεν αντιδρούν πάντα με τον ίδιο τρόπο στην παρουσία μιας αναπαράστασης. Η αντίδρασή τους ρυθμίζεται, επίσης, από τη προϊστορία τους και το πλαίσιο εντός του οποίου χρησιμοποιούν την αναπαράσταση.

Όμως, θα μπορούσε κανείς να αντιτείνει ότι ακόμα και τα ζώα, όταν ακούν ένα σήμα κινδύνου δεν αντιδρούν πάντα με τον ίδιο τρόπο. Αν, για παράδειγμα, το σήμα για την ύπαρξη λιονταριού προέρχεται από το βάθος ενός γκρεμού, τότε, ενδεχομένως, τα ζώα που βρίσκονται στο χέλιος του γκρεμού να μην τραπεούν σε φυγή. Ή, ο σκύλος που έχει εκπαιδευτεί να αντιδρά στον ήχο ενός βομβητή όπως στην παρουσία φαγητού, ενδεχομένως, να μην αντιδράσει καθόλου αν είναι χορτάτος¹¹⁶. Κατά συνέπεια, η διαφορά ανάμεσα στα σήματα και τις αναπαραστάσεις με την πλήρη έννοια του όρου δεν έγκειται στο ότι τα πρώτα προκαλούν μια μοναδική και ανελαστική αντίδραση, ενώ οι δεύτερες ένα πλήθος

¹¹⁴ Βλ. Judge 1983, σελ. 40 και Taylor 1985α οι οποίοι παραπέμπουν στον Piaget (1946/1962). Βλ., επίσης, Sebeok 1994, σελ.22: "Το σήμα είναι ένα σημείο το οποίο μηχανικά (φυσικά) ή συμβατικά (τεχνητά) προκαλεί κάποια αντίδραση από τη μεριά του δέκτη".

¹¹⁵ Η συγκεκριμένη συμπεριφορά δεν είναι πάντα έμφυτη, μπορεί να είναι και επίκτητη. Ο συμπεριφορισμός μάς έχει παράσχει αμέτρητα παραδείγματα επίκτητων αντιδράσεων σε σήματα. Αρκεί να σκεφτεί κανείς τα σκυλιά του Ραβλον που έμαθαν να αντιδρούν στον ήχο ενός βομβητή όπως αντιδρούσαν στον εντοπισμό τροφής.

¹¹⁶ Να διευκρινίσουμε ότι αναφερόμαστε στον τρόπο που τα ανώτερα θηλαστικά αντιδρούν στα σήματα, διότι όσο προχωρούμε προς τους κατώτερους οργανισμούς αυτή η ελαστικότητα στην αντίδραση χάνεται.

αντιδράσεων ανάλογα με την προϊστορία του χρήστη και το πλαίσιο χρήσης. Σε τι έγκειται τότε η διαφορά σημάτων και αναπαραστάσεων με την πλήρη έννοια του όρου;

Τα σήματα πρωταρχικά είναι σήματα για την παρουσία κάποιου πράγματος και ως τέτοια πρέπει να κατανοηθούν. Υπό αυτή την έννοια, τα σήματα για την παρουσία λιονταριού προκαλούν στα ζώα την ίδια συμπεριφορά που θα προκαλούσε ο εντοπισμός της οσμής του λιονταριού ή η οπτική αντίληψη του λιονταριού. Με άλλα λόγια, τα σήματα αντιμετωπίζονται από τα ζώα που τα λαμβάνουν ως ένα χαρακτηριστικό του ίδιου του λιονταριού, ως κάτι που ανήκει εγγενώς σε αυτό¹¹⁷. Έτσι, η αντίληψη του σήματος προκαλεί το ίδιο σύνολο συμπεριφορών που θα προκαλούσε η αντίληψη κάποιου φυσικού χαρακτηριστικού του λιονταριού¹¹⁸. Αντίθετα, σε μια αναπαράσταση -με την πλήρη έννοια του όρου- ενός λιονταριού, η αναπαραστασιακή μορφή της δεν εκλαμβάνεται ως κάποιο φυσικό χαρακτηριστικό του ίδιου του λιονταριού¹¹⁹. Η αναπαραστασιακή μορφή εκλαμβάνεται ως κάτι που αναφέρεται στο λιοντάρι ως *μη σωματικά παρόν*. Η θέση αυτή δεν έρχεται σε αντίθεση με την παρατήρηση που έγινε λίγες σελίδες πιο πάνω, ότι, δηλαδή, η αναπαράσταση δεν μπορεί να θεωρηθεί ως κάτι το οποίο παίρνει τη θέση κάποιου πράγματος που δεν είναι παρόν. Διότι αυτό που είναι σημαντικό, εν προκειμένω, δεν είναι το εάν ή όχι το αναπαριστώμενο είναι *όντως παρόν*, αλλά το ότι *εκλαμβάνεται από τον χρήστη ως μη παρόν*. Συγκεκριμένα, στο παράδειγμα με την προσωπογραφία του κοριτσιού, το αναπαριστώμενο κορίτσι παρουσιάζεται ως *μη σωματικά παρόν στον χρήστη της προσωπογραφίας*, σε αντίθεση με το ίδιο το κορίτσι που

¹¹⁷ Μια τέτοιου τύπου χρήση των σημάτων ενδέχεται να ισχύει και στους πρωτόγονους λαούς (πρβλ Heidegger 1962, σελ. 113/82: "... για τον πρωτόγονο άνθρωπο το σημείο συμπίπτει με το ενδεικνυόμενο. Το σημείο μπορεί να αντιπροσωπεύσει όχι μόνο με την έννοια του υποκατάστατου, παρά κατά τρόπο τρόπο ώστε αυτό τούτο το σημείο πάντα είναι το ενδεικνυόμενο.")

¹¹⁸ Το γεγονός ότι ένα σήμα, όπως τα ήχη στο έδαφος, ενδέχεται να είναι πολύ μακριά από το αντικείμενο με το οποίο συνδέεται δεν αλλάζει την κατάσταση. Η περίπτωση μπορεί να θεωρηθεί ακριβώς ανάλογη με την περίπτωση εντοπισμού της οσμής ενός απομακρυσμένου αρπακτικού ή ακόμα και των ήχων που παράγει ένα απομακρυσμένο αρπακτικό.

¹¹⁹ Είναι σκόπιμο να τονίσουμε ότι η διαφορά που εντοπίζουμε μεταξύ σημάτων και αναπαραστάσεων με την πλήρη έννοια του όρου δεν προτείνεται ως μια διαφορά που ισχύει *όντως* μεταξύ ζώων και ανθρώπων. Στόχος μας είναι απλώς να προσδιορίσουμε *εννοιολογικά* το τι σημαίνει για έναν οργανισμό να χρησιμοποιεί αναπαραστάσεις με την πλήρη έννοια του όρου και γι' αυτό το λόγο αντιδιαστέλλουμε αυτή τη χρήση με τη χρήση σημάτων. Ενδεχομένως, στο ζωικό βασίλειο υπάρχουν ζώα που χρησιμοποιούν αναπαραστάσεις με την πλήρη έννοια του όρου, αλλά και αντίστροφα, ενδεχομένως οι άνθρωποι χρησιμοποιούν ορισμένες φορές σήματα με την παραπάνω έννοια του όρου. Αυτό είναι ένα εμπειρικό ζήτημα που μπορεί να απαντηθεί μόνο με εμπειρική έρευνα.

στέκεται δίπλα στην προσωπογραφία του και το οποίο εμφανίζεται ως σωματικά παρόν. Στο συμπεριφοριστικό επίπεδο αυτή η διαφορά εκδηλώνεται από τις διαφορετικές συμπεριφορές που μπορεί να προκαλέσει η προσωπογραφία του κοριτσιού και το ίδιο το κορίτσι. Χαρακτηριστικό προς αυτή την κατεύθυνση είναι ένα γνωστό συμβάν από τις πρώτες μέρες του κινηματογράφου. Όταν προβλήθηκε η ταινία των αδελφών Lumière *Η είσοδος του τραίνου στο σταθμό της La Ciotat* οι θεατές, βλέποντας το τρένο να έρχεται καταπάνω τους, πανικοβλήθηκαν και άρχισαν να τρέχουν για να μην τους πατήσει. Στην προκειμένη περίπτωση, οι θεατές αντέδρασαν όπως θα αντιδρούσαν αν έβλεπαν το ίδιο το τρένο. Με άλλα λόγια, δεν κατάλαβαν ότι επρόκειτο για μια αναπαράσταση του τραίνου, και ως εκ τούτου, για ένα μη σωματικά παρόν τρένο.

Για να διακρίνουμε τον τρόπο με τον οποίο λειτουργούν οι αναπαραστάσεις με την πλήρη έννοια του όρου από τον τρόπο που λειτουργούν τα σήματα θα λέμε ότι οι αναπαραστάσεις με την πλήρη έννοια του όρου *παραπέμπουν* (denote¹²⁰, designate) σε κάτι, ενώ τα σήματα *ενδεικνύουν* (indicate) κάτι. Με άλλα λόγια, μια αναπαράσταση παραπέμπει σε κάτι όταν αναφέρεται σε αυτό ως μη σωματικά παρόν¹²¹, ενώ ένα σήμα ενδεικνύει κάτι όταν αναφέρεται σε κάτι που εκλαμβάνεται ως σωματικά παρόν - όπου αυτό σημαίνει ότι το σήμα εκλαμβάνεται ως κάποιο φυσικό χαρακτηριστικό του ίδιου του αναφερόμενου.

Με την ικανότητα χρήσης αναπαραστάσεων με την πλήρη έννοια του όρου δημιουργείται για πρώτη φορά η δυνατότητα αποσύνδεσης του χρήστη από το άμεσο περιβάλλον του. Δημιουργείται, έτσι, ένα δεύτερο

¹²⁰ Το "denote" έχει αποδοθεί με τον όρο "δηλώνω" (βλ., ενδεικτικά, Ricoeur 1996 (μετάφραση Κ. Παπαγιώργη) και Συμεωνίδου-Χριστίδου 1998). Ωστόσο, δεν υιοθετώ αυτή τη μετάφραση γιατί το "δηλώνω" αποτελεί ορθώς μετάφραση και του "state" που εκφράζει ένα είδος επιτελεστικού ομιλιακού ενεργήματος (βλέπε την τρίτη επισήμανση στο τέλος της τρέχουσας ενότητας). Επίσης, ως μετάφραση του "denote" έχει χρησιμοποιηθεί ο όρος "υποδηλώνω" (βλ. Wittgenstein 1953/1977 (μετάφραση Π. Χριστοδουλίδη)). Δεν υιοθετώ ούτε αυτή τη μετάφραση γιατί το "υποδηλώνω" σημαίνει συνήθως "εκφράζω κάτι υπαινικτικά, με έμμεσο τρόπο". Όταν, όμως, λέμε ότι μια αναπαράσταση denotes κάτι άλλο σίγουρα δεν υπάρχει τίποτε το υπαινικτικό σε αυτή τη σχέση.

¹²¹ Το αντικείμενο στο οποίο μια αναπαράσταση παραπέμπει ονομάζεται συχνά denotatum ή designatum. Οι δύο αυτοί όροι δεν είναι συνώνυμοι, ωστόσο, για τους σκοπούς του κεφαλαίου, η διαφορά τους δεν θα μας απασχολήσει. Ο Morris (βλ. υποσημείωση 113) τούς διαφοροποιεί ως εξής: "Ένα designatum δεν είναι ένα πράγμα, αλλά ένα είδος αντικειμένων ή μια κλάση αντικειμένων - και μια κλάση μπορεί να έχει πολλά μέλη, ή ένα μέλος ή κανένα μέλος. Τα denotata είναι τα μέλη των κλάσεων" (Morris 1938, σελ. 83). "Έτσι, γίνεται σαφές ότι, ενώ κάθε σημείο έχει ένα designatum, δεν έχει κάθε σημείο ένα denotatum" (ibid.). Θα αναφερθούμε εμμέσως στη διάκριση μεταξύ denotatum και designatum λίγο παρακάτω κατά την παρουσίαση της θεωρίας του Frege.

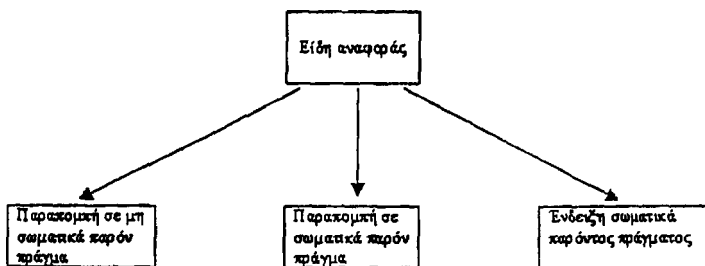
επίπεδο, το επίπεδο παραπομπής στο περιβάλλον. Αυτό, σε πρώτη προσέγγιση, προϋποθέτει μια απόσπαση-απεικλωτική του αναπαραστασιακού φορέα από ό,τι αναριστά - πράγμα που δεν συνεπάγεται, φυσικά, ότι η αναπαράσταση παύει να αναφέρεται σε κάτι συγκεκριμένο. Ο αναπαραστασιακός φορέας εξακολουθεί να συμπλέκεται με το αναφερόμενο αλλά με έναν πιο πλούσιο τρόπο, απελευθερώνοντας έτσι τον χρήστη από το περιορισμένο σύνολο συμπεριφορών που μπορούν να προκαλέσουν τα σήματα. Αυτό συμβαίνει γιατί ο αναπαραστασιακός φορέας δεν εκλαμβάνεται ως ένα φυσικό χαρακτηριστικό του αναφερόμενου, ως κάτι, δηλαδή, που ανήκει εγγενώς στο αναφερόμενο. Έτσι, ο αναπαραστασιακός φορέας αποκτά μια σχετική ανεξαρτησία από το αναφερόμενο και το άμεσο περιβάλλον. Η αποστασιοποίηση από το άμεσο περιβάλλον που επιτυγχάνεται μέσω των αναπαραστάσεων επιτρέπει στον χρήστη να κατανοήσει κάτι, μεταξύ άλλων, ως ενεργεία υπαρκτό (actual), παρελθοντικό, μελλοντικό ή και φανταστικό¹²². Γενικότερα, όταν ένα πράγμα παύει να λειτουργεί ως σήμα και λειτουργεί ως αναπαράσταση με την πλήρη έννοια του όρου, το αναφερόμενο μπορούμε να πούμε ότι *εμφανίζεται* με κάποιο τρόπο στον χρήστη του, δηλαδή *κατανοείται* με κάποιο τρόπο από αυτόν.

Δεν θα επεκταθούμε περισσότερο σε αυτού του τύπου την ανάλυση γιατί υπερβαίνει τους στόχους του κεφαλαίου. Αυτό που θέλουμε απλώς να κρατήσουμε είναι ότι η λειτουργία των αναπαραστάσεων, σε αντίθεση με εκείνη των σημάτων, δεν εξαντλείται στην αναφορά ενός σωματικά παρόντος πράγματος, δηλαδή σε αυτό που θα ονομάζουμε ένδειξη (indication). Αν επιλέξουμε να θεωρούμε και τα σήματα ως ένα είδος αναπαραστάσεων, τότε αυτά είναι, σίγουρα, μια πολύ απλούστερη μορφή αναπαράστασης. Το περιεχόμενο των αναπαραστάσεων με την πλήρη έννοια του όρου περιλαμβάνει μια επιπλέον *διάσταση* - ένα πρόσθετο βαθμό ελευθερίας - που αφορά τον τρόπο εμφάνισης, κατανόησης και, γενικότερα χρήσης, του αναφερόμενου.

Ωστόσο, θα μπορούσε κανείς να αντιτείνει ότι αν κάποιος μπορεί να κατανοήσει αναπαραστάσεις με την πλήρη έννοια του όρου, δηλαδή αναπαραστάσεις που αναφέρονται σε ένα μη σωματικά παρόν πράγμα, τότε μπορεί να κατανοήσει και αναπαραστάσεις που αναφέρονται σε ένα σωματικά παρόν πράγμα, όπως, για παράδειγμα, η έκφραση "Να η Άννα" η οποία αναφέρεται στην Άννα ως σωματικά παρούσα. Σημαίνει, άραγε, αυτό ότι η εν λόγω έκφραση λειτουργεί ως σήμα; Θεωρούμε πως όχι, γιατί η έκφραση "Να η Άννα", όταν χρησιμοποιείται από ένα γνώστη της

¹²² Η ικανότητα απόσπασης από το άμεσο περιβάλλον προϋποτίθεται επίσης της ικανότητας κατανόησης *κανονιστικών* σχέσεων, όπως το ορθό και το λανθασμένο. Για παράδειγμα, μια λανθασμένη πεποίθηση είναι μια πελοίτηση που δεν ανταποκρίνεται στα πράγματα και, υπό αυτή την έννοια, προϋποθέτει την ικανότητα κατανόησης κάποιας κατάστασης που δεν υφίσταται.

ελληνικής γλώσσας, δεν ενδεικνύει απλώς ένα σωματικά παρόν πράγμα, αλλά αναφέρεται σε αυτό το πράγμα ως σωματικά παρόν. Με άλλα λόγια, στη συγκεκριμένη αναπαράσταση υπάρχει ήδη η επιπλέον διάσταση που αφορά τον τρόπο εμφάνισης του αναφερόμενου: Ένας επαρκής ομιλητής της ελληνικής γλώσσας όταν ακούει την έκφραση "Να η Άννα" δεν την εκλαμβάνει ως ένα (ιδιότυπο) φυσικό χαρακτηριστικό της ίδιας της Άννας, όπως, δηλαδή, ένα ζώο, ενδεχομένως, εκλαμβάνει το σήμα κινδύνου για την παρουσία λιονταριού ως ένα φυσικό χαρακτηριστικό του ίδιου του λιονταριού. Το γεγονός ότι η φυσική γλώσσα διαθέτει τρόπους για να αναφέρεται σε σωματικά παρόντα πράγματα δεν σημαίνει ότι αυτοί οι τρόποι πρέπει να εξισωθούν ως προς τη λειτουργία τους με τη χρήση σημάτων. Οι δεκτικές αντωνυμίες της φυσικής γλώσσας που αναφέρονται σε σωματικά παρόντα πράγματα προϋποθέτουν την ικανότητα παραπομπής στα πράγματα -αποτελούν ουσιαστικά μια τροποποιημένη χρήση της παραπομπής: αντί να παραπέμπουν σε ένα μη σωματικά παρόν πράγμα παραπέμπουν σε κάποιο σωματικά παρόν. Για αυτό το λόγο η παραπομπή σε ένα σωματικά παρόν πράγμα διαφέρει ριζικά από την ένδειξη ενός σωματικά παρόντος πράγματος: η πρώτη σε αντίθεση με τη δεύτερη δεν εκλαμβάνεται ως ένα φυσικό χαρακτηριστικό του αναφερόμενου (βλ. παρακάτω σχήμα).



Όμως, αν τελικά ορίσουμε την έννοια της ένδειξης ενός πράγματος με ένα τόσο ασθενή τρόπο, θα μπορούσε κανείς να αντιτείνει ότι αυτή καταλήγει να μην διακρίνεται από την άμεση αντίληψη του πράγματος. Αν το σήμα κινδύνου για την παρουσία λιονταριού εκλαμβάνεται από τον χρήστη ως κάποιο φυσικό χαρακτηριστικό του ίδιου του λιονταριού, τότε η αντίληψη του σήματος είναι της ίδιας τάξης με την αντίληψη του λιονταριού (ή κάποιου φυσικού του χαρακτηριστικού). Η ένσταση αυτή είναι σωστή. Από την οπτική του ίδιου του χρήστη (που δεν έχει την ικανότητα χρήσης αναπαραστάσεων με την πλήρη έννοια του όρου) η ένδειξη ενός πράγματος δεν διαφέρει από την αντίληψη του ίδιου του πράγματος ή κάποιου εγγενούς χαρακτηριστικού αυτού. Η διαφορά

μεταξύ ένδειξης και άμεσης αντίληψης ενός πράγματος μπορεί να εντοπισθεί μόνο από την οπτική ενός παρατηρητή που έχει τη δυνατότητα χρήσης αναπαραστάσεων με την πλήρη έννοια του όρου¹²³.

Για να μιλήσουμε με έναν πιο οικείο και φιλοσοφικά αναγνωρίσιμο τρόπο για την επιπλέον διάσταση που καθιστά ένα σήμα αναπαράσταση με την πλήρη έννοια του όρου χρειάζεται να εισαγάγουμε την έννοια του νοήματος, όπως αυτή χρησιμοποιήθηκε από τον Frege. Ο Frege (1892/1960) εισήγαγε την έννοια του νοήματος (Sinn) κατ' αντιδιαστολή προς την έννοια της αναφοράς (Bedeutung)¹²⁴. Σύμφωνα με τη διάκριση του Frege, η έννοια του περιεχομένου μιας αναπαράστασης μπορεί να χρησιμοποιηθεί με δύο διαφορετικούς τρόπους: μπορεί να σημαίνει είτε το αναφερόμενο της αναπαράστασης, (εναλλακτικά, το denotatum ή designatum¹²⁵) είτε το νόημα της αναπαράστασης. Για να χρησιμοποιήσουμε ένα πασίγνωστο παράδειγμα, οι εκφράσεις "ο Αυγερινός" και "ο Αποσπερίτης" έχουν το ίδιο αναφερόμενο: τον πλανήτη Αφροδίτη. Ωστόσο, οι δύο αυτές αναπαραστάσεις δεν ταυτίζονται πλήρως ως προς το περιεχόμενο, διότι η έκφραση "ο Αυγερινός" έχει διαφορετικό νόημα από την έκφραση "ο Αποσπερίτης". Η πρώτη αναπαράσταση περιγράφει την Αφροδίτη όπως αυτή παρουσιάζεται την αυγή, ενώ η δεύτερη όπως παρουσιάζεται κατά τη δύση. Το νόημα των εκφράσεων "ο

¹²³ Επίσης, δεν είναι, ίσως, ορθό το να θεωρούμε ότι όντα που δεν έχουν την ικανότητα χρήσης αναπαραστάσεων με την πλήρη έννοια του όρου, μπορούν παρόλα αυτά να αντιλαμβάνονται έναν αντικειμενικό κόσμο, δηλαδή ένα κόσμο αντικειμένων και ιδιοτήτων. Εφόσον, τα όντα αυτά δεν μπορούν να αποσπαστούν από την άμεση εμπειρία, δεν έχουν τη δυνατότητα αφαίρεσης και γενίκευσης: το λιοντάρι-εδώ-και-τώρα δεν μπορεί να αναγνωριστεί ως του ίδιου τύπου με το λιοντάρι που εμφανίστηκε λίγες μέρες πριν, δηλαδή με κάτι που δεν είναι σωματικά παρόν.

¹²⁴ Ο Frege χρησιμοποίησε την εν λόγω διάκριση για τις γλωσσικές αναπαραστάσεις. Ωστόσο, σε μια πρώτη προσέγγιση, μπορούμε να την εφαρμόσουμε σε όλα τα είδη εξωτερικών αναπαραστάσεων. Μια αντίστοιχη διάκριση εισήγαγε ο Goodman όσον αφορά τις εικονιστικές αναπαραστάσεις (ή, απλώς, αναπαραστάσεις (representations), όπως αυτές τις ονόμαζε) αναφερόμενος στην αναπαράσταση και της αναπαράστασης-ως: "Η καθημερινή χρήση συχνά δεν διακρίνει μεταξύ της αναπαράστασης και της αναπαράστασης-ως ... Εάν σου πω ότι έχω ένα συγκεκριμένο μαύρο άλογο και κατόπιν βγάλω μια φωτογραφία στην οποία το άλογο εμφανίζεται ως μια μικρή φωτεινή κηλίδα, πολύ δύσκολα μπορείς να με κατηγορήσεις ότι ψεύδομαι αλλά μπορεί κάλλιστα να αισθάνεσαι ότι σε παραπλάνησα. Εσύ εντελώς κατανοητά ... ανέμενες ότι η εικόνα ... ήταν μια εικόνα-μαύρου-αλόγου, ότι αναπαριστά το ... εν λόγω πράγμα ως ένα [μαύρο] άλογο" (Goodman 1968, σελ. 29). Έτσι, λοιπόν, για τον Goodman μια εικόνα έχει μια διπλή λειτουργία: παραπέμπει (denotes) σε κάποιο αντικείμενο και το ταξινομεί: μια εικονιστική αναπαράσταση πάντα παραπέμπει σε κάτι ως κάτι (δηλαδή κατά ένα συγκεκριμένο τρόπο).

¹²⁵ Βλ. υποσημείωση 121.

Αυγερινός" και "ο Αποσπερίτης" είναι η περιγραφή - ή, με γενικότερους όρους, η προοπτική - υπό την οποία προσεγγίζεται ό,τι αναπαριστούν. Επίσης, μια άλλη συναφής μεταφορά με την οποία ο Frege προσπάθησε να χαρακτηρίσει το νόημα -τουλάχιστον στην περίπτωση των ονομάτων- είναι εκείνη του τρόπου παρουσίασης: νόημα μιας έκφρασης είναι ο τρόπος παρουσίασης αυτού στο οποίο παραπέμπει.

Ένας βασικός λόγος για τον οποίον ο Frege εισήγαγε τη διάκριση νοήματος/αναφοράς ήταν η διαφορά πληροφοριακότητας μεταξύ προτάσεων με την ίδια σημασιολογική τιμή. Αν τα *a* και *b* έχουν την ίδια αναφορά, τότε οι γλωσσικές προτάσεις "το *a* είναι το *b*" και "το *a* είναι το *a*" έχουν την ίδια σημασιολογική τιμή. Ωστόσο, είναι προφανές ότι η πρώτη πρόταση σε αντίθεση με τη δεύτερη είναι πληροφοριακή, δηλαδή είναι κάτι παραπάνω από μια απλή ταυτολογία. Αυτός είναι ο βασικός λόγος που ο Frege ανέπτυξε τη θεωρία νοήματος ως συμπλήρωμα της θεωρίας αναφοράς. Ωστόσο, εκεί που η αναγκαιότητα της διάκρισης σε νόημα και αναφορά αναδεικνύεται ξεκάθαρα είναι στην περίπτωση αναπαράστασης φανταστικών αντικειμένων: η έκφραση "ο Κένταυρος" δεν έχει αναφερόμενο γιατί δεν υπάρχει τέτοιο ζώο στη φύση, ωστόσο αυτό δεν σημαίνει ότι η έκφραση δεν έχει και περιεχόμενο. Στην προκειμένη περίπτωση το περιεχόμενο της έκφρασης "ο Κένταυρος" είναι αποκλειστικά το νόημά της¹²⁶. Σε συμφωνία προς το παραπάνω παράδειγμα, με την έννοια του περιεχομένου μιας αναπαράστασης θα αναφερόμαστε συνολικά στο νόημα και το αναφερόμενο της αναπαράστασης¹²⁷. Γενικά, στο Frege, ενώ μια έκφραση μπορεί να έχει

¹²⁶ Ωστόσο, η μεταφορά που χρησιμοποιεί ο Frege για να μιλήσει για το νόημα (δηλαδή, νόημα=τρόπος παρουσίασης του αναφερόμενου) δεν είναι συμβατή με την περίπτωση των μη-υπαρχόντων αντικειμένων: αν δεν υπάρχει αναφερόμενο τότε δεν υπάρχει και τρόπος παρουσίασής του. Αν τελικά το νόημα κατανοηθεί ως τρόπος παρουσίασης του αναφερόμενου, τότε αυτό είναι ουσιαδώς εξαρτώμενο από την ύπαρξη του αναφερόμενου.

¹²⁷ Ένας τρίτος λόγος για τον οποίον ο Frege εισήγαγε την έννοια του νοήματος είναι για να επιλύσει το πρόβλημα της αντικατάστασης όρων μέσα σε αδιαφανή συμφρασιακά πλαίσια, όπως εκείνα που δημιουργούν τα ρήματα προτασιακής στάσης, οι τροπικοί τελεστές και τα εισαγωγικά. Έστω, για παράδειγμα, η πρόταση "Ο Γιάννης πιστεύει ότι ο Αυγερινός είναι ο Αυγερινός" (1). Αν αντικαταστήσουμε την δεύτερη εμφάνιση της έκφρασης "ο Αυγερινός" με την έκφραση "ο Αποσπερίτης", εφόσον και οι δύο έχουν το ίδιο αναφερόμενο, θα πρέπει η πρώτη πρόταση να έχει την ίδια τιμή αληθείας με τη δεύτερη. Ωστόσο, αυτό δεν ισχύει διότι αν ο Γιάννης δεν γνωρίζει το αναφερόμενο της έκφρασης "ο Αποσπερίτης", η πρόταση "Ο Γιάννης πιστεύει ότι ο Αυγερινός είναι ο Αποσπερίτης" (2) είναι ψευδής. Η λύση που έδωσε ο Frege σε αυτό το πρόβλημα με την εισαγωγή της έννοιας του νοήματος είναι η εξής: στο συμφρασιακό πλαίσιο μιας πεποιθήσης τα κύρια ονόματα δεν παραπέμπουν στο σύνθηδες αναφερόμενό τους αλλά σε ένα έμμεσο (ή πλάγιο) αναφερόμενο που δεν είναι άλλο από το νόημά τους. Έτσι, σύμφωνα με αυτή τη θεώρηση, η πρόταση (1) δεν έχει την ίδια

νόημα χωρίς αναφορά, το αντίθετο δεν είναι δυνατό. Αυτό συμβαίνει γιατί το νόημα καθορίζει την αναφορά.

Ενσωματώνοντας τη διάκριση νοήματος/αναφοράς (ή, ακριβέστερα, αναφερόμενου) στον προκαταρκτικό ορισμό της έννοιας της αναπαράστασης μπορούμε να πούμε ότι:

Ορισμός 4:	Ένα αντικείμενο Α συνιστά αναπαράσταση κάποιου πράγματος Ο, δηλαδή έχει ως περιεχόμενο το Ο, αν για κάποιον επαρκή χρήστη Χ η μορφή Μ του Α μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο Ο (αναφερόμενο) υπό μια συγκεκριμένη προοπτική Ν (νόημα).
------------	---

Προτού προχωρήσουμε θα κάνουμε τρεις επισημάνσεις που θεωρούμε σημαντικές.

1. Δεν αναφέρονται όλες οι γλωσσικές αναπαράστασεις σε αντικείμενα. Ο Frege διακρίνει, γενικά, τρεις γραμματικές κατηγορίες: τους ενικούς όρους, τις "εννοιακές λέξεις" (ό,τι με σημερινούς όρους θα ονομάζαμε κατηγορήματα) και τις γλωσσικές προτάσεις. Οι γλωσσικές προτάσεις μπορούν να αναλυθούν σε ενικούς όρους και κατηγορήματα, ενώ οι ενικοί όροι και τα κατηγορήματα δεν είναι περαιτέρω αναλύσιμα. Από τα τρία παραπάνω είδη γλωσσικών αναπαράστασεων μόνο οι ενικοί όροι έχουν ως αναφερόμενο μεμονωμένα αντικείμενα (άτομα). Ποιο είναι, λοιπόν, το αναφερόμενο των κατηγορημάτων και των γλωσσικών προτάσεων; Ας πάρουμε πρώτα τα κατηγορήματα: σύμφωνα με τον Frege, το αναφερόμενο του κατηγορήματος "είναι κόκκινος" είναι το σύνολο των ατόμων που έχουν την ιδιότητα του κόκκινου¹²⁸. Το αναφερόμενο, λοιπόν, σε αυτή την περίπτωση δεν είναι κάποιο μεμονωμένο άτομο αλλά ένα σύνολο ατόμων (ακόμα και αν πρόκειται για μονοσύνολο¹²⁹). Γενικότερα, αν το κατηγορημα αναφέρεται σε κάποια ν-μελή σχέση Κ, τότε το αναφερόμενο του κατηγορήματος είναι το σύνολο των διατεταγμένων ν-άδων που πληρούν τη σχέση Κ. Αντίστοιχα, στις γλωσσικές προτάσεις το αναφερόμενο είναι η τιμή αληθείας τους, το εάν δηλαδή είναι αληθείς ή ψευδείς. Με άλλα λόγια, σε επίπεδο γλωσσικής πρότασης υπάρχουν μόνο

τιμή αληθείας με την πρόταση (2), διότι η έκφραση "ο Αυγερινός" έχει διαφορετικό έμμεσο αναφερόμενο από τη έκφραση "ο Αποσπερίτης" και, κατά συνέπεια, η αντικατάσταση της μιας από την άλλη δεν διατηρεί την αλήθεια (Frege 1892/1977, σελ. 28. Βλέπε, επίσης, Miller 1998, σελ. 31-32).

¹²⁸ Με τον ίδιο τρόπο μπορούμε να προσδιορίσουμε το αναφερόμενο και των γενικών όρων, όπως, για παράδειγμα, του όρου "λιοντάρι".

¹²⁹ Σύμφωνα με την διάκριση μεταξύ designatum και denotatum που εισαγάγαμε στην υποσημείωση 121, τα κατηγορήματα έχουν μόνο designatum και όχι denotatum.

δύο δυνατές αναφορές. Εδώ πλέον καθίσταται εντελώς προφανής η ανεπάρκεια της θεωρίας αναφοράς για την περιγραφή αυτού που προθεωρητικά ονομάζουμε περιεχόμενο μιας γλωσσικής πρότασης: το σύνολο των γλωσσικών προτάσεων συνδέεται με δύο μόνο είδη αναφοράς! Έτσι, λοιπόν, γίνεται φανερό για μια ακόμα φορά η αναγκαιότητα ανάπτυξης μιας θεωρίας νοήματος.

	Αναφερόμενο	Νόημα
Ενικός όρος	Άτομα ¹³⁰	Τρόπος παρουσίασης ¹³¹
Κατηγορήμα - εννοιακή λέξη (Begriffswort)	Σύνολο ατόμων ¹³² ή, γενικότερα, διατεταγμένων ν-άδων	Έννοια (Begriff / concept)
Γλωσσική πρόταση	Αληθοτιμή	Διανόημα ¹³³ (Gedanke /thought)

Ο ανωτέρω πίνακας συγκεφαλαιώνει τα διαφορετικά είδη αναφοράς αλλά και τα αντίστοιχα είδη νοήματος για κάθε μία από τις τρεις διαφορετικές περιπτώσεις γλωσσικών αναπαραστάσεων. Φυσικά δεν θα πρέπει να θεωρηθεί ότι τα συγκεκριμένα τρία είδη γλωσσικών αναπαραστάσεων εξαντλούν όλους τους τύπους αναπαράστασης που απαιτούνται στη φυσική γλώσσα. Ένα προφανές παράδειγμα είναι οι λογικοί σύνδεσμοι. Το "και" δεν είναι ούτε ενικός όρος, ούτε κατηγορήμα, ούτε γλωσσική πρόταση. Επίσης δεν έχει αναφορά¹³⁴ και το νόημά του καθορίζεται από τον τρόπο που χρησιμοποιείται. Το τι είδη αναφοράς και

¹³⁰ Με την ορολογία του Morris πρόκειται για τα denotata (βλ. υποσημείωση 121)

¹³¹ Ο Evans (1982) και ο McDowell (1984, 1986), ερμηνεύοντας τον Frege, θεώρησαν ότι το νόημα των ενικών όρων, ειδικότερα των κυρίων ονομάτων και των δεικτικών και προσωπικών αντωνυμιών, δεν είναι περιγραφικό. Γι' αυτό μίλησαν για μη-περιγραφικούς τρόπους παρουσίασης ή de re νοήματα (βλ. Recanati 1993, σελ. 98). Επειδή συμφωνούμε με αυτή τη προσέγγιση στον ορισμό 4 χρησιμοποίησαμε το γενικότερο όρο "προοπτική" αντί του όρου "περιγραφή" για να αναφερθούμε στο νόημα Ν της αναπαράστασης.

Επίσης, στο υπόλοιπο της διατριβής, όταν δεν αναφερόμαστε ρητά στη θεωρία του Frege, θα χρησιμοποιούμε τους όρους "τρόπος εμφάνισης" και "τρόπος φανέρωσης" αντί του όρου "τρόπος παρουσίασης". Ο όρος "παρουσίαση" θα χρησιμοποιείται για εκείνες τις περιπτώσεις που κάτι εμφανίζεται / φανερώνεται ως σωματικά παρόν.

¹³² Με την ορολογία του Morris πρόκειται για το designatum (βλ. υποσημείωση 121)

¹³³ Το διανόημα είναι ό,τι μετέπειτα στην γλωσσουαλυτική παράδοση ονομάστηκε πρόταση (proposition).

¹³⁴ Πρβλ. Wittgenstein 1921/1971: "Η βασική μου σκέψη είναι πώς οι 'λογικές σταθερές' δεν υποκαθιστούν (represent, vertreten)" (§ 4.0312).

νοήματος τελικά υπάρχουν είναι ένα ανοικτό ερευνητικό ζήτημα (βλ. Evans 1982).

Αν τώρα κανείς εξετάσει τις διάφορες μη-γλωσσικές αναπαραστάσεις παραμένει ένα ανοιχτό ερευνητικό ζήτημα το εάν αυτές είναι σε θέση να εκφράσουν διαφορετικά είδη αναφοράς (και, αντίστοιχα, νοήματος). Αν συνέβαινε κάτι τέτοιο, αυτό θα σήμαινε ότι και στα μη-γλωσσικά συστήματα αναπαραστάσεων υφίστανται διαφορετικές γραμματικές κατηγορίες. Θα επανέλθουμε σε αυτό το ζήτημα σε επόμενα κεφάλαια. Το μόνο που θέλουμε να κρατήσουμε προς στιγμήν από την πρώτη επισήμανση είναι ότι όταν μιλάμε για αναφερόμενο μιας αναπαράστασης δεν σημαίνει κατ' ανάγκη ότι αναφερόμαστε σε κάποιο αντικείμενο. Χωρίς να υιοθετήσουμε τις επιμέρους λύσεις του Frege, μπορούμε να δεχτούμε ότι μια αναπαράσταση ενδέχεται να αναφέρεται, όχι μόνο σε αντικείμενα, αλλά και σε ιδιότητες, σχέσεις ή καταστάσεις πραγμάτων. Έτσι, ο προκαταρκτικός ορισμός της έννοιας της αναπαράστασης μπορεί να τροποποιηθεί ως εξής:

Ορισμός 5:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου αντικειμένου, ιδιότητας, σχέσης ή κατάστασης πραγμάτων O , δηλαδή έχει ως περιεχόμενο το O , αν για κάποιον επαρκή χρήστη X η μορφή M του A μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο O (αναφερόμενο) υπό μια συγκεκριμένη προοπτική N (νόημα).
------------	--

2. Μια δεύτερη επισήμανση που θεωρούμε πολύ σημαντική σε σχέση με τη διάκριση νοήματος/αναφοράς είναι ότι ο Frege, σύμφωνα τουλάχιστον με την ερμηνεία του Evans, ανέπτυξε τη θεωρία αναφοράς για να περιγράψει τη σχέση των γλωσσικών εκφράσεων με τον κόσμο και τη θεωρία νοήματος για να περιγράψει τη σχέση των γλωσσικών εκφράσεων με το νου, δηλαδή ως μια θεωρία της *κατανόησης* των γλωσσικών αναπαραστάσεων: "Μπορούμε να θεωρήσουμε ότι η θεωρία αναφοράς, πολύ γενικά, αφορά τις σχέσεις μεταξύ εκφράσεων της γλώσσας και οντοτήτων του κόσμου. Τώρα, ο Frege συνειδητοποίησε την ανάγκη για ένα άλλο επίπεδο περιγραφής και θεωρίας που αφορούσε, θα μπορούσαμε εξίσου γενικά να πούμε, τις σχέσεις μεταξύ εκφράσεων της γλώσσας και της κατανόησης αυτών από επαρκείς ομιλητές" (Evans 1982, σελ. 13)¹³⁵. Σύμφωνα λοιπόν με αυτή την ερμηνεία, η θεωρία αναφοράς

¹³⁵ Επίσης, ο McCulloch, αναφερόμενος στο ίδιο ζήτημα, επισημαίνει: "Η θεωρία νοήματος του Frege είναι πολύ λιγότερο επεξεργασμένη απ' ό,τι η θεωρία αναφοράς. Αλλά, ενώ η τελευταία αφορά τους αναπαραστατικούς δεσμούς μεταξύ γλώσσας και κόσμου, η πρώτη αφορά τους δεσμούς μεταξύ γλώσσας και νου." (McCulloch 1995, σελ. 63).

έχει να κάνει με τη σχέση αναπαράστασης - κόσμος και η θεωρία νοήματος με τη σχέση αναπαράστασης - υποκειμένου που την κατανοεί. Οι γλωσσικές εκφράσεις, θα μπορούσαμε να πούμε, έχουν δύο λειτουργίες: αφενός, *παραπέμπουν* σε κάποιο αναφερόμενο και, αφετέρου, *εκφράζουν* (express) τις σκέψεις (ήτοι, τα διανοήματα) του υποκειμένου που τις παράγει.

Ωστόσο, αν κανείς επιχειρήσει να προσδιορίσει το οντολογικό καθεστώς του νοήματος έρχεται αντιμέτωπος με ανυπέρβλητα προβλήματα. Σύμφωνα με τον Frege, τα νοήματα ανήκουν σε έναν πλατωνικού τύπου κόσμο που διαχωρίζεται τόσο από τη σφαίρα του ψυχολογικού είναι όσο και από τη σφαίρα των υλικών πραγμάτων: "Ένα διανόημα δεν ανήκει ούτε στον εσωτερικό μου κόσμο ως μια [υποκειμενική] παράσταση (Vorstellung), ούτε όμως στον εξωτερικό κόσμο, τον κόσμο των πραγμάτων που γίνονται αντιληπτά από τις αισθήσεις" (Frege 1988, σελ. 52)¹³⁶. Τα νοήματα είναι αφηρημένες οντότητες που υπάρχουν αιώνια ανεξάρτητα από το εάν συλλαμβάνονται ή όχι: "Το έργο της επιστήμης δεν συνίσταται στη δημιουργία, αλλά στην ανακάλυψη αληθών διανοημάτων ... η αλήθεια ενός διανοήματος είναι άχρονη" (Frege 1988, σελ. 51). Έτσι, τα νοήματα δεν έχουν ανάγκη το κατανοούν υποκειμενο για να υπάρχουν -είναι αντικειμενικά και ανεξάρτητα του νου: "Όταν κανείς συλλαμβάνει ... ένα διανόημα δεν το δημιουργεί αλλά έρχεται να σταθεί σε μια συγκεκριμένη σχέση ... προς κάτι που ήδη προϋπήρχε" (Frege 1988, σελ. 55, σημ. 7)¹³⁷.

Η θέση του Frege για το οντολογικό καθεστώς των νοημάτων δημιουργεί ανυπέρβλητα προβλήματα, ειδικότερα στις μέρες μας που οι περισσότερες φιλοσοφικές προσεγγίσεις θέλουν να είναι φυσιοκρατικές. Κυριότερο απ' όλα είναι το πρόβλημα της αλληλεπίδρασης: δεν φαίνεται να υπάρχει καμιά ικανοποιητική εξήγηση για το πώς κάτι άυλο, όπως το νόημα που ανήκει σε έναν τρίτο κόσμο, μπορεί να αλληλεπιδρά αιτιακά με τα πράγματα του υλικού κόσμου και εν προκειμένω τη ψυχολογική κατάσταση του στοχαστή. Ο Frege διατείνεται ότι "όταν ένα διανόημα συλλαμβάνεται, αρχικά επιφέρει αλλαγές μόνο στον εσωτερικό κόσμο εκείνου που το συλλαμβάνει· ωστόσο παραμένει ανέγγιχτο ως προς τον

¹³⁶ Βλ. επίσης Frege 1892/1977, σελ. 21-22: "Πρέπει να ξεχωρίζουμε την αναφορά [ήτοι, το αναφερόμενο] και το νόημα ενός σημείου από την [υποκειμενική] παράσταση (Vorstellung) που είναι συνδεδεμένη με το σημείο.... Η αναφορά [ήτοι, το αναφερόμενο] ενός κύριου ονόματος είναι το ίδιο το αντικείμενο στο οποίο παραπέμπουμε [designate] μέσω του ονόματος· η παράσταση που έχουμε σ' αυτή την περίπτωση είναι εντελώς υποκειμενική· ανάμεσά τους βρίσκεται το νόημα που δεν είναι πα υποκειμενικό όπως η παράσταση, ούτε είναι όμως και το ίδιο το αντικείμενο".

¹³⁷ Βλ. επίσης Frege 1988, σελ. 55: "...ο στοχαστής δεν δημιουργεί [τα διανοήματα]...αλλά οφείλει να τα εκλάβει όπως είναι. Μπορούν να είναι αληθή χωρίς να συλλαμβάνονται από ένα στοχαστή".

πυρήνα της ουσίας του" (Frege 1988, σελ. 54). Η έννοια της σύλληψης του νοήματος μιας αναπαράστασης αποτελεί μια μυστηριώδη ενέργεια που κάθε άλλο παρά εξηγεί τον τρόπο κατανόησης της αναπαράστασης.

Επίσης, ένα άλλο πρόβλημα με τη θέση του Frege είναι ότι πολλαπλασιάζονται οι οντότητες: στις δύο οντότητες που είχαμε αρχικά, δηλαδή, την αναπαραστασιακή μορφή και το αντικείμενο στο οποίο αυτή παρέπεμπε, τώρα προστίθεται μια τρίτη οντότητα με ακαθόριστο οντολογικό καθεστώς, η οποία, με τη σειρά της, τροποποιεί μια τέταρτη οντότητα που είναι η ψυχολογική κατάσταση του στοχαστή. Αν το νόημα είναι αυτό που συνδέει τον αναπαραστασιακό φορέα με το αναφερόμενο και τη ψυχολογική κατάσταση του στοχαστή, τότε χρειαζόμαστε μια εξήγηση για το πώς συνδέεται το νόημα με κάθε μία από αυτές τις τρεις οντότητες -εξήγηση που δεν δίνεται¹³⁸.

Τέλος, αν το νόημα ανήκει σε έναν άχρονο και αμετάβλητο κόσμο, τότε όλα τα νοήματα είναι προεδομένα και δεν εξαρτώνται από τις ιστορικές εξελίξεις. Σε αυτή την περίπτωση, όμως, καθίσταται προβληματική η δυνατότητα κατανόησης του νοήματος νέων όρων.

Εξαιτίας των πολλών προβλημάτων που δημιουργεί η φρεγκεανή ερμηνεία του νοήματος, αρκετοί φιλόσοφοι προτίμησαν να εξηγήσουν το νόημα σχεσιακά, δηλαδή να το θεωρήσουν όχι ως αυθύπαρκτη οντότητα αλλά ως ένα σύνολο σχέσεων του χρήστη με την αναπαράσταση, ως ένα σύνολο χρήσεων της αναπαράστασης. Βασικός εκπρόσωπος αυτής της θέσης μπορεί να θεωρηθεί ο ύστερος Wittgenstein¹³⁹, ο οποίος επισημαίνει χαρακτηριστικά: "Για μια μεγάλη κλάση περιπτώσεων - μολονότι όχι για όλες τις περιπτώσεις¹⁴⁰ - όπου τη μεταχειριζόμαστε, η λέξη "σημασία" (meaning) μπορεί να οριστεί έτσι: η σημασία μιας λέξης είναι η χρήση της

¹³⁸ Βλ. ένα συναφές επιχείρημα στο Judge 1983, σελ. 90.

¹³⁹ Βλ. επίσης Ryle (1957, σελ. 145): "οι σημασίες δεν είναι πράγματα, ούτε καν πολύ παράξενα πράγματα. Η εκμάθηση της σημασίας μιας έκφρασης είναι περισσότερο σαν την εκμάθηση μιας άσκησης στη γυμναστική παρά σαν τη συνάντησή με ένα αντικείμενο που δεν έχουμε ξανασυναντήσει. Είναι το να μαθαίνει κανείς να λειτουργεί ορθά με μια έκφραση. ... Αλλά [η χρήση μιας έκφρασης] δεν είναι μια πρόσθετη υπόσταση ή υποκείμενο κατηγορήσεως. Δεν είναι ένα μη-φυσικό, μη-νοητικό αντικείμενο -αλλά όχι επειδή είναι είτε φυσικό είτε νοητικό αντικείμενο, αλλά επειδή δεν είναι αντικείμενο. Καθώς δεν είναι αντικείμενο, δεν είναι κάτοικος μιας πλατωνικής επικράτειας αντικειμένων."

¹⁴⁰ Ο Glock (1996, σελ. 377) επισημαίνει ότι αυτός ο περιορισμός που θέτει ο Wittgenstein δεν αφορά κάποιον συγκεκριμένο τύπο εκφράσεων αλλά μια διαφορετική έννοια του όρου "σημασία": "Δεδομένου ότι ο Wittgenstein δεν είχε κανέναν ενδοιασμό στο να αποδώσει σημασία υπό αυτή την έννοια [δηλ. την έννοια της χρήσης], για παράδειγμα, στα κύρια ονόματα, ο περιορισμός πιθανώς δεν εξαιρεί κάποιους τύπους εκφράσεων, αλλά μια συγκεκριμένη έννοια της 'σημασίας', δηλαδή τη φυσική σημασία (significance), όπως στη πρόταση 'Αυτά τα σύννεφα σημαίνουν βροχή'".

μέσα στη γλώσσα." (Wittgenstein 1953, § 43)¹⁴¹. Ο Wittgenstein, ουσιαστικά, υιοθετεί μια νέα μεταφορά για να μιλήσει για το νόημα: αντί της "οπτικής" μεταφοράς που υποστασιοποιούσε το νόημα περιγράφοντας το ως μια ενδιάμεση οντότητα μεταξύ της γλωσσικής έκφρασης και του αναφερόμενου, υιοθετεί μια μεταφορά από τον χώρο της πράξης¹⁴². Έτσι, το νόημα μιας λέξης κατανοείται ως ο κανόνας χρήσης της εντός ενός γλωσσικού παιχνιδιού. Δεν πρόκειται, ωστόσο, για κάποιους αυστηρούς και αμετάβλητους κανόνες που διέπουν τη χρήση των λέξεων, ούτε πολύ περισσότερο για κάποιους κανόνες που χρειάζεται να διατυπωθούν γλωσσικά προκειμένου να εφαρμοστούν. Η κατανοήση μιας λέξης είναι μια πρακτική γνώση, μια γνώση του πώς πρέπει να χρησιμοποιείται η λέξη σε διάφορα πλαίσια¹⁴³. Υπ' αυτή την έννοια, η "γνώση" του νοήματος μιας λέξης είναι παρόμοια με τη γνώση της χρήσης ενός εργαλείου, δηλαδή την επιδέξια χρήση του για την επίτευξη ενός έργου¹⁴⁴.

Μια παρόμοια, ως ένα βαθμό, κατεύθυνση για την προσέγγιση του νοήματος συναντάμε και στον Morris (1938), ο οποίος εισήγαγε την κλασική διάκριση της σημειωτικής σε σύνταξη, σημασιολογία καιπραγματολογία. Σύμφωνα με αυτή τη διάκριση, η σύνταξη έχει να κάνει με τη σχέση των σημείων μεταξύ τους, η σημασιολογία με τη σχέση των σημείων με τα αντικείμενα στα οποία εφαρμόζονται και η πραγματολογία

¹⁴¹ Σύμφωνα με τον McCulloch (1995, σελ. 80) ο Wittgenstein χρησιμοποιεί εδώ τον όρο "σημασία" με την έννοια που ο Frege χρησιμοποιεί τον όρο "νόημα". Παρότι το έργο του Wittgenstein γενικά εγείρει μεγάλα ερμηνευτικά ζητήματα θα δεχθούμε τη θέση του McCulloch χωρίς περαιτέρω διερεύνηση.

¹⁴² Βλ. Cobb-Stevens 1990, σελ. 37.

¹⁴³ Στο τέταρτο κεφάλαιο θα αναφερθούμε αναλυτικά σε αυτό το πρακτικό είδος γνώσης που αποτελεί όπως θα ισχυρισθούμε ένα είδος πρακτικής γνώσης δεξιοτήτων (know-how)

¹⁴⁴ Μια παρεμφερή ανάλυση του νοήματος μπορεί να συναντήσει κανείς και στην ερμηνευτική παράδοση (βλ. Heidegger 1927/1962). Ωστόσο, αυτό που ανέδειξε κυρίως η ερμηνευτική παράδοση είναι η εκφραστική διάσταση του περιεχομένου των αναπαραστάσεων που συνίσταται στη φανέρωση των όντων (ibid, σελ. 193, 203-204, 205 κ.α.). Χαρακτηριστική είναι η θέση του Taylor ο οποίος, συγκρίνοντας τον τρόπο που κατανοήθηκε η γλώσσα στην αναλυτική και την ηπειρωτική φιλοσοφία, διαπιστώνει ότι η πρώτη έδωσε έμφαση στην παραδειγματική (designative) διάσταση των αναπαραστάσεων, ενώ η δεύτερη στην εκφραστική (expressive) διάσταση των αναπαραστάσεων. Πάντως και η εκφραστική διάσταση του περιεχομένου των αναπαραστάσεων εντοπίζεται, σύμφωνα με τον Taylor, στη σχέση των αναπαραστάσεων με τους χρήστες τους. Αναφερόμενος στις δύο διαστάσεις του περιεχομένου μιας πρότασης επισημαίνει: "Σε κάθε διάσταση σχετίζουμε την πρόταση με κάτι διαφορετικό: με τα αντικείμενα περί των οποίων είναι, στη μία [δηλ. την παραδειγματική διάσταση] και με τη σκέψη που εκφράζει, στην άλλη [δηλ. την εκφραστική διάσταση]" (Taylor 1985β, σελ. 219).

με τη σχέση των σημείων προς τους ερμηνευτές τους¹⁴⁵. Αν παραμείνουμε σε ένα γενικό επίπεδο προσέγγισης, ο τρόπος που ο Morris διακρίνει τη σημασιολογία από την πραγματολογία αντιστοιχεί στον τρόπο που ο Evans διακρίνει την θεωρία αναφοράς από τη θεωρία νοήματος. Με άλλα λόγια, μπορούμε να θεωρήσουμε ότι η πραγματολογία αποτελεί ουσιαστικά μια θεωρία νοήματος η οποία διαφοροποιείται από τη φρεγκεανή θεωρία νοήματος στον τρόπο που αναλύει το νόημα. Σύμφωνα με τον Morris, η πραγματολογική διάσταση μιας αναπαράστασης –ή, με τη δική του ορολογία, ενός σημείου– πρέπει να αναζητηθεί στους τρόπους συμπεριφοράς ενός οργανισμού όταν ανταποκρίνεται στην αναπαράσταση. Πιο συγκεκριμένα, ο Morris, ακολουθώντας τον Peirce, θεωρεί ότι η πραγματολογική διάσταση της αναπαράστασης, ή, με την φρεγκεανή ορολογία, το νόημά της, έγκειται "σε μια συνήθεια (habit) και όχι στην άμεση φυσιολογική αντίδραση που ο φορέας του σημείου προκαλεί ή στα συνοδευόμενα νοητικά είδωλα (mental images¹⁴⁶) ή συναισθήματα" (Morris 1938, σελ. 109).

Το πώς ακριβώς θα πρέπει να κατανοηθεί αυτή η συνήθεια και οι συνακόλουθοι κανόνες χρήσης των αναπαραστάσεων είναι ένα ανοικτό φιλοσοφικό ερώτημα¹⁴⁷. Αυτό που θα κρατήσουμε απλώς από τη επισήμανση του Morris είναι ότι η σχετική απάντηση δεν πρέπει να αναζητηθεί ούτε σε μια νοησιωρχική κατεύθυνση (βλ. νοητικά είδωλα ή συναισθήματα) ούτε σε μια φυσικαλιστική (physicalistic) κατεύθυνση (βλ. άμεση φυσιολογική αντίδραση που ο φορέας του σημείου προκαλεί).

3. Τέλος, μια τρίτη σημαντική επισήμανση είναι ότι δεν έχουν όλες οι γλωσσικές προτάσεις αναφορά. Με αυτό δεν εννοούμε τις γλωσσικές προτάσεις που περιέχουν όρους οι οποίοι αναφέρονται σε μη υπάρχοντα αντικείμενα, αλλά γλωσσικές προτάσεις που δεν είναι ούτε αληθείς ούτε ψευδείς και παρόλα αυτά είναι απόλυτα κατανοητές. Υπάρχουν αμέτρητες τέτοιες γλωσσικές προτάσεις που χρησιμοποιούνται

¹⁴⁵ "[Σύνταξη είναι] η μελέτη των συντακτικών σχέσεων μεταξύ των σημείων απομονώνοντάς τις από τις σχέσεις των σημείων προς τα αντικείμενα ή τους ερμηνευτές" (Morris 1938, σελ. 91). "Μπορεί κανείς να μελετήσει τις σχέσεις των σημείων προς τα αντικείμενα στα οποία τα σημεία είναι εφαρμόσιμα. Αυτή η σχέση θα ονομαστεί *σημασιολογική διάσταση της σημείωσης* ... η μελέτη αυτής της διάστασης θα ονομαστεί *σημασιολογία*. Η το αντικείμενο μελέτης θα μπορούσε να είναι η σχέση των σημείων προς τους ερμηνευτές. Αυτή η σχέση θα ονομαστεί *πραγματολογική διάσταση της σημείωσης* ... και η μελέτη αυτής της διάστασης θα ονομαστεί *πραγματολογία*." (Morris 1938/1971, σελ. 84). Επίσης, "Η σημασιολογία ασχολείται με τη σχέση των σημείων προς ...τα αντικείμενα στα οποία μπορεί ή όχι να παραπέμπουν. Η πραγματολογία αφορά τις σχέσεις των σημείων προς τους ερμηνευτές τους" (Morris 1938/1971, σελ. 35, 43).

¹⁴⁶ Βλ. κεφάλαιο 5, υποσημείωση 222.

¹⁴⁷ Βλ., ενδεικτικά, McDowell 1995, Haugeland 1998a, Τσινόρεμα 1995. Βλ., επίσης, κεφάλαιο 4, ενότητα 4.3.

συνεχώς στο λόγο¹⁴⁸. Για παράδειγμα, η πρόταση "Πήγαινε να βρεις το λιοντάρι" είναι μια καθ' όλα κατανοητή πρόταση. Ωστόσο, δεν έχει κανένα νόημα να ρωτήσουμε αν η συγκεκριμένη πρόταση είναι αληθής ή ψευδής. Μια προσταγή, σε αντίθεση με μια απόφαση (assertion), δεν αναφέρεται σε κάτι ως ισχύον για τον κόσμο και, κατά συνέπεια, δεν έχει νόημα το να εξεταστεί εάν ανταποκρίνεται ή όχι στην πραγματικότητα αυτό στο οποίο υποτίθεται ότι αναφέρεται. Το ίδιο ισχύει όχι μόνο όταν προστάζουμε αλλά και όταν υποσχόμαστε, όταν ρωτάμε, όταν ανακηρύσσουμε, όταν προειδοποιούμε, όταν επικρίνουμε, όταν επιδοκιμάζουμε, όταν απολογούμαστε, όταν στοιχηματίζουμε, όταν βαπτίζουμε κ.λπ. Όλα αυτά τα ρήματα, όταν χρησιμοποιούνται σε μια πρόταση στο πρώτο ενικό πρόσωπο του ενεστώτα, δεν αναφέρουν (report) την επιτέλεση μιας πράξης (π.χ. μια υπόσχεση), αλλά αποτελούν τα ίδια, όταν εκφωνούνται, την επιτέλεση της πράξης (Austin 1995, σελ. 217). Όταν, υπό τις κατάλληλες συνθήκες, πω "κηρύσσω την έναρξη των εργασιών" δεν αναφέρομαι το γεγονός ότι κήρυξα την έναρξη των εργασιών, αλλά επιτελώ (perform) την πράξη της κήρυξης των εργασιών. Η διαφορά γίνεται φανερή αν χρησιμοποιήσουμε το ρήμα σε άλλο πρόσωπο ή σε παρελθοντικό χρόνο. Αν πούμε "ο Χ κήρυξε την έναρξη των εργασιών την 1^η Οκτωβρίου" δεν επιτελούμε την πράξη της ανακήρυξης γιατί δεν είμαστε ο Χ, αλλά ακόμα και αν ήμασταν ο Χ πάλι δεν θα μπορούσαμε να επιτελέσουμε την πράξη γιατί η πράξη είναι συντελεσμένη. Με την πρόταση "ο Χ κήρυξε την έναρξη των εργασιών την 1^η Οκτωβρίου" απλώς αναφερόμαστε στο γεγονός της κήρυξης των εργασιών.

Θα μπορούσε, λοιπόν, κανείς να ισχυριστεί ότι με μια γλωσσική πρόταση έχουμε τη δυνατότητα να κάνουμε δύο πράγματα: είτε να αναφερθούμε σε μια κατάσταση πραγμάτων, είτε να επιτελέσουμε μια πράξη - ένα ομιλιακό ενέργημα (speech act). Στην πρώτη περίπτωση, η πρόταση έχει αναφορά, με την έννοια ότι μπορεί να είναι είτε αληθής είτε ψευδής, ενώ στην δεύτερη περίπτωση δεν αναφέρεται σε κάτι και, υπό αυτή την έννοια, δεν είναι ούτε αληθής ούτε ψευδής. Οπότε θα μπορούσαμε να πούμε ότι οι προτάσεις της δεύτερης κατηγορίας είναι προτάσεις με νόημα αλλά χωρίς αναφορά.

Ωστόσο, τα πράγματα δεν είναι τόσο απλά. Όπως διαπίστωσε ο Austin, η διάκριση μεταξύ καθαρά επιτελεστικών (performative) ρημάτων

¹⁴⁸ Πρόκειται για ένα εμπειρικό δεδομένο που αδυνατεί να εξηγήσει η επαληθευσιοκρατική θεωρία της γλωσσικής σημασίας (verification theory of meaning). Σύμφωνα με αυτή τη θεωρία, "μια πρόταση θεωρείται ότι έχει κυριολεκτικά σημασία αν και μόνο αν είναι είτε αναλυτική είτε εμπειρικά επαληθεύσιμη" (Ayer 1936, σελ. 12) (βλ. και υποσημείωση 42). Αλλά και γενικότερα, κάθε θεωρία αναπαραστασιακού περιεχομένου που βασίζεται στην έννοια της αλήθειας -όπως κατανοείται στο πλαίσιο της σημασιολογίας συνθηκών αλήθειας - αντιμετωπίζει προβλήματα στην εξήγηση του περιεχομένου εκφράσεων που, ενώ δεν είναι ούτε αληθείς ούτε ψευδείς, είναι κατανοητές.

και διαπιστωτικών (constative) ρημάτων, δηλαδή ρημάτων που χρησιμοποιούνται απλώς για να αναφέρουν κάποιο γεγονός, δεν ισχύει. Ακόμα και όταν χρησιμοποιούμε τα υποτιθέμενα διαπιστωτικά ρήματα επιτελούμε κάποιο ενέργημα. Όταν λέω "δηλώνω ότι το λιοντάρι είναι πάνω στο τραπέζι" δεν αναφέρομαι (refer) απλώς στο προτασιακό περιεχόμενο *p*—ότι, δηλαδή, το λιοντάρι είναι πάνω στο τραπέζι-, αλλά και αναφέρω (report) ότι το *p* ισχύει, δηλαδή επιτελώ ένα ομιλιακό ενέργημα γιατί δεσμεύομαι ότι το συγκεκριμένο προτασιακό περιεχόμενο είναι αληθές¹⁴⁹.

Όμως θα μπορούσε κανείς να αντιτείνει ότι όλα τα παραπάνω περί επιτελεστικών ενεργημάτων αφορούν κάποια πολύ εκλεπτυσμένα γλωσσικά φαινόμενα και κάθε άλλο παρά χαρακτηρίζουν το σύνολο των εξωτερικών αναπαραστάσεων. Με άλλα λόγια, φαίνεται να έχουμε οδηγηθεί στην εξέταση ορισμένων εξειδικευμένων χρήσεων των αναπαραστάσεων οι οποίες απαντώνται μόνο σε πολύ αναπτυγμένα αναπαραστασιακά συστήματα όπως είναι οι φυσικές γλώσσες. Αυτό, όμως, που μας ενδιαφέρει στην παρούσα ενότητα είναι ένας χαρακτηρισμός των αναπαραστάσεων ως γένους και, κατά συνέπεια, η μελέτη της επιτελεστικής χρήσης των ομιλιακών ενεργημάτων δεν συμβάλλει καθόλου προς αυτή την κατεύθυνση. Γι' αυτό είναι σκόπιμο να επιστρέψουμε και να παραμείνουμε στο αρχικό χαρακτηριστικό που εντοπίσαμε ότι διέπει τις αναπαραστάσεις, ήτοι, ότι όλες οι αναπαραστάσεις έχουν μια αναφορική λειτουργία.

Ωστόσο, η παραπάνω αντίρρηση δεν είναι ορθή. Καταρχάς, όπως ήδη είδαμε, δεν μπορούν να διακριθούν οι διαπιστωτικές χρήσεις των γλωσσικών προτάσεων από τις επιτελεστικές γιατί ακόμα και μια δήλωση (statement) ή μια αναφορά (report) αποτελούν επιτελεστικά ενεργήματα. Δηλαδή, σε κάθε γλωσσική χρήση υπάρχει μια επιτελεστική διάσταση. Αντίθετα, όπως είδαμε, δεν είναι όλες οι γλωσσικές προτάσεις είτε αληθείς είτε ψευδείς, ή, με άλλα λόγια, δεν έχουν όλες οι προτάσεις αναφορική διάσταση. Κατά συνέπεια, αν υπάρχει κάτι που χαρακτηρίζει γενικά τις εκφράσεις της φυσικής γλώσσας αυτό είναι η επιτελεστική τους λειτουργία και όχι η αναφορική.

Ισχύει, άραγε, το ίδιο και για λιγότερο εκλεπτυσμένα αναπαραστασιακά συστήματα όπως τα μη γλωσσικά αναπαραστασιακά συστήματα, οι πρωτόγονες γλώσσες ή οι γλώσσες των ζώων; Αυτό

¹⁴⁹ Η διαπίστωση αυτή οδήγησε τον Austin να αναλύσει κατά έναν διαφορετικό τρόπο τα ομιλιακά ενεργήματα. Σύμφωνα με τη νέα πρόταση, κάθε ομιλιακό ενέργημα αποτελείται από το λεκτικό (locutionary) ή προτασιακό περιεχόμενο και την ελλεκτική δύναμη (illocutionary force). Έτσι, το ίδιο προτασιακό περιεχόμενο μπορεί, ανάλογα με την ελλεκτική δύναμη, να συνιστά διαφορετικό είδος ομιλιακού ενεργήματος. Ωστόσο, αυτή η λύση δεν είναι χωρίς προβλήματα (βλ. Lycan 1999, σελ. 181-184).

αποτελεί ένα ανοικτό εμπειρικό ζήτημα. Ένα σημαντικό ερώτημα, εν προκειμένω, είναι αν τέτοιου είδους συστήματα μπορούν καν να εκφράσουν προτασιακά περιεχόμενα ή συστατικά αυτών. Ας δούμε και πάλι το παράδειγμα με τα σήματα κινδύνου των ζώων. Υπό το φως όσων είπαμε σε αυτή την τρίτη επισήμανση, προκύπτει μια εναλλακτική ερμηνεία για το τι προσλαμβάνουν τα ζώα που ακούν το σήμα κινδύνου. Ενδεχομένως μια κραυγή κινδύνου δεν αναφέρεται σε ένα αρπαικτικό που πλησιάζει ούτε καν το ενδεικνύει, αλλά απλά προστάζει το ζώο που ακούει αυτή την κραυγή να κρυφτεί¹⁵⁰. Σύμφωνα με αυτή την ερμηνεία λοιπόν, μια κραυγή θα μπορούσε να είναι μια προσταγή χωρίς προτασιακό περιεχόμενο. Δύο διαφορετικά προστακτικά σήματα αλλά επιτάσσουν δύο διαφορετικές συμπεριφορές και μόνο καταχρηστικά θα μπορούσαμε να πούμε ότι "περγράφουν" δύο διαφορετικές συμπεριφορές. Μάλιστα, ορισμένοι ερευνητές¹⁵¹ υποστηρίζουν ότι τα προστακτικά σήματα προηγούνται οντογενετικά και φυλογενετικά των ενδεικτικών σημάτων. Η θέση αυτή είναι αρκετά εύλογη για δύο τουλάχιστον λόγους: α) τα προστακτικά σήματα, σε αντίθεση με τα ενδεικτικά, συνδέονται άμεσα με τη συμπεριφορά και β) τα ενδεικτικά σήματα προϋποθέτουν τη δυνατότητα αναγνώρισης ενός αντικειμενικού κόσμου, δηλαδή προϋποθέτουν μια πολύ εξελιγμένη νοητική λειτουργία.

Δεν θα επεκταθούμε περισσότερο σε αυτό το ζήτημα, απλά θα υιοθετήσουμε ως εύλογη τη διαπίστωση ότι αν υπάρχει κάτι που χαρακτηρίζει γενικά τα γλωσσικά συστήματα (εξελιγμένα ή μη) αυτό είναι η επιτελεστική τους λειτουργία και όχι η αναφορική. Σύμφωνα με όσα είπαμε στη δεύτερη επισήμανση, η επιτελεστική λειτουργία αφορά τη σχέση του χρήστη με τις αναπαραστάσεις, δηλαδή την πραγματολογική διάσταση των αναπαραστάσεων¹⁵². Και αν όντως το νόημα αφορά τη σχέση του χρήστη με τις αναπαραστάσεις, τότε η επιτελεστική λειτουργία μπορεί να εκληφθεί ως ένα είδος νοήματος. Σε αυτή τη περίπτωση, όμως, πρόκειται για ένα νόημα που μπορεί να αναλυθεί ανεξάρτητα από κάθε

¹⁵⁰ Αυτή η ερμηνεία είναι αρκετά εύλογη τουλάχιστον για την περίπτωση των κατώτερων οργανισμών. Βέβαια παραμένει ανοικτό το ζήτημα όσον αφορά το ποιες συνθήκες πρέπει να ικανοποιούνται προκειμένου μια κραυγή να συνιστά προσταγή. Αυτό είναι ένα μεγάλο ζήτημα που δεν θα διερευνήσουμε περαιτέρω. Θα επιστημόνουμε απλώς ότι μια βασική συνθήκη για να συνιστά μια έκφραση προσταγή είναι ενδεχομένως το να μπορεί ο οργανισμός που τη δέχεται να αρνηθεί να την εκτελέσει.

¹⁵¹ Βλ., ενδεικτικά, Skinner 1957, Papineau 1998, Place 2000, Hooker (προσωπική επικοινωνία). Πρβλ., επίσης, Wittgenstein (1980, σελ. 31): "Η καταγωγή και η πρωτογενής μορφή του γλωσσικού παιχνιδιού είναι μια αντίδραση [Reaktion]: μόνο από αυτή μπορούν να αναπτυχθούν πιο πολύπλοκες μορφές. Η γλώσσα, θέλω να πω, είναι μια εκλέπτυνση και εν αρχή ήταν η πράξη", πρβλ. μετάφραση Κ. Κωβαίου στο Wittgenstein 1986, σελ. 56.

¹⁵² Αλλά δεν την εξαντλεί. Βλ., για παράδειγμα, Grice 1989.

είδους αναφορά. Η έννοια της αναπαράστασης, ωστόσο, είναι άρρηκτα συνδεδεμένη με την έννοια της αναφοράς. Υπ' αυτή την έννοια, εκφράσεις που έχουν επιτελεστική λειτουργία χωρίς να έχουν αναφορική λειτουργία δεν μπορούν να ονομαστούν αναπαραστάσεις. Η έκφραση "Ζήτω!" δεν είναι αναπαράσταση. Αναδεικνύεται έτσι ένα όριο της έννοιας της αναπαράστασης. Η έννοια της (αυτόνομης και κατανοητής) γλωσσικής έκφρασης δεν ταυτίζεται με την έννοια της αναπαράστασης: υπάρχουν γλωσσικές εκφράσεις που δεν είναι αναπαραστάσεις¹⁵³.

3.3.1.3 Η προέλευση της αναπαραστασιακής σχέσης

Μέχρι στιγμής έχουμε προσδιορίσει - ως ένα βαθμό - τι σημαίνει για κάτι να έχει αναπαραστασιακό περιεχόμενο. Είδαμε ότι η έννοια του περιεχομένου μπορεί να αναλυθεί περαιτέρω με βάση τις έννοιες της αναφοράς και του νοήματος. Μπορούμε τώρα να εξετάσουμε το δεύτερο σημαντικό ερώτημα που θέσαμε στην αρχή της προηγούμενης ενότητας αναφορικά με την έννοια του περιεχομένου: πώς μπορεί ένα πράγμα να αποκτήσει αναπαραστασιακό περιεχόμενο; Με άλλα λόγια, ποια είναι η προέλευση του περιεχομένου των αναπαραστάσεων;

Με δεδομένη την ανάλυση της προηγούμενης ενότητας, το ερώτημα αυτό αναλύεται σε δύο υποερωτήματα: (α) πώς μπορεί ένα πράγμα να αποκτήσει αναφορά και (β) πώς μπορεί ένα πράγμα να αποκτήσει νόημα; Ωστόσο, το να αναλύουμε κατ' αυτόν τον τρόπο το ερώτημα, ίσως, οδηγεί σε μια υπερβολικά σχηματική και απλουστευτική αντιμετώπιση του προβλήματος γιατί φαίνεται σαν να αποδεχόμαστε ότι κάτι μπορεί να αποκτήσει αναφορά ανεξάρτητα από το αν έχει αποκτήσει νόημα και το αντίστροφο. Αν λάβουμε υπόψη την ανάλυση που κάναμε για να διακρίνουμε τα ενδεικτικά σήματα από τις αναπαραστάσεις με την πλήρη έννοια του όρου, προκύπτει ότι η "ύπαρξη" αναφοράς (με την έννοια της παραπομπής) προϋποθέτει την "ύπαρξη" νοήματος. Πιο συγκεκριμένα, είδαμε ότι προκειμένου κάτι να έχει αναφορά, δηλαδή να είναι αναπαράσταση με την πλήρη έννοια του όρου, θα πρέπει ο χρήστης να μην το εκλαμβάνει ως ένα εγγενές χαρακτηριστικό του αναφερόμενου. Έτσι, το τρένο στην ταινία των αδελφών Lumière δεν αποτελεί αναπαράσταση για τους θεατές που το βάζουν στα πόδια γιατί, απλούστατα, δεν το διακρίνουν από ένα πραγματικό τρένο. Με γενικότερους όρους, διαπιστώσαμε ότι συνθήκη δυνατότητας για την κατανόηση αναφορικών σχέσεων είναι η ικανότητα απόσπασης από το άμεσο περιβάλλον, μια ικανότητα, ακριβώς, που μας διανοίγει στον χώρο του νοήματος. Αν δεχθούμε, λοιπόν, αυτή την ανάλυση, τότε προκύπτει ότι συνθήκη δυνατότητας για να "αποκτήσει" κάτι αναφορά είναι να έχει

¹⁵³ Όπως, αντίθετα, υπάρχουν αναπαραστάσεις που δεν είναι γλωσσικές. Βλ., παραδείγματος χάρι, τις εικόνες και τους ενδείκτες.

νόημα¹⁵⁴. Το αντίστροφο δεν ισχύει, όπως δείξαμε στην τρίτη επισήμανση της προηγούμενης ενότητας: η γλώσσα αφθονεί από προτάσεις που έχουν νόημα και για τις οποίες η διάκριση αληθής/ψευδής δεν είναι εφαρμόσιμη.

3.3.1.3.1 Θεωρίες προέλευσης του νοήματος

Αν δεχθούμε την παραπάνω θέση, τότε έχουμε άμεσα και την απάντηση στο ερώτημα πώς μπορεί ένα πράγμα να αποκτήσει αναφορά: με το να αποκτήσει νόημα. Το κρίσιμο, λοιπόν, ερώτημα για όσους επιλέγουν αυτή την οδό καθίσταται το πώς ένα πράγμα αποκτά νόημα. Όπως είδαμε στη δεύτερη επισήμανση της προηγούμενης ενότητας, η θεωρία νοήματος έχει να κάνει με τη σχέση αναπαράστασης - υποκειμένου που την κατανοεί. Εκεί, λοιπόν, πρέπει να ψάξουμε την απάντηση στο ερώτημα πώς ένα πράγμα αποκτά νόημα. Ωστόσο, η θεωρία του Frege δεν δίνει απάντηση σε αυτό το ερώτημα, καθότι τα φρεγκεανά νοήματα προϋπάρχουν και δεν είναι σαφές το πώς "συμπλέκονται" με συγκεκριμένα αντικείμενα προκειμένου να τα καταστήσουν αναπαραστάσεις. Επιπλέον, όπως είδαμε, η φρεγκεανή θεωρία για το τι είναι το νόημα παρουσιάζει από μόνη της ανυπερβλήτα προβλήματα που μας υποχρεώνουν να την εγκαταλείψουμε. Αντίθετα, η βιτγκενσταϊνική προσέγγιση, η οποία, ως ένα βαθμό, συγγενεύει και με την πραγματιστική προσέγγιση του Μοτίς, παρέχει άμεσα μια απάντηση στο ερώτημα πώς ένα πράγμα αποκτά νόημα: από τον τρόπο χρήσης του. Υιοθετώντας αυτή την προσέγγιση, ο προκαταρκτικός ορισμός της έννοιας της αναπαράστασης μπορεί να συμπληρωθεί ως εξής:

Ορισμός 6:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου αντικειμένου, ιδιότητας, σχέσης ή κατάστασης πραγμάτων O , δηλαδή έχει ως περιεχόμενο το O , αν για κάποιον επαρκή χρήστη X η μορφή M του A μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο O (αναφερόμενο) υπό μια συγκεκριμένη προοπτική N (νόημα), δυνάμει του τρόπου που ο X χρησιμοποιεί το A .
------------	---

Μια δεύτερη προσέγγιση για το πώς κάτι αποκτά νόημα στηρίζεται στην έννοια της ερμηνείας. Για πολλούς φιλοσόφους θεωρείται σχεδόν αυτονόητο ότι μια αναπαράσταση είναι αναπαράσταση δυνάμει κάποιου ερμηνευτή¹⁵⁵. Χωρίς ερμηνευτές δεν θα υπήρχαν αναπαραστάσεις

¹⁵⁴ Η θέση αυτή σχετίζεται με τη φρεγκεανή θέση ότι το νόημα καθορίζει την αναφορά.

¹⁵⁵ Ορισμένοι διαβλέπουν αυτή τη θέση στο έργο του Locke (βλ. Blackburn 1984, σελ. 41), ωστόσο χαρακτηριστικότερος εκφραστής της είναι ο Grice (1957) ο οποίος ορίζει τη (μη-φυσική) σημασία των συμβόλων με βάση τις επικοινωνιακές προθέσεις των ομιλητών. Βλ., επίσης, Searle 1983. Από το χώρο της

στον κόσμο. Ο ερμηνευτής είναι αυτός που αποδίδει πρωτογενώς το νόημα σε ορισμένα αντικείμενα καθιστώντας τα αναπαράστασις. Αν υιοθετηθεί κανείς αυτή τη θεωρία, ο προκαταρκτικός ορισμός της έννοιας της αναπαράστασης μπορεί, εναλλακτικά, να συμπληρωθεί ως εξής:

Ορισμός 7:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου αντικειμένου, ιδιότητας, σχέσης ή κατάστασης πραγμάτων O , δηλαδή έχει ως περιεχόμενο το O , αν για κάποιον επαρκή χρήση X η μορφή M του A μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο O (αναφερόμενο) υπό μια συγκεκριμένη προοπτική N (νόημα), δυνάμει του ίδιου του X που λειτουργεί ως ερμηνευτής του A .
------------	--

Σε μια τέτοιου τύπου προσέγγιση η έννοια της αποβλεπτικότητας θεωρείται εξηγητικά πρότερη της έννοιας της αναπαράστασης: οι προτασιακές στάσεις και, ειδικότερα, οι προθέσεις του ερμηνευτή είναι εκείνες που καθιστούν κάτι αναπαράσταση. Σύμφωνα με την ορολογία του Searle (1992), οι εξωτερικές αναπαράστασις έχουν *παράγωγη* αποβλεπτικότητα και μόνο ο νους έχει *εγγενή* αποβλεπτικότητα¹⁵⁶.

Υπάρχει, τέλος, μια τρίτη προσέγγιση για το πώς κάτι αποκτά νόημα. Η προσέγγιση αυτή βασίζεται στην έννοια της συστημικής λειτουργίας ή της δομικής σχέσης: ένα αντικείμενο αναπαριστά κάποιο άλλο χάρις στον λειτουργικό ρόλο ή τις δομικές σχέσεις που έχει μέσα σε ένα σύστημα άλλων αντικειμένων. Η προσέγγιση αυτή ονομάζεται, μεταξύ άλλων, "σημασιολογία λειτουργικού ρόλου" και χαρακτηριστικότερος σύγχρονος υποστηρικτής της είναι ο Block (1986)¹⁵⁷. Η συγκεκριμένη

φαινομενολογίας χαρακτηριστική είναι προς αυτή την κατεύθυνση η ακόλουθη επισήμανση του Merleau-Ponty (1962, σελ. 413): "Αυτό το τραπέζι φέρει ίχνη της παρελθούσας ζωής μου γιατί έχω χαράξει κάποια αρχικά πάνω σ' αυτό και έχω χρώσει μελάνι πάνω του. Αλλά αυτά τα ίχνη από μόνα τους δεν αναφέρονται στο παρελθόν: είναι παρόντα και, στο βαθμό που βρίσκα σε αυτά σημεία κάποιου 'προηγούμενου' συμβάντος, είναι επειδή εγώ μεταφέρω αυτή τη συγκεκριμένη σημασία (significance) εντός μου".

¹⁵⁶ Βλ. κεφάλαιο 7, ενότητα 7.2.

¹⁵⁷ Βλ. κεφάλαιο 7, ενότητα 7.2.5.1. Στοιχεία της συγκεκριμένης προσέγγισης μπορεί να βρει κανείς και σε θεωρίες που δεν είναι, τουλάχιστον εκ πρώτης όψεως, φυσικοποιητικές, όπως ο γαλλικός δομισμός και ο συναγωγισμός (inferentialism) (Sellars 1980, Brandom 1994). Χαρακτηριστική διατύπωση της δομιστικής προσέγγισης είναι η ακόλουθη: "Αυτό που κάνει τη λέξη 'ήλιος' να παραπέμπει (designate) στον ήλιο δεν είναι η ομοιότητα μεταξύ της λέξης και του πράγματος, ούτε είναι ο εσωτερικός χαρακτήρας του ενός ή του άλλου. Είναι, μάλλον, η σχέση μεταξύ της λέξης 'ήλιος' και του συνόλου όλων των ελληνικών λέξεων είναι ο τρόπος με τον οποίον διαφοροποιείται από αυτές. Η λέξη έχει σημασία (meaning) μόνο μέσω όλου του θεσμού της γλώσσας.... Το πιο ακριβές χαρακτηριστικό μιας λέξης είναι 'τι οι άλλες δεν είναι'. Σημασία (signification) υφίσταται όχι για μια

προσέγγιση είναι εγγενώς ολιστική, διότι επιχειρεί να εξηγήσει το πώς ένα αντικείμενο αποκτά την αναπαραστασιακή του λειτουργία ανατρέχοντας στη συντακτική σχέση που αυτό έχει με ένα σύνολο άλλων αντικειμένων που λειτουργούν ως αναπαραστάσεις και όχι στη σχέση του με το αναπαριστώμενο. Πρόκειται, ουσιαστικά, για μια προσέγγιση που δίνει έμφαση στη συντακτική διάσταση των αναπαραστάσεων, στον τρόπο, δηλαδή, που οι διάφορες αναπαραστασιακές μορφές σχετίζονται μεταξύ τους. Σύμφωνα με τη σημασιολογία λειτουργικού ρόλου, δεν είναι δυνατόν να υπάρξει μία μεμονωμένη αναπαράσταση: για να υπάρξει μία αναπαράσταση θα πρέπει να υπάρχουν και άλλες.

Μια έννοια που εισάγεται από τους υποστηρικτές της συγκεκριμένης προσέγγισης είναι η έννοια του αναπαραστασιακού συστήματος: ένα αντικείμενο αποκτά την αναπαραστασιακή του λειτουργία μόνο στο πλαίσιο ενός αναπαραστασιακού συστήματος, δηλαδή ενός συστήματος που απαρτίζεται από ένα σύνολο χαρακτήρων και ένα σύνολο κανόνων για τον χειρισμό αυτών των χαρακτήρων. Χαρακτηριστικό παράδειγμα αποτελούν οι γλωσσικές αναπαραστάσεις (ή, τουλάχιστον, ορισμένες από αυτές) το νόημα των οποίων προκύπτει, σύμφωνα με τους υποστηρικτές της εν λόγω προσέγγισης, από τον τρόπο που χρησιμοποιούνται εντός του γλωσσικού συστήματος. Για παράδειγμα, το νόημα του λογικού συνδέσμου "και" προσδιορίζεται από τις συναγωγές τις οποίες επιτρέπει (π.χ. από τις προκειμένες p και q μπορούμε να συναγάγουμε το $p \wedge q$)¹⁵⁸. Αν υιοθετήσει κανείς αυτή την τρίτη προσέγγιση, προκύπτει ο εξής προκαταρκτικός ορισμός για το τι είναι αναπαράσταση:

Ορισμός 8:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου αντικειμένου, ιδιότητας, σχέσης ή κατάστασης πραγμάτων O , δηλαδή έχει ως περιεχόμενο το O , αν για κάποιον επαρκή χρήση X η μορφή M του A μέσα στο πλαίσιο ενός αναπαραστασιακού συστήματος μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο O (αναφερόμενο) υπό μια συγκεκριμένη προοπτική N (νόημα), χάρις στην δομική ή λειτουργική σχέση του Φ με τους υπόλοιπους φορείς του αναπαραστασιακού συστήματος.
------------	---

λέξη αλλά για όλες τις λέξεις στη μεταξύ τους σχέση" (Merleau-Ponty 1973, σελ. 80, 98).

¹⁵⁸ Ωστόσο, όπως ήδη διαπιστώσαμε στην ενότητα 3.3.1.2, δεν έχει νόημα να μιλήσουμε για αναφορά στην περίπτωση του "και". Αυτό είναι ένα γενικότερο πρόβλημα της εξεταζόμενης προσέγγισης. Αν το αναπαραστασιακό περιεχόμενο μιας έκφρασης καθορίζεται πλήρως από τις συντακτικές σχέσεις της με τις υπόλοιπες εκφράσεις του αναπαραστασιακού συστήματος, πώς μπορεί να προσδιοριστεί η σχέση της έκφρασης με τον κόσμο;

Η τρίτη αυτή προσέγγιση συγχέεται, κακώς κατά τη γνώμη μου, με την πρώτη προσέγγιση, δηλαδή τη βιτγκεσταϊνική προσέγγιση που εξηγεί το νόημα μιας αναπαράστασης με βάσει τον τρόπο χρήσης. Ωστόσο, η σημασιολογία λειτουργικού ρόλου αφορά το συντακτικό επίπεδο γιατί ενδιαφέρεται για το ποιες αναπαραστασιακές μορφές σχετίζονται με μια συγκεκριμένη αναπαραστασιακή μορφή, ενώ η βιτγκεσταϊνική προσέγγιση αφορά το πραγματολογικό επίπεδο γιατί μελετά σε ποια πρακτικά πλαίσια και υπό ποιες συνθήκες χρησιμοποιείται μια συγκεκριμένη αναπαράσταση¹⁵⁹.

3.3.1.3.2 Θεωρίες προέλευσης της αναφοράς

Μέχρι στιγμής αναζητήσαμε την απάντηση στο ερώτημα πώς κάτι αποκτά αναπαραστασιακό περιεχόμενο στο πώς το ίδιο αποκτά νόημα. Γι' αυτό αναφερθήκαμε στις θεωρίες προέλευσης του αναπαραστασιακού περιεχομένου που δίνουν προτεραιότητα στο νόημα έναντι της αναφοράς, υπό την έννοια ότι το νόημα καθορίζει την αναφορά. Ωστόσο, υπάρχουν αρκετοί που υιοθετούν άλλη αφετηρία. Για την ακρίβεια, θεωρούν ότι μια αναπαράσταση μπορεί να έχει αναφορά χωρίς να έχει νόημα. Κατά συνέπεια, στην προκειμένη περίπτωση, το ερώτημα πώς ένα πράγμα αποκτά αναπαραστασιακό περιεχόμενο ισοδυναμεί με το ερώτημα πώς αυτό αποκτά αναφορά. Η αναφορική σχέση, όπως φάνηκε με τη δεύτερη επισήμανση της προηγούμενης ενότητας, έχει να κάνει με τη σχέση του αναπαραστασιακού φορέα με το αναπαριστώμενο. Έτσι, οι προσεγγίσεις που επιχειρούν να απαντήσουν άμεσα στο ερώτημα πώς κάτι αποκτά αναφορά, δηλαδή χωρίς την επίκληση της έννοιας του νοήματος, στοχεύουν στην εξήγηση της αναφορικής σχέσης με βάση κάποια ειδικού τύπου σχέση μεταξύ του αναπαραστασιακού φορέα και του αναπαριστώμενου. Υπάρχουν, τουλάχιστον, δύο σχετικές προσεγγίσεις.

Η πρώτη βασίζεται στη σχέση ομοιότητας: ένα αντικείμενο καθίσταται αναπαράσταση ενός άλλου αντικείμενου στο βαθμό που ομοιάζει ή, γενικότερα, παρουσιάζει ένα δομικό ισομορφισμό προς αυτό. Χαρακτηριστικό παράδειγμα θεωρούνται οι εικόνες¹⁶⁰, οι οποίες αποκτούν την αναπαραστασιακή τους λειτουργία χωρίς στην (υποτιθέμενη¹⁶¹) ομοιότητά τους προς το αναπαριστώμενο¹⁶². Αν υιοθετήσει κανείς αυτή

¹⁵⁹ Το εκάστοτε πλαίσιο δεν είναι κατ' ανάγκη αποκλειστικά εξωγλωσσικό: η γνώση της χρήσης μιας αναπαράστασης περιλαμβάνει και τη γνώση των συμφραστικών πλαισίων στα οποία μπορεί να χρησιμοποιείται.

¹⁶⁰ Ωστόσο, μια παρόμοιος τύπου θεωρία έχει προταθεί, σύμφωνα με μια ερμηνεία (Haugeland 1991), ακόμα και για τις γλωσσικές αναπαραστάσεις από τον Wittgenstein 1921/1971.

¹⁶¹ Βλ. επόμενη ενότητα και παράρτημα 1.

¹⁶² Βλ. Cummins 1996.

την προσέγγιση, προκύπτει ο εξής προκαταρκτικός ορισμός για το τι είναι αναπαράσταση:

Ορισμός 9:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου αντικειμένου, ιδιότητας, σχέσης ή κατάστασης πραγμάτων O , δηλαδή έχει ως περιεχόμενο το O , αν για κάποιον επαρκή χρήστη X η μορφή M του A μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο O (αναφερόμενο) χάρις σε μια σχέση ομοιότητας μεταξύ χαρακτηριστικών του αναπαραστασιακού φορέα Φ και του αναφερόμενου O .
------------	--

Η δεύτερη προσέγγιση βασίζεται στη σχέση αιτιότητας. Ένα αντικείμενο αναπαριστά κάποιο άλλο στο βαθμό που έχει μια ειδικού τύπου αιτιακή σχέση με αυτό: ο καπνός είναι σημάδι φωτιάς και οι δακτύλιοι στον κορμό του δένδρου είναι ενδείκτες (indices) της ηλικίας του¹⁶³. Σε αυτή την περίπτωση, ο προκαταρκτικός ορισμός της αναπαράστασης τροποποιείται ως εξής:

Ορισμός 10:	Ένα αντικείμενο A συνιστά αναπαράσταση κάποιου αντικειμένου, ιδιότητας, σχέσης ή κατάστασης πραγμάτων O , δηλαδή έχει ως περιεχόμενο το O , αν για κάποιον επαρκή χρήστη X η μορφή M του A μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο O (αναφερόμενο) χάρις σε μια σχέση αιτιότητας μεταξύ του αναφερόμενου O και του αναπαραστασιακού φορέα Φ .
-------------	--

3.3.1.4 Είδη αναπαραστάσεων

Οι εναλλακτικοί ορισμοί που διατυπώσαμε στην προηγούμενη ενότητα, δηλαδή οι ορισμοί 6-10 δημιουργούν άμεσα ένα πρόβλημα: εάν στόχος μας είναι να προσδιορίσουμε το τι συνιστά (εξωτερική) αναπαράσταση εν γένει, τότε, εκ πρώτης όψεως, μόνο ένας από αυτούς θα ισχύει και άρα οι υπόλοιποι θα πρέπει να είναι λανθασμένοι. Τα πράγματα, ωστόσο, δεν είναι τόσο απλά. Καταρχάς, δεν είναι σαφές ότι ο τρόπος με τον οποίο μια αναπαράσταση αποκτά περιεχόμενο χαρακτηρίζει την αναπαράσταση ως γένος. Με άλλα λόγια, παραμένει υπό διευκρίνηση το εάν θα πρέπει να ενσωματωθεί στον ορισμό της αναπαράστασης ως γένους ο τρόπος με τον οποίον οι αναπαραστάσεις αποκτούν περιεχόμενο. Οι

¹⁶³ Βλ., για παράδειγμα, τη θέση του Grice (1957) περί φυσικής σημασίας και Dretske 1988 (πρβλ. υποσημείωση 165). Επίσης, στην ίδια γενική προσέγγιση θα μπορούσε πιθανόν να ενταχθεί και η αιτιακή θεωρία αναφοράς του Putnam (1975) (βλ. κεφάλαιο 7, ενότητα 7.2.3).

πέντε προαναφερθείσες προτάσεις για την προέλευση της αναπαραστασιακής σχέσης δεν είναι, κατ' ανάγκη, αποκλειστικές διότι υπάρχει η πιθανότητα να ισχύουν όλες: κάθε μια, για διαφορετικό είδος αναπαραστάσεων. Για παράδειγμα, ο ορισμός 9 υποτίθεται ότι ισχύει για τις εικόνες, τα μοντέλα και, γενικότερα, τις αναλογικές αναπαραστάσεις, ενώ ο ορισμός 10 για τους ενδείκτες¹⁶⁴. Ανεξάρτητα από την ορθότητα των συγκεκριμένων προτάσεων (βλέπε τη συνέχεια αυτής της ενότητας), αυτό που μπορούμε άμεσα να παρατηρήσουμε είναι ότι δεν αφορούν όλες τις αναπαραστάσεις, αλλά μόνο εκείνο το υποσύνολο τους που έχει αναφορά χωρίς νόημα. Κατά συνέπεια, οι ορισμοί 9 και 10 δεν μπορούν να αποτελέσουν ορισμό της έννοιας της αναπαράστασης ως γένους.

Η ίδια παρατήρηση ισχύει και για τον ορισμό 8 γιατί προϋποθέτει ότι για να υπάρξει μία αναπαράσταση θα πρέπει να υπάρχουν και άλλες. Αν, λοιπόν, δεχθούμε αυτόν τον ορισμό ως ορισμό του γένους αναπαράσταση, τότε είμαστε αναγκασμένοι να απορρίψουμε την ιδέα ότι μπορεί κάτι να λειτουργήσει μεμονωμένα ως αναπαράσταση. Το ζήτημα είναι πολύ σύνθετο και σίγουρα δεν μπορεί να λυθεί με αφορισμούς. Ωστόσο, αν παραμείνουμε στην προθεωρητική χρήση της έννοιας της αναπαράστασης, ορισμένες φορές θεωρούμε κάτι αναπαράσταση ακόμα κι αν δεν ανήκει σε κάποιο αναπαραστασιακό σύστημα. Με άλλα λόγια, υπάρχουν περιπτώσεις που θεωρούμε ότι ένα πράγμα αναφέρεται σε κάτι άλλο χωρίς να σχετίζεται λειτουργικά ή δομικά με τους αναπαραστασιακούς φορείς άλλων αναπαραστάσεων. Χαρακτηριστικό παράδειγμα θεωρούνται οι ενδείκτες. Επίσης, όπως επισημάναμε στην υποσημείωση 158, ο ορισμός 8, ακόμα κι' αν δεχθούμε ότι προσδιορίζει το νόημα μιας έκφρασης με βάση τη σχέση της με τις υπόλοιπες εκφράσεις του αναπαραστασιακού συστήματος, δεν φαίνεται να επαρκεί για τον καθορισμό της σχέσης της έκφρασης με τον κόσμο. Με άλλα λόγια, ακόμα κι αν δεχθούμε ότι η σημασιολογία λειτουργικού ρόλου εξηγεί την προέλευση του νοήματος των αναπαραστάσεων, στη συνέχεια δεν βοηθά στην εξήγηση του πώς οι αναπαραστάσεις αποκτούν αναφορά.

Οι παραπάνω λόγοι συνηγορούν υπέρ της άποψης ότι ο ορισμός 8 δεν αφορά όλες τις αναπαραστάσεις, δηλαδή δεν αφορά τις αναπαραστάσεις ως γένος. Μάλιστα, οι ίδιοι οι υποστηρικτές της σημασιολογίας λειτουργικού ρόλου θεωρούν ότι η συγκεκριμένη προσέγγιση χαρακτηρίζει τις γλωσσικές αναπαραστάσεις. Ωστόσο, εκ πρώτης όψεως, δεν φαίνεται αυτονόητο ότι κάθε γλωσσική έκφραση αποκτά το νόημα της από τις συντακτικές σχέσεις που έχει με τις υπόλοιπες εκφράσεις του γλωσσικού συστήματος. Για παράδειγμα, η ίδια έκφραση μπορεί να αποτελεί άλλοτε απειλή και άλλοτε υπόσχεση ανάλογα με τον τρόπο εκφοράς και το πλαίσιο χρήσης, ακόμα κι αν έχει τις ίδιες συντακτικές σχέσεις με τις υπόλοιπες εκφράσεις του γλωσσικού

¹⁶⁴ Πρόκειται για ό,τι στην ενότητα 3.3.1.2 ονομάσαμε "ενδεικτικά σήματα".

συστήματος. Οπότε, τουλάχιστον, το πραγματολογικό νόημα δεν καθορίζεται από τις συντακτικές σχέσεις της έκφρασης με τις υπόλοιπες εκφράσεις του γλωσσικού συστήματος.

Λόγω όλων αυτών των ενστάσεων θεωρούμε ότι ο ορισμός 8 δεν ισχύει για τις αναπαραστάσεις ως γένος. Απομένουν οι ορισμοί 6 και 7. Αυτοί, εκ πρώτης όψεως, δεν φαίνεται να πάσχουν από το πρόβλημα ότι χαρακτηρίζουν κάποιο συγκεκριμένο είδος αναπαραστάσεων. Ωστόσο, υπάρχουν ορισμένοι φιλόσοφοι που προτείνουν τη διάκριση μεταξύ μη-φυσικών (ή συμβατικών) και φυσικών αναπαραστάσεων¹⁶⁵ η οποία, αν ισχύει, καθιστά τους ορισμούς 6 και 7 ορισμούς των μη-φυσικών αναπαραστάσεων. Πιο συγκεκριμένα, σύμφωνα με την παραπάνω προσέγγιση, οι μη-φυσικές αναπαραστάσεις είναι αντικείμενα ή καταστάσεις που δεν διαθέτουν τις αναπαραστασιακές τους ιδιότητες εγγενώς, αλλά τις αποκτούν κατά ένα παράγωγο τρόπο από τους χρήστες τους. Οι φυσικές αναπαραστάσεις από την άλλη μεριά, υποτίθεται ότι υφίστανται ως τέτοιες από μόνες τους στην φύση.

Υπό μια έννοια, η διάκριση μεταξύ μη-φυσικών και φυσικών αναπαραστάσεων αναδιατυπώνει τη διάκριση μεταξύ αναπαραστάσεων που αποκτούν αναφορά με το να αποκτήσουν νόημα και αναπαραστάσεων που αποκτούν άμεσα αναφορά από τον τρόπο που ο αναπαραστασιακός φορέας σχετίζεται με το αναπαριστώμενο. Στην ενότητα 3.3.1.2 υποστηρίξαμε την άποψη ότι συνθήκη δυνατότητας για να έχει μια αναπαράσταση αναφορά είναι να έχει νόημα. Με βάση αυτή τη θέση προκύπτει ότι δεν έχει νόημα να μιλάμε για φυσικές αναπαραστάσεις υπό την παραπάνω έννοια. Η αναφορική σχέση της αναπαράστασης με το αναπαριστώμενο δεν μπορεί να προκύψει ούτε από τη σχέση ομοιότητας αλλά ούτε και από τη σχέση αιτιότητας μεταξύ της αναπαράστασης και του αναπαριστώμενου.

Ας εξετάσουμε καταρχάς τη σχέση ομοιότητας. Βασιζόμενοι στον Goodman (1968) θα δείξουμε ότι η ύπαρξη ομοιότητας ανάμεσα σε δύο αντικείμενα δεν αποτελεί ούτε επαρκή ούτε αναγκαία συνθήκη προκειμένου το ένα να συνιστά απεικόνιση (δηλαδή, εικονιστική αναπαράσταση) του άλλου¹⁶⁶. Η σχέση ομοιότητας σε αντίθεση με τη σχέση απεικόνισης είναι συμμετρική και ανακλαστική. Συγκεκριμένα, αν το α ομοιάζει προς το β, τότε και το β ομοιάζει προς το α, ενώ αν το α απεικονίζει το β τότε το β δεν απεικονίζει το α. Επίσης, κάθε αντικείμενο

¹⁶⁵ Βλ. Dretske 1988, σελ. 54: "...φυσικά σημεία: συμβάντα και συνθήκες που αντλούν τις ενδεικτικές τους δυνάμεις, όχι (όπως στην περίπτωση των συμβόλων) από εμάς, από τη χρήση που τους κάνουμε να ενδεικνύουν, αλλά από τον τρόπο που αντικειμενικά συνδέονται με τις συνθήκες που σηματοδοτούν". Επίσης πρβλ. Grice 1957.

¹⁶⁶ Στο παράρτημα 1 παρουσιάζεται αναλυτικά η κριτική του Goodman καθώς και η θετική του πρόταση για το τι είναι μια εικόνα και σε τι διαφέρει από τις γλωσσικές αναπαραστάσεις.

ομοιάζει με τον εαυτό του, δεν ισχύει όμως ότι και κάθε αντικείμενο απεικονίζει τον εαυτό του. Άρα, λυιπόν, δεν είναι κάθε σχέση ομοιότητας σχέση απεικόνισης. Μήπως όμως ισχύει το αντίθετο, δηλαδή, κάθε σχέση απεικόνισης είναι σχέση ομοιότητας; Ούτε και αυτό μπορεί να υποστηριχθεί γιατί ορισμένες εικόνες αναφέρονται σε φανταστικά αντικείμενα και, κατά συνέπεια, δεν υπάρχει κάτι προς το οποίο να ομοιάζουν. Επίσης, το γεγονός ότι μια καρικατούρα μπορεί να διαφέρει εντελώς από το πρόσωπο που απεικονίζει δεν μας εμποδίζει να τη θεωρούμε απεικόνιση του προσώπου.

Ας έρθουμε τώρα στη δεύτερη σχέση με τη οποία επιχειρείται να εξηγηθεί η αναφορική σχέση της αναπαράστασης με το αναπαριστώμενο: τη σχέση αιτιότητας. Καταρχάς είναι προφανές ότι υπάρχει μια απειρία αιτιακών σχέσεων που δεν θεωρούνται σχέσεις αναπαράστασης: η κίνηση της μπάλας Β που προκαλείται αιτιακά από την κρούση της με την μπάλα Α δεν καθιστά, από μόνη της, την κίνηση της μπάλας Β (ή την ίδια την μπάλα Β) αναπαράσταση (ειδικότερα, ενδεικτική) της κίνησης της μπάλας Α (ή της ίδιας της μπάλας Α). Επίσης, αντίστροφα, το Γ μπορεί να ενδεικνύει το Δ χωρίς το Δ να αποτελεί αιτία του Γ. Ο καπνός ενδεικνύει τη φωτιά ακόμα και όταν δεν έχει παραχθεί από κάποια φωτιά. Πρόκειται για την περίπτωση της λανθασμένης ένδειξης. Με άλλα λόγια, αν η αναφορική σχέση του ενδεικτικού οφειλόταν στην αιτιακή σχέση του με το ενδεικνύμενο δεν θα υπήρχε περιθώριο λάθους. Όμως η δυνατότητα λάθους είναι αναγκαίο χαρακτηριστικό των αναπαριστάσεων¹⁶⁷. Κατά συνέπεια, η αιτιακή σχέση δεν μπορεί να εξηγήσει τη σχέση ένδειξης ή, γενικότερα, την αναφορική σχέση των αναπαριστάσεων.

Με τα παραπάνω επιχειρήματα δείξαμε ότι η ύπαρξη είτε σχέσης ομοιότητας είτε σχέσης αιτιότητας μεταξύ δύο πραγμάτων δεν αποτελεί ούτε αναγκαία ούτε επαρκή συνθήκη για να συνιστά το ένα πράγμα απεικόνιση ή ενδεικτική του άλλου αντίστοιχα. Δύο πορίσματα προκύπτουν άμεσα από αυτό το συμπέρασμα: α) το διακριτικό γνώρισμα που καθιστά ένα πράγμα απεικόνιση ή ενδεικτική δεν πρέπει να αναζητηθεί στη σχέση του πράγματος με το αναπαριστώμενο και β) δεν υπάρχουν φυσικές αναπαριστάσεις αλλά μόνο μη-φυσικές ή, ισοδύναμα, δεν υπάρχουν αναπαριστάσεις που αποκτούν την αναφορά τους άμεσα από τον τρόπο που ο αναπαραστασιακός φορέας σχετίζεται με το αναπαριστώμενο – όλες οι αναπαριστάσεις αποκτούν την αναφορική τους λειτουργία από το νόημά τους.

Οι επονομαζόμενες "φυσικές αναπαριστάσεις" παρότι δεν είναι πραγματικές αναπαριστάσεις, συχνά στην πράξη, εκλαμβάνονται ως αναπαριστάσεις. Αυτή θεωρούμε ότι αποτελεί μια από τις βασικότερες αιχμές λαθών στην χρήση της έννοιας της αναπαράστασης. Για παράδειγμα, πολύ συχνά στην επιστημονική έρευνα χρησιμοποιούνται

¹⁶⁷ Θα επανέλθουμε σε αυτό το ζήτημα αναλυτικά στο έβδομο κεφάλαιο.

μοντέλα κλίμακας για τη μελέτη ενός φαινομένου. Όμως τα μοντέλα κλίμακας δεν συνιστούν από μόνα τους αναπαραστάσεις. Η ομοιότητα του μοντέλου με το μοντελοποιούμενο δεν αρκεί για να εγκαθιδρύσει την αναφορική σχέση μεταξύ αυτών. Το μοντέλο από μόνο του θα λέμε ότι αποτελεί μια *προσομοίωση* ή, γενικότερα, μια *εγγραφή* (record)¹⁶⁸ του μοντελοποιούμενου. Το ίδιο συμβαίνει και στην περίπτωση της σχέσης αιτιότητας. Η ταινία που προκύπτει από τη μαγνητοσκόπηση ενός πλοίου δεν συνιστά αναπαράσταση του πλοίου, αλλά απλά μια εγγραφή του πλοίου. Το ίδιο συμβαίνει και με τις αναλογικές μετρήσεις: η ένδειξη ενός θερμομέτρου δεν αποτελεί από μόνη της μια αναπαράσταση της θερμοκρασίας αλλά απλά μια εγγραφή της θερμοκρασίας. Η εγγραφή από μόνη της, δηλαδή ανεξάρτητα από τον τρόπο που χρησιμοποιείται από τον χρήστη ή τον ερμηνευτή¹⁶⁹, δεν έχει αναφορική λειτουργία - δεν ενδεικνύει τη θερμοκρασία.

Ένας τρόπος να φανεί η διαφορά μεταξύ εγγραφών και αναπαραστάσεων είναι να συγκρίνουμε τη διαδικασία μετάβασης από ένα είδος εγγραφής σε ένα άλλο με τη διαδικασία μετάβασης από ένα είδος αναπαράστασης σε ένα άλλο¹⁷⁰. Θα ξεκινήσουμε με τις εγγραφές. Υπάρχουν δύο, τουλάχιστον, είδη εγγραφών: η ψηφιακή εγγραφή και η αναλογική εγγραφή. Η πρόοδος της τεχνολογίας έχει δείξει ότι μπορούμε να περνάμε από το ένα είδος εγγραφής στο άλλο με καθαρά μηχανικό τρόπο. Έτσι, μπορούμε πλέον να "εγγράψουμε" ένα τοπίο με μια φωτογραφική μηχανή και στη συνέχεια να ψηφιοποιήσουμε τη φωτογραφία με τη βοήθεια ενός σαρωτή. Με εξίσου μηχανικό τρόπο μπορούμε να κάνουμε και την αντίστροφη διαδικασία. Η μετάβαση, λοιπόν, από ένα είδος εγγραφής σε κάποιο άλλο -ελλείπει καλύτερου όρου, ας την ονομάσουμε "μετεγγραφή" - είναι μια καθαρά μηχανική διαδικασία. Αντίθετα, η μετάβαση από ένα είδος αναπαράστασης σε κάποιο άλλο είναι μια διαδικασία που δεν μπορεί να επιτελεστεί μηχανικά. Αρκεί να σκεφτούμε το πόσο δύσκολο είναι να περιγράψουμε με λόγια μια εικόνα ή αντίστροφα να φτιάξουμε από περιγραφές το σκίτσο ενός προσώπου. Εδώ δεν πρόκειται πλέον για μια μηχανική διαδικασία αλλά για μια *μετάφραση*.

Γιατί η μετάφραση σε αντίθεση με τη μετεγγραφή δεν μπορεί να εκτελεστεί μηχανικά; Η διάκριση που υιοθετήσαμε μεταξύ αναπαράστασης και εγγραφής δίνει άμεσα την απάντηση: η μετεγγραφή βασίζεται αποκλειστικά στη *μορφή* των εγγραφών ανεξάρτητα από το τι κάθε φορά εγγράφεται. Εξάλλου, μια εγγραφή από μόνη της δεν έχει

¹⁶⁸ Βλ. Haugeland 1991.

¹⁶⁹ Το εάν τελικά αυτός που αποδίδει νόημα σε μια αναπαράσταση πρέπει να κατανοηθεί ως χρήστης ή ως ερμηνευτής είναι ένα ερώτημα που θα εξεταστεί στη συνέχεια της παρούσας ενότητας.

¹⁷⁰ Βλ. Haugeland 1991.

αναπαραστασιακό περιεχόμενο, δεν αναφέρεται σε ό,τι εγγράφει. Αντίθετα, στην περίπτωση της μετάφρασης μεταβαίνουμε από ένα είδος αναπαραστασιακού περιεχομένου σε κάποιο άλλο.

Μια ένσταση σε αυτό το σημείο θα ήταν ότι υπάρχουν συγκεκριμένες περιπτώσεις που μεταβαίνουμε από ένα είδος αναπαράστασης σε κάποιο άλλο με καθαρά μηχανικό τρόπο και άρα υπάρχουν μεταφράσεις που δεν διαφοροποιούνται ως προς αυτό το χαρακτηριστικό από τις μετεγγραφές. Ας δούμε ένα παράδειγμα. Έστω ότι χωρίζουμε την ασπρόμαυρη φωτογραφία ενός λιονταριού σε πολύ μικρά τετραγωνίδια και στη συνέχεια φτιάχνουμε μια λίστα όπου καταγράφουμε για κάθε τετράγωνο το χρώμα του. Έτσι, στη θέση της φωτογραφίας έχουμε τώρα ένα σύνολο γλωσσικών περιγραφών του τύπου:

"το τετράγωνο (1,1) είναι μαύρο"

"το τετράγωνο (1,2) είναι άσπρο"

κ.λπ.

Στην προκειμένη περίπτωση έχουμε μια διαδικασία μετάβασης από μια απεικονιστική αναπαράσταση σε μια γλωσσική αναπαράσταση. Πρόκειται, λοιπόν, για μια διαδικασία μετάφρασης. Η διαδικασία αυτή, όμως, μπορεί να εκτελεστεί καθαρά μηχανικά. Άρα, ορισμένες μεταφράσεις είναι μηχανικές και, κατά συνέπεια, οι μετεγγραφές δεν διακρίνονται από τις μεταφράσεις με βάση τη μηχανικότητά τους.

Η αντίρρηση αυτή δεν ευσταθεί διότι το προαναφερθέν παράδειγμα δεν αποτελεί στην πραγματικότητα παράδειγμα μετάφρασης. Ο λόγος είναι απλός: διαβάζοντας τις περιγραφές δεν θα καταλάβουμε ότι αναφέρονται σε ένα λιοντάρι. Σε μια μετάφραση μεταφράζεται το αναπαραστασιακό περιεχόμενο. Για παράδειγμα, μια μετάφραση της φωτογραφίας του λιονταριού θα ήταν μια περιγραφή των χαρακτηριστικών του λιονταριού όπως εμφανίζονται στην φωτογραφία. Αντίθετα, η γλωσσική περιγραφή του χρώματος των τετραγωνιδίων της φωτογραφίας αποτελεί μια αναπαράσταση της αναπαραστασιακής μορφής της φωτογραφίας. Αυτή η αναπαράσταση είναι "τυφλή" απέναντι στο περιεχόμενο της φωτογραφίας. Γι' αυτό άλλωστε η συγκεκριμένη μέθοδος μετάβασης από τη φωτογραφία στη γλωσσική περιγραφή δεν επιφέρεται από μεταβολές στο περιεχόμενο της φωτογραφίας. Αυτό γίνεται φανερό αν σκεφτούμε ότι αντί της φωτογραφίας θα μπορούσε κάλλιστα να υπάρχει ένας τυχαίος σχηματισμός χρωμάτων που δεν απεικονίζει τίποτε. Από τα παραπάνω συνάγεται ότι η διαδικασία που περιγράφει το παράδειγμα δεν αποτελεί μετάφραση αλλά μια μέθοδο παραγωγής (γλωσσικών) αναπαραστάσεων των χρωμάτων κάποιας επιφάνειας.

Ωστόσο, αν δεχθούμε τελικά ότι πρόκειται για γλωσσικές αναπαραστάσεις των χρωμάτων κάποιας επιφάνειας πάλι υπονομεύεται η προτεινόμενη διάκριση μεταξύ αναπαραστάσεων και εγγραφών. Ο λόγος είναι απλός: η διαδικασία παραγωγής τέτοιου είδους αναπαραστάσεων μπορεί πλήρως να μηχανοποιηθεί. Πιο συγκεκριμένα, αρκεί ένα μηχάνημα που έχει τη δυνατότητα ανίχνευσης του χρώματος κάθε τετραγωνιδίου. Ένα τέτοιο μηχάνημα θα μπορούσε αυτόματα να παραγάγει γλωσσικές

προτάσεις του τύπου "το τετράγωνο (1,1) είναι μαύρο" από εικονιστικές αναπαραστάσεις.

Η αντίρρηση αυτή αναδεικνύει εκ νέου την αναγκαιότητα προσφυγής στον χρήστη προκειμένου να διαπιστωθεί αν ένα πράγμα λειτουργεί ως αναπαράσταση. Όπως ήδη διαπιστώσαμε, το ίδιο το μηχάνημα που μετατρέπει την φωτογραφία σε ένα σύνολο περιγραφών βασίζεται αποκλειστικά στη μορφή της φωτογραφίας. Κατά τον ίδιο τρόπο, αν επρόκειτο να εκτελέσει την αντίστροφη διαδικασία, πάλι θα βασιζόταν αποκλειστικά στη μορφή των γλωσσικών αναπαραστάσεων. Με άλλα λόγια, το μηχάνημα δεν μπορεί να χρησιμοποιήσει τη φωτογραφία και τις περιγραφές ως αναπαραστάσεις γιατί δεν έχει πρόσβαση στο αναπαραστασιακό τους περιεχόμενο. Γι' αυτό και δεν έχει καμία σημασία τελικά για την ορθή λειτουργία του μηχανήματος το εάν η φωτογραφία απεικονίζει όντως κάτι ή αποτελεί μια τυχαία σύνθεση χρωμάτων, όπως επίσης και το εάν τα εξαγόμενα σχήματα αποτελούν όντως ελληνικές προτάσεις ή είναι κάποια ακατανόητα ορνιθοσκαλισματα. Τόσο η φωτογραφία όσο και οι περιγραφές, από την άποψη του μηχανήματος, αποτελούν διαφορετικά είδη εγγραφών. Από την άποψη ενός επαρκούς χρήστη αναπαραστάσεων, ωστόσο, η μεν φωτογραφία αποτελεί αναπαράσταση ενός lionταριού, οι δε γλωσσικές περιγραφές αποτελούν αναπαραστάσεις της αναπαραστασιακής μορφής της φωτογραφίας¹⁷¹.

Ανακεφαλαιώνοντας, θα λέγαμε ότι η έννοια των φυσικών αναπαραστάσεων είναι παραπλανητική διότι οι επονομαζόμενες "φυσικές αναπαραστάσεις" δεν διαθέτουν από μόνες τους αναπαραστασιακό περιεχόμενο. Υπ' αυτή την έννοια δεν πρέπει να θεωρούνται ένα ιδιαίτερο είδος αναπαραστάσεων—είναι απλώς εγγραφές¹⁷².

¹⁷¹ Η διαφορά της μετάφρασης από τη μετεγγραφή αναδεικνύεται επίσης αν επιχειρήσουμε να αντιστρέψουμε αυτές τις δύο διαδικασίες. Ενώ από μια ψηφιακή εγγραφή μπορούμε με πολύ μεγάλη ακρίβεια να ξαναπάρουμε την αρχική αναλογική εγγραφή, δεν συμβαίνει το ίδιο όταν μεταφράσουμε το σύνολο των γλωσσικών περιγραφών σε εικόνα. Τίποτε δεν εγγυάται ότι η νέα εικόνα θα είναι ίδια με την αρχική.

¹⁷² Στην καθημερινή χρήση της γλώσσας υπάρχουν περιπτώσεις που μεταχειριζόμαστε τη λέξη "αναπαράσταση" για να αναφερθούμε σε ό,τι εδώ ονομάσαμε "εγγραφές". Φυσικά, μια τέτοια χρήση δεν είναι δυνατόν να δειχθεί εκ των προτέρων ότι είναι λανθασμένη και ούτε κάτι τέτοιο αποτελεί στόχο του παρόντος κεφαλαίου. Αυτό που μας ενδιαφέρει είναι να καταδείξουμε ότι η χρήση του όρου "αναπαράσταση" υπό την έννοια της εγγραφής, δεν προϋποθέτει την αναφορική λειτουργία. Κατά συνέπεια, όταν στο πλαίσιο της φιλοσοφίας και, ειδικότερα, της φιλοσοφίας του νου, η λέξη "αναπαράσταση" χρησιμοποιείται ως τεχνικός όρος με στόχο να εξηγηθεί η αποβλεπτικότητα των νοητικών φαινομένων με βάση την αναφορικότητα των εσωτερικών αναπαραστάσεων, τότε η συγκεκριμένη εξήγηση δεν είναι επιτυχής αν η αναπαράσταση κατανοηθεί ως εγγραφή.

Εφόσον, λοιπόν, υπάρχουν μόνο μη-φυσικές αναπαραστάσεις, οι ορισμοί 6 και 7 δεν αποτελούν, ως προς αυτήν την άποψη, ορισμούς κάποιου συγκεκριμένου είδους αναπαραστάσεων. Δεχόμενοι, μέχρι αποδείξεως του εναντίου, ότι δεν υπάρχει κάποιο άλλο είδος αναπαράστασης για το οποίο δεν ισχύουν οι ορισμοί 6 και 7, μπορούμε να συμπεράνουμε ότι οι εν λόγω ορισμοί αφορούν την έννοια της αναπαράστασης ως γένος. Όμως σε αυτή την περίπτωση μόνο ένας από τους δύο ορισμούς πρέπει να ισχύει. Το ζητούμενο, εν προκειμένω, είναι να εξετάσουμε αν οι αναπαραστάσεις αποκτούν το νόημά τους από έναν ερμηνευτή ή από έναν χρήστη. Προκειμένου να ξεκινήσει αυτή η έρευνα πρέπει να διευκρινιστεί σε τι διαφέρει ένας χρήστης από έναν ερμηνευτή. Το συγκεκριμένο ερώτημα δεν είναι καθόλου εύκολο να απαντηθεί διότι προϋποθέτει μια ολοκληρωμένη θεωρία για το τι καθιστά ένα ον νοήμον. Γι' αυτό θα προβούμε απλώς σε ορισμένες γενικές επισημάνσεις.

Καταρχάς, ένας ερμηνευτής, προτού αποδώσει αναπαραστασιακό περιεχόμενο σε ένα εξωτερικό αντικείμενο, "κατέχει" από μόνος του αυτό το περιεχόμενο υπό τη μορφή νοητικού περιεχομένου. Με τη διαδικασία της ερμηνείας το αντικείμενο αποκτά το αναπαραστασιακό περιεχόμενο που προϋπάρχει στο νου του ερμηνευτή. Αντίθετα, στην περίπτωση του χρήστη, το αναπαραστασιακό περιεχόμενο δεν προϋπάρχει υπό μια άλλη μορφή αλλά εγκαθιδρύεται ακριβώς από την αλληλεπίδραση του χρήστη με το αντικείμενο. Με γενικότερους όρους, στην περίπτωση του ερμηνευτή υπάρχει μια προτεραιότητα του νοητικού περιεχομένου έναντι του περιεχομένου των εξωτερικών αναπαραστάσεων, κάτι που δεν ισχύει στην περίπτωση του χρήστη. Η προτεραιότητα αυτή δεν πρέπει να κατανοηθεί χρονικά αλλά υπερβατολογικά: η κατοχή νοητικού περιεχομένου εκ μέρους του ερμηνευτή αποτελεί συνθηκική δυνατότητα για να αποκτήσει ένα εξωτερικό αντικείμενο αναπαραστασιακό περιεχόμενο. Μιλώντας εντελώς σχηματικά, ο ορισμός 7 προϋποθέτει την υπερβατολογική προτεραιότητα της σκέψης έναντι της γλώσσας¹⁷³.

Ωστόσο, αν ισχύει η παραδοχή ότι η σκέψη (το νοητικό περιεχόμενο) προηγείται υπερβατολογικά της γλώσσας και, γενικότερα, του αναπαραστασιακού περιεχομένου των εξωτερικών αναπαραστάσεων, τότε δημιουργείται το ερώτημα με ποιο τρόπο μπορεί ο ερμηνευτής να

¹⁷³ Βλ. Searle 1983, σελ. vii: "Εφόσον, οι προτάσεις – οι ήχοι που βγαίνουν από το στόμα κάποιου ή τα σημάδια που φτιάχνει κανείς στο χαρτί – είναι, θεωρημένες υπό μια έννοια, απλώς αντικείμενα του κόσμου όπως κάθε άλλο αντικείμενο, η ικανότητά τους να αναπαριστούν δεν είναι εγγενής αλλά παράγωγη από την αποβλεπτικότητα του νου. Η αποβλεπτικότητα των νοητικών καταστάσεων, από την άλλη μεριά, δεν είναι παράγωγη από κάποιες προγενέστερες μορφές αποβλεπτικότητας αλλά είναι εγγενής στις ίδιες τις [νοητικές] καταστάσεις". Ένα πρόσθετο επιχείρημα για το τι διαφοροποιεί τον (επιδέξιο) χρήστη από τον ερμηνευτή αναπαραστάσεων (και, γενικότερα, εργαλείων) θα δοθεί στο τέταρτο κεφάλαιο (ενότητα 4.3.2).

αποκτήσει νοητικά περιεχόμενα χωρίς τη βοήθεια εξωτερικών αναπαραστάσεων; Το ερώτημα αυτό έχει επιχειρηθεί να απαντηθεί κυρίως με δύο τρόπους. Ο πρώτος τρόπος προαναφέρθηκε ήδη στο δεύτερο κεφάλαιο. Πρόκειται για την προσέγγιση της κλασσικής γνωσιακής επιστήμης: η σκέψη δεν είναι τίποτε άλλο από τον χειρισμό νοητικών αναπαραστάσεων του λειτουργικού επιπέδου. Ο δεύτερος τρόπος είναι η νοησαρχική-καρτεσιανή προσέγγιση, σύμφωνα με την οποία σκέψη είναι η κατοχή και η χρήση (υποκειμενικών) παραστάσεων (Vorstellungen).

Αν ακολουθήσει κανείς την πρώτη προσέγγιση, ο ερμηνευτής αποσυντίθεται σε ένα σύνολο υποπροσωπικών γνωσιακών υπομηχανισμών, ενώ αν ακολουθήσει τη δεύτερη προσέγγιση ο ερμηνευτής κατανοείται ως ένα καρτεσιανό υποκείμενο. Στα αμέσως επόμενα κεφάλαια θα ασκήσουμε κριτική στην πρώτη προσέγγιση και στο όγδοο κεφάλαιο θα παρουσιάσουμε επιχειρήματα κατά της δεύτερης προσέγγισης. Αν οι κριτικές αυτές αποδειχθούν βάσιμες, τότε ο ορισμός 6 παραμένει η μόνη εύλογη πρόταση για το τι είναι εξωτερική αναπαράσταση εν γένει.

Ωστόσο, στην παρούσα φάση, δηλαδή πριν γίνει η κριτική στις δύο προαναφερθείσες προσεγγίσεις, δεν δικαιούμαστε να εκλάβουμε τον ορισμό 6 ως ορισμό του γένους αναπαράσταση. Γι' αυτό θα παραμείνουμε στον αμέσως προηγούμενο ορισμό, δηλαδή τον ορισμό 5:

Ορισμός 5:	Ένα αντικείμενο Α συνιστά αναπαράσταση κάποιου αντικειμένου, ιδιότητας, σχέσης ή κατάστασης πραγμάτων Ο, δηλαδή έχει ως περιεχόμενο το Ο, αν για κάποιον επαρκή χρήστη Χ η μορφή Μ του Α μπορεί να αναφέρεται (ακριβέστερα, να παραπέμπει) στο Ο (αναφερόμενο) υπό μια συγκεκριμένη προοπτική Ν (νόημα).
------------	--

3.4 *Επιλογές*

Η εξέταση της έννοιας των εξωτερικών αναπαραστάσεων μάς παρείχε ένα πλούτο διακρίσεων και διαπιστώσεων που θα χρησιμοποιήσουμε στην εξέταση της έννοιας των νοητικών αναπαραστάσεων. Το βασικό συμπέρασμα ήταν ότι αυτό που χαρακτηρίζει τις εξωτερικές αναπαραστάσεις ως γένος είναι η αναφορική τους λειτουργία. Επίσης, δείξαμε ότι συνθήκη δυνατότητας για να έχει ένα πράγμα αναφορά είναι να έχει νόημα. Υπ' αυτό το πρίσμα, η αναπαραστασιακή σχέση μπορεί να αναλυθεί σε δύο αλληλένδετες σχέσεις: αφενός την αναφορική σχέση της αναπαράστασης με το αναπαριστώμενο και αφετέρου τη νοηματική σχέση του χρήστη με την αναπαράσταση. Όπως θα δούμε στα επόμενα κεφάλαια κατά την εξέταση της έννοιας της νοητικής αναπαράστασης, η γνωσιακή επιστήμη βασίζεται σχεδόν αποκλειστικά στην αναφορική σχέση των αναπαραστάσεων και, υπ' αυτή την έννοια, κατανοεί τις νοητικές αναπαραστάσεις ως διμελείς

σχέσεις μεταξύ του αναπαραστασιακού φορέα και του αναπαριστώμενου. Στην προηγούμενη ενότητα δείξαμε ότι αν κανείς επιχειρήσει να εξηγήσει τις αναπαραστάσεις κατ' αυτόν τον τρόπο, δεν κατορθώνει να συλλάβει την έννοια της αναπαράστασης αλλά αυτό που αποκαλέσαμε "εγγραφή". Στα επόμενα κεφάλαια δεν θα προϋποθέσουμε αυτή την κριτική, αλλά μάλλον θα την επαναλάβουμε πολύ πιο συστηματικά και λεπτομερειακά. Επίσης, θα αναπτύξουμε νέα επιχειρήματα που υπονομεύουν την έννοια της νοητικής αναπαράστασης από διαφορετικές σκοπιές.

Στο δεύτερο κεφάλαιο, διατυπώνοντας την ΑΥΘΝ, εξετάσαμε, σε μια πρώτη προσέγγιση, τι σημαίνει για μια νοητική αναπαράσταση να είναι γλωσσική-συμβολική. Στο επόμενο κεφάλαιο θα μελετήσουμε λεπτομερέστερα αυτήν την ιδέα μέσα από το πρόγραμμα αναπαράστασης της γνώσης στην κλασική Τεχνητή Νοημοσύνη και θα δείξουμε ότι υπάρχουν είδη γνώσης που δεν μπορούν να εκφραστούν με συμβολικές αναπαραστάσεις. Ακολούθως, στα κεφάλαια 5 και 6 θα εξετάσουμε αντίστοιχα δύο διαφορετικά είδη νοητικών αναπαραστάσεων που προτείνονται εναλλακτικά από υποστηρικτές της αναπαραστασιακής θεωρίας του νου. Πρόκειται για τις εικονιστικές και τις κατανεμημένες αναπαραστάσεις αντίστοιχα. Θα δείξουμε ότι και οι δύο αυτές έννοιες παρουσιάζουν ανυπέρβλητα προβλήματα. Τέλος, στο έβδομο κεφάλαιο θα ασχοληθούμε με το πρόβλημα της προέλευσης του περιεχομένου των νοητικών αναπαραστάσεων, δηλαδή θα εξετάσουμε το ερώτημα δυνάμει τίνος πράγματος εσωτερικές καταστάσεις ενός οργανισμού αποκτούν αναπαραστασιακό περιεχόμενο; Θα δείξουμε ότι καμία από τις προτεινόμενες λύσεις του προβλήματος δεν ευσταθεί.

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ

**Το αδιέξοδο του παραδοσιακού προγράμματος αναπαράστασης της
γνώσης και η αναγκαιότητα ύπαρξης μη προτασιακής γνώσης**

4.1 Εισαγωγή

Στο κεφάλαιο αυτό θα εξετάσουμε τα αδιέξοδα στα οποία οδηγήθηκε το πρόγραμμα αναπαράστασης της γνώσης στην Τέχνητή Νοημοσύνη (ΤΝ)¹⁷⁴ -- έναν κλάδο που διαπνέεται από την Αναπαραστασιακή - Υπολογιστική Θεωρία του Νου (ΑΥΘΝ) και αποτελεί μαζί με την γνωσιακή ψυχολογία τη ραχοκοκαλιά της γνωσιακής επιστήμης. Θα ξεκινήσουμε με την παρουσίαση των βασικότερων τεχνικών που αναπτύχθηκαν για την αναπαράσταση της γνώσης. Όπως θα δούμε, η ανθρώπινη γνώση κατανοήθηκε στο πλαίσιο της ΤΝ ως αποκλειστικά προτασιακή γνώση και γι' αυτό θεωρήθηκε ότι μπορεί να εκφραστεί με συμβολικές αναπαραστάσεις. Ωστόσο, αυτό το εγχείρημα δεν ευδοκίμησε. Ειδικότερα, η κοινότοπη γνώση και η περιστασιακή γνώση - μορφές γνώσης που, όπως θα δούμε, εμπλέκονται σχεδόν σε κάθε γνωσιακή δραστηριότητα - αποδείχθηκε ότι δεν αποτελούν είδος προτασιακής γνώσης και ότι κατά συνέπεια δεν μπορούν να εκφραστούν με συμβολικές αναπαραστάσεις. Η αποτυχία του προγράμματος (συμβολικής) αναπαράστασης της γνώσης αναδεικνύει καλύτερα ίσως από κάθε άλλο πρόγραμμα της κλασικής γνωσιακής επιστήμης α) την αναγκαιότητα ύπαρξης μη προτασιακής γνώσης και β) την εξάρτηση της προτασιακής γνώσης από τη μη προτασιακή.

Όπως είδαμε στο δεύτερο κεφάλαιο, μια βασική θέση στην οποία στηρίζεται η ΑΥΘΝ και, κατ' επέκταση, η γνωσιακή επιστήμη είναι ότι οι νοητικές αναπαραστάσεις που υπόκεινται των γνωσιακών φαινομένων είναι συμβολικές. Σύμφωνα με τον Fodor, μόνο οι συμβολικές αναπαραστάσεις μπορούν να εξηγήσουν την παραγωγικότητα και τη συστηματικότητα της σκέψης. Επίσης, μόνο σε τέτοιου τύπου αναπαραστάσεις μπορούν να στηριχθούν οι υπολογιστικές λειτουργίες που, σύμφωνα με την ΑΥΘΝ, υλοποιούν τις αποβλεπτικές λειτουργίες. Με άλλα λόγια, η έννοια του υπολογισμού, κατά τον Fodor, μπορεί να οριστεί μόνο με τη βοήθεια συμβολικών αναπαραστάσεων.

Ωστόσο, κατά την παρουσίαση της ΑΥΘΝ στο δεύτερο κεφάλαιο, κάθε αναφορά στις συμβολικές αναπαραστάσεις στηρίχτηκε σε μια προθεωρητική κατανόηση του όρου. Ο μόνος τρόπος με τον οποίο προσδιορίστηκαν οι συμβολικές αναπαραστάσεις ήταν μέσω παραδειγμάτων. Χαρακτηριστικότερο παράδειγμα, όπως είδαμε, θεωρήθηκαν οι γλωσσικές εκφράσεις, ενώ παράδειγμα μη-συμβολικών αναπαραστάσεων θεωρήθηκαν οι εικόνες. Ο λόγος που ακολουθήσαμε μια τέτοιου τύπου προσέγγιση στον προσδιορισμό των συμβολικών νοητικών αναπαραστάσεων είναι ότι η ίδια η θεωρία του Fodor έχει στηριχθεί σε μια προθεωρητική κατανόηση της έννοιας. Αυτή η στάση είναι ιδιαίτερα χρήσιμη για να καταφέρει κανείς να ολοκληρώσει μια θεωρία τόσο γενική

¹⁷⁴ Βλ. κεφάλαιο 1, ενότητα 1.3.8.

όσο η ΑΥΘΝ. Αν ο Fodor επιχειρούσε πρώτα να διασαφηνίσει και να περιχαρακώσει θεωρητικά την έννοια της συμβολικής αναπαράστασης, είναι πιθανό να μην ολοκλήρωνε ποτέ την ΑΥΘΝ. Όπως φάνηκε, εξάλλου, και από το τρίτο κεφάλαιο, κάθε απόπειρα διατύπωσης μίας γενικής θεωρίας περί αναπαράστασης συναντά ανυπέρβλητες δυσκολίες¹⁷⁵. Γι' αυτό σε εκείνο το κεφάλαιο περιοριστήκαμε σε ένα πρώτο προσδιορισμό της έννοιας της αναπαράστασης ως γένους και όχι στο πώς διαφοροποιούνται μεταξύ τους τα διάφορα είδη αναπαράστασεων. Ωστόσο, προκειμένου πλέον να ελέγξουμε την ορθότητα της ΑΥΘΝ θα πρέπει να συγκεκριμενοποιήσουμε τι είναι οι συμβολικές αναπαράστασεις και πώς διαφοροποιούνται από τα άλλα είδη αναπαράστασεων. Μόνο μέσα από μια τέτοια πορεία θα διαφανεί εάν η βασική έννοια γύρω από την οποία χτίζεται η ΑΥΘΝ ευσταθεί και επαρκεί για να εξηγήσει όλο το εύρος των γνωσιακών φαινομένων.

4.2 Το πρόγραμμα αναπαράστασης της γνώσης στην κλασική τεχνητή νοημοσύνη

Ο κλάδος της γνωσιακής επιστήμης που έχει συμβάλλει ίσως περισσότερο από κάθε άλλο στην διερεύνηση της έννοιας της συμβολικής αναπαράστασης είναι η Τεχνητή Νοημοσύνη. Ειδικότερα η κλασική ΤΝ¹⁷⁶ αποτελεί μια έκφραση της γνωσιακής επιστήμης διότι ουσιαστικά βασίζεται στην ΑΥΘΝ. Πιο συγκεκριμένα, η βασική παραδοχή στην οποία στηρίζεται η κλασική ΤΝ είναι ότι η σκέψη, και γενικότερα οι νοητικές διαδικασίες, δεν είναι τίποτα άλλο από υπολογιστικός χειρισμός συμβολικών αναπαράστασεων. Σύμφωνα, λοιπόν, με αυτή την προσέγγιση, οι νοητικές διαδικασίες εξισώνονται με τον χειρισμό συμβολικών αναπαράστασεων μέσω τυπικών κανόνων, δηλαδή, κανόνων που αφορούν την *μορφή* των αναπαράστασεων. Στην κλασική ΤΝ οι συμβολικές αναπαράστασεις χρησιμοποιούνται για να κωδικοποιήσουν την ανθρώπινη γνώση, κατ' αναλογία προς τον τρόπο που η *φυσική* γλώσσα κωδικοποιεί τη γνώση.

Το ενδιαφέρον μας για την ΤΝ θα επικεντρωθεί αποκλειστικά στην χρήση αυτής ως προτύπου για την εξήγηση της *ανθρώπινης* νοημοσύνης. Θα ασχοληθούμε ειδικότερα με την αναπαράσταση της γνώσης στον υπολογιστή η οποία αποτελεί ένα βασικό πεδίο έρευνας της ΤΝ και έχει σαν στόχο την ανάπτυξη νέων και αποτελεσματικότερων

¹⁷⁵ Απόδειξη τούτου αποτελεί ο εξαιρετικά περιορισμένος αριθμός φιλοσόφων που έχουν αναλάβει ένα τέτοιο εγχείρημα. Κυριότεροι, εξ αυτών, είναι οι: Peirce, Cassirer, Morris, Langer και Goodman: Επίσης γενικές θεωρίες περί αναπαράστασεων ή σημείων μπορεί κανείς να συναντήσει στο χώρο της γλωσσολογίας και μιας παράγωγης από αυτήν επιστήμης, της σημειολογίας. Σημαντικοί θεωρητικοί σε αυτόν τον χώρο είναι οι: Saussure, Benveniste και Eco.

¹⁷⁶ Βλ. Haugeland 1985.

τεχνικών και φορμαλισμών για την αποθήκευση και διαχείριση της γνώσης. Θα ξεκινήσουμε λοιπόν με την παρουσίαση των βασικότερων τεχνικών που αναπτύχθηκαν για την αναπαράσταση της γνώσης. Μέσα από αυτήν την εξέταση θα διαφανεί ότι στο πλαίσιο της κλασικής ΤΝ η ανθρώπινη γνώση κατανοήθηκε ως αποκλειστικά προτασιακή γνώση και γι' αυτό θεωρήθηκε ότι μπορεί να εκφραστεί με συμβολικές αναπαράστασεις.

4.2.1 Από τη συντακτική επεξεργασία της φυσικής γλώσσας στους μικρόκοσμους

Το αίτημα για την αναπαράσταση της γνώσης στον υπολογιστή αναδύθηκε κυρίως μέσα από την έρευνα για την μηχανική επεξεργασία της φυσικής γλώσσας. Τα πρώτα συστήματα επεξεργασίας της φυσικής γλώσσας στηρίζονταν στην αναγνώριση λέξεων ή φράσεων χωρίς συντακτική ή σημασιολογική ανάλυση. Για παράδειγμα, το πρόγραμμα ELIZA¹⁷⁷ μπορούσε μεταξύ άλλων να απαντά σε προτάσεις της μορφής "Εγώ εσένα" με ερωτήσεις της μορφής "Γιατί εσύ εμένα;". Αυτό το πετύχαινε απλά με τον εντοπισμό των λέξεων κλειδιών στα όρια μιας πρότασης.

Η επόμενη εξέλιξη στην επεξεργασία της φυσικής γλώσσας ήταν η χρήση αυτόματων συντακτικών αναλυτών. Αφετηρία αποτέλεσε η Μετασημματιστική Γραμματική του Chomsky (1957). Στο έργο αυτό γινόταν διάκριση μεταξύ της επιφανειακής και της βαθιάς δομής της πρότασης. Η πρώτη είναι η ακολουθία των συγκεκριμένων λέξεων που απαρτίζουν την πρόταση, ενώ η δεύτερη περιγράφεται με βάση ιεραρχικές γραμματικές κατηγορίες όπως ονοματική φράση, ρηματική φράση, ουσιαστικό κ.λ.π. Τα κύρια στοιχεία της θεωρίας του Chomsky που ενίσχυσαν την υπολογιστική μεθοδολογία στην μελέτη της γλώσσας είναι:

1. Η γλώσσα θεωρείται ένα τυπικά καθορισμένο γενετικό σύστημα
2. Χρησιμοποιείται η πολύ ισχυρή μαθηματική έννοια της αναδρομής
3. Η σύνταξη θεωρείται αυτόνομη από τη σημασιολογία¹⁷⁸

Από τα τέλη του '60 έγινε φανερό ότι για τη μηχανική επεξεργασία της γλώσσας -είτε πρόκειται για μετάφραση, απάντηση ερωτήσεων ή ανάκτηση πληροφοριών- η συντακτική ανάλυση δεν επαρκούσε. Έπρεπε να δημιουργηθούν προγράμματα που να βασίζονται στο τι σημαίνουν οι λέξεις και οι προτάσεις. Τα πρώτα προγράμματα που

¹⁷⁷ Weizenbaum 1966.

¹⁷⁸ Με βάση τις ιδέες του Chomsky δημιουργήθηκαν τα δίκτυα αναδρομικής μετάβασης (recursive transition networks) και αργότερα τα δίκτυα επανυξημένης μετάβασης (augmented transition networks).

δημιουργήθηκαν με αυτό το σκεπτικό αφορούσαν μία πολύ μικρή και σαφώς οριοθετημένη περιοχή γνώσης (οικογενειακά δένδρα, βαθμολογίες στο baseball κλπ).

Το πιο γνωστό από τα προγράμματα που δημιουργήθηκαν με αυτό τον τρόπο είναι το SHRDLU του Winograd (1972), ένα πρόγραμμα κατανόησης της φυσικής γλώσσας το οποίο μπορούσε να απαντά σε ερωτήσεις, να εκτελεί διαταγές και να αποθηκεύει νέα γεγονότα σχετικά με έναν κόσμο στερεών σχημάτων. Για την ακρίβεια, επρόκειτο για στερεά σχήματα που εμφανίζονταν στην οθόνη ενός υπολογιστή και τα οποία μετακινούνταν με τη βοήθεια ενός προσομοιωμένου μηχανικού βραχίονα. Το σημαντικό με το SHRDLU ήταν ότι επικοινωνούσε με τον χρήστη του σε φυσική γλώσσα. Το SHRDLU ήταν ένα ολοκληρωμένο σύστημα που χρησιμοποιούσε σύνταξη, σημασιολογία και πραγματολογική γνώση γύρω από τα στερεά σχήματα. Επίσης, μια βασική καινοτομία ήταν ότι η γνώση δεν κωδικοποιούνταν δηλωτικά (declaratively), αλλά διαδικαστικά (procedurally). Αυτό σημαίνει ότι τα λεξιλογικά και συντακτικά στοιχεία της γλώσσας ορίστηκαν ως προγράμματα και όχι ως δεδομένα. Έτσι, ουσιαστικά δεν υπήρχε αναπαράσταση αυτών των στοιχείων ανεξάρτητη από τον τρόπο που το σύστημα τα επεξεργαζόταν. Οι πληροφορίες ελέγχου και χρήσης της γνώσης περιλαμβάνονταν μέσα στην ίδια την αναπαράσταση της γνώσης.

Οι διάλογοι που διεξήγαγε το SHRDLU με τους χρήστες είχαν μια εξαιρετική φυσικότητα. Ωστόσο, αυτή η φυσικότητα εξαφανιζόταν κάθε φορά που οι εντολές του χρήστη παρέκκλιναν από αυτές που ορίζονταν μέσα στο κόσμο των στερεών. Βέβαια, οι ερευνητές δεν έκρυβαν το γεγονός ότι η επιτυχία του SHRDLU οφειλόταν κυρίως στο ότι αφορούσε ένα πολύ περιορισμένο τμήμα του κόσμου - έναν "μικρόκοσμο" όπως χαρακτηριστικά τον ονόμαζαν¹⁷⁹. Αυτό, ωστόσο, δεν φαινόταν να τους ανησυχεί γιατί προϋπέθεταν ότι ο πραγματικός κόσμος αποτελείται προσεγγιστικά από ένα άθροισμα τέτοιων μικρόκοσμων. Σε ποιο βαθμό όμως οι μικρόκοσμοι αποτελούν τμήμα του πραγματικού κόσμου;

Σύμφωνα με τον Haugeland (1985, σελ. 185), "[E]νας μικρόκοσμος είναι ένα επινοημένο τεχνητό πεδίο στο οποίο όλα τα δυνατά αντικείμενα, ιδιότητες και συμβάντα έχουν οριστεί εκ των προτέρων με ρητό και περιορισμένο τρόπο". Υπό μια έννοια, οι μικρόκοσμοι θα μπορούσαν να θεωρηθούν ως κάτι αντίστοιχο με τα μοντέλα που χρησιμοποιούνται στη φυσική για την προσομοίωση φυσικών

¹⁷⁹ "Κάθε μοντέλο ή 'μικρόκοσμος', όπως θα τον ονομάζουμε, είναι πολύ σχηματικός: μιλά για ένα παραμυθένιο τόπο στον οποίο τα πράγματα είναι τόσο απλοποιημένα ώστε σχεδόν κάθε δήλωση περί αυτών θα ήταν κυριολεκτικά ψευδής εάν αποδιδόταν στον πραγματικό κόσμο." Minsky, M. & Papert, S.: "Προσχέδιο για πρόταση προς την ARPA για την έρευνα στην τεχνητή νοημοσύνη στο MIT" 1970-1971, σελ. 39 (αναφέρεται στο Dreyfus 1992, σελ. 9).

καταστάσεων. Όπως, δηλαδή, ένα μαθηματικό εκκρεμές βοηθά στη μελέτη της κίνησης του πραγματικού εκκρεμούς, έτσι και ο μικρόκοσμος προσομοιώνει ένα τμήμα του πραγματικού κόσμου¹⁸⁰. Αυτή η αναλογία, όμως, είναι παραπλανητική γιατί το ζητούμενο στην τεχνητή νοημοσύνη δεν είναι η προσομοίωση του κόσμου, αλλά η προσομοίωση της κατανόησης του κόσμου. Ουσιαστικά, λοιπόν, αυτό που πρέπει να εξεταστεί είναι εάν η κατανόηση ενός τμήματος του κόσμου (ή καλύτερα μιας γνωστικής περιοχής) μπορεί να αποκοπεί από την κατανόηση του κόσμου συνολικά. Αυτό, σε πρώτη προσέγγιση, θα σήμαινε ότι το συγκεκριμένο τμήμα του κόσμου θα μπορούσε να εκφραστεί ως ένα σύνολο γεγονότων και σχέσεων που δεν εξαρτώνται από γεγονότα και σχέσεις άλλων τμημάτων του κόσμου.

Είναι, όμως, κάτι τέτοιο δυνατό, ειδικότερα μάλιστα στην περίπτωση που οι μικρόκοσμοι αναφέρονται σε κοινωνικές δραστηριότητες, όπως για παράδειγμα, ο κόσμος των συναλλαγών ή της πολιτικής; Επίσης, ένα βαθύτερο ερώτημα που πρέπει να τεθεί είναι εάν η κατανόηση ενός τμήματος του κόσμου ισοδυναμεί με την κατοχή, αφενός, ενός συνόλου αναπαραστάσεων γεγονότων και, αφετέρου, ενός συνόλου κανόνων για τον χειρισμό των αναπαραστάσεων. Το δεύτερο αυτό ερώτημα προϋποτίθεται του πρώτου και αφορά, ουσιαστικά, την επάρκεια της χρησιμοποιούμενης οντολογίας¹⁸¹.

Ο P. Winston και οι υπόλοιποι ερευνητές του MIT, μετά από πολλές προσπάθειες να συνθέσουν μικρόκοσμους με στόχο την προσομοίωση ευρύτερων κόσμων, ομολογούν σε μια εσωτερική αναφορά τα τεράστια προβλήματα που δημιουργεί αυτή η στρατηγική:

"Η ΤΝ τα πήγε καλά σε αυστηρά περιορισμένες επικράτειες - ο Winograd για παράδειγμα, εντυπωσίασε τους πάντες με την εμπειρογνωμοσύνη (expertise) του συστήματος φυσικής γλώσσας του κόσμου των στερεών. Η επέκταση μιας τέτοιου είδους ικανότητας σε μεγαλύτερους κόσμους δεν αποδείχθηκε ωστόσο άμεση.... Ήρθε η ώρα να αντιμετωπίσουμε τα εμπλεκόμενα προβλήματα ως κεντρικά ζητήματα"¹⁸²

Η αποτυχία των μικρόκοσμων ανέδειξε ως κεντρικό ζήτημα για την ΤΝ το πρόβλημα της αναπαράστασης της πραγματολογικής γνώσης.

¹⁸⁰ Ο ίδιος ο Winograd περιγράφοντας την εργασία του μεταχειρίζεται όρους της φυσικής, θέλοντας έτσι να επισημάνει την αναλογία: "Ενδιαφερόμαστε να αναπτύξουμε έναν φορμαλισμό ή μια 'αναπαράσταση', με την οποία να περιγράψουμε τη γνώση. Αναζητούμε τα 'άτομα' και τα 'σωματίδια' από τα οποία είναι χτισμένος, καθώς και τις 'δυνάμεις' που ενεργούν επ' αυτού", Winograd (1976), σελ. 9.

¹⁸¹ Το ζήτημα αυτό θα εξετασθεί στην ενότητα 4.3.

¹⁸² Winston, P. and the Staff of the MIT, AI Laboratory, AI Memo No 366 (May 1976), σελ. 22 (αναφέρεται στο Dreyfus 1992, σελ.14).

Πιο συγκεκριμένα, οι ερευνητές συνειδητοποίησαν ότι η στρατηγική των μικρόκοσμων απέτυχε ακριβώς επειδή με αυτή, στην ουσία, μάλλον απέφρευαν παρά αντιμετώπιζαν την εμπλεκόμενη πραγματολογική γνώση.

Ειδικότερα, όσον αφορά το ζήτημα της επεξεργασίας φυσικής γλώσσας, υπάρχουν πολλά επιμέρους προβλήματα που ανακύπτουν κυρίως από τη μη καταγραφή της πραγματολογικής γνώσης. Θα αναφέρουμε ενδεικτικά ορισμένα από αυτά:

α) πολυσημία

- *λεκτική*: ι) ομωνυμία πχ. "όργανο" (της τάξης ή μουσικό), ιι) διαφορετική γραμματική κατηγορία πχ. "ήσυχα" = επίρρημα ή επίθετο
- *συντακτική* πχ. "το κορίτσι δέχτηκε τις προτάσεις με ενδιαφέρον" (δέχτηκε εκείνες τις προτάσεις που είχαν ενδιαφέρον ή δέχτηκε με ενδιαφέρον όλες τις προτάσεις;)
- *σημσιολογική* πχ. "Η ιστορία του Wittgenstein ήταν συγκινητική" (Ο Wittgenstein είναι ο αφηγητής ή το θέμα της ιστορίας;)

β) Προσδιορισμός της αναφοράς των αντωνυμιών και των επιρρημάτων πχ. "δώσαμε στον Νορβηγό τον αναπτήρα γιατί αυτός τον ήθελε" (ποιος ήθελε ποιόν;), "ο καιρός εδώ είναι όπως και πριν από δύο μέρες" (πού ισχύει αυτό και πότε;)

γ) ελλειπτικές προτάσεις πχ. "Το μπρίκι είναι [μέσα] στο ντουλάπι" (ποιο μπρίκι; το καράβι ή το σκεύος της κουζίνας);¹⁸³

δ) Προσδιορισμός της επιτελεστικής δύναμης σε ένα ομιλιακό ενέργημα όταν αυτή δεν αναφέρεται ρητά π.χ. "Θα σε επισκεφτώ σύντομα" (πρόκειται για υπόσχεση ή για απειλή);¹⁸⁴

ε) ειρωνεία, υπερβολή και γενικότερα περιπτώσεις όπου η σημασία με την οποία εκφωνεί ο ομιλητής μια πρόταση είναι διαφορετική από την κυριολεκτική σημασία της πρότασης. Π.χ. "Η εργασία του Πέτρου είναι εξαιρετική" (είναι όντως εξαιρετική ή πολύ κακή;), "Να η πόρτα!" (μπορεί να σημαίνει, μεταξύ άλλων, ότι αυτή εδώ είναι η πόρτα ή να σημαίνει βγες έξω)¹⁸⁵

στ) μεταφορά, μετωνυμία πχ. Η Σάλλυ είναι παγόβουνο¹⁸⁶

4.2.2 Αναπαράσταση της γνώσης

Στα μέσα του 1970, έγινε μια σημαντική στροφή στον τρόπο κατανόησης της ανθρώπινης νοημοσύνης προκειμένου να αντιμετωπιστεί το πρόβλημα της πραγματολογικής γνώσης. Οι ερευνητές δεν μιλούσαν πλέον για μια διαδικασία επίλυσης προβλημάτων αλλά για μια διαδικασία αναγνώρισης προτύπων που βασίζονταν στην προγενέστερη κατάσταση του υποκειμένου και τις προσδοκώμενες από αυτό εξελίξεις. Με αφετηρία

¹⁸³ Βλ. Dreyfus 1992, σελ. 215 και Haugeland 1985, σελ. 197.

¹⁸⁴ Βλ. κεφάλαιο 3, ενότητα 3.3.1.2, τρίτη επισήμανση.

¹⁸⁵ Βλ. Grice 1989.

¹⁸⁶ Βλ. Searle 1983, σελ. 95-96.

αυτή την καθοδηγητική ιδέα δημιουργήθηκε ένα σύνολο νέων τρόπων αναπαράστασης της γνώσης: beta structures (Moore & Newell, 1973), πλαίσια (frames) (Minsky, 1975), σχήματα (schemes) (Bobrow & Norman, 1975), εννοιολογικές εξαρτήσεις (conceptual dependencies) (Schank, 1973), σενάρια (scripts) (Schank & Abelson, 1977), πρωτότυπα (prototypes) και προοπτικές (perspectives) (Bobrow & Winograd, 1977)). Κάθε μία από τις προαναφερθείσες τεχνικές κωδικοποίησης της γνώσης είχε ως αφετηρία συγκεκριμένα προβλήματα, όπως αυτά που αναφέρονται στο τέλος της προηγούμενης ενότητας. Θα εξετάσουμε ενδεικτικά ορισμένες από αυτές τις νέες προσεγγίσεις.

4.2.2.1 Πλαίσια (frames)

Τα πλαίσια αποτέλεσαν ουσιαστικά το πρότυπο για όλες τις τεχνικές αναπαράστασης της γνώσης που επακολούθησαν. Ο Minsky (1975, σελ. 212) προσδιορίζει το τι είναι ένα πλαίσιο ως εξής: "Ένα πλαίσιο είναι μια δομή δεδομένων για την αναπαράσταση στερεοτυπικών καταστάσεων, όπως το να είναι κανείς σε ένα συγκεκριμένου είδους καθιστικό ή το να πηγαίνει σε ένα παιδικό πάρτυ γενεθλίων ... Μπορούμε να σκεφτούμε ένα πλαίσιο ως ένα δίκτυο κόμβων και σχέσεων. Τα 'ανώτερα επίπεδα' ενός πλαισίου είναι παγιωμένα και παριστούν πράγματα που είναι πάντα αληθή για την υποτιθέμενη κατάσταση. Τα κατώτερα επίπεδα έχουν πολλά τερματικά-'υποδοχές' ('slots') που πρέπει να συμπληρωθούν με συγκεκριμένα στιγμιότυπα ή δεδομένα. Κάθε τερματικό μπορεί να προσδιορίζει ορισμένες συνθήκες τις οποίες οφείλουν να πληρούν οι τιμές του".

Η εφαρμογή των πλαισίων (frames) για τον προσδιορισμό της σημασίας των λέξεων της φυσικής γλώσσας οδηγεί στη συσχέτιση κάθε λέξης με ένα πρότυπο. Αυτό το πρότυπο μοιάζει με ορισμό στο βαθμό που περιλαμβάνει περιγραφές των αντικειμένων για τα οποία χρησιμοποιείται η λέξη. Ωστόσο, σε αντίθεση με τους συνήθεις ορισμούς, αυτές οι περιγραφές δεν αποτελούν ικανές και αναγκαίες συνθήκες των χρήσεων της λέξης. Το πλαίσιο μπορεί να περιλαμβάνει περιγραφές που ισχύουν σε συνήθεις περιπτώσεις ή σε σχέση με ορισμένα συμφραζόμενα, οπότε για να χρησιμοποιηθεί αποδοτικά, για παράδειγμα σε μια διαδικασία κατανόησης κειμένου, θα πρέπει να λαμβάνεται υπόψη τι είναι ήδη γνωστό για το περιγραφόμενο αντικείμενο.

Τα πλαίσια μπορούν να αξιοποιηθούν από συστήματα που δεν χρησιμοποιούν παραγωγικούς κανόνες συναγωγής μιας και επιτρέπουν τη δυνατότητα απόδοσης αντιφατικών περιγραφών. Για παράδειγμα, το πλαίσιο "πελαργός" περιλαμβάνει ως προκαθορισμένη τιμή (default) το ότι πετάει, ωστόσο αυτό δεν αποκλείει τη δυνατότητα να υπάρχει ένας πελαργός που δεν πετάει, επειδή, για παράδειγμα, είναι χτυπημένος. Τα

συμπεράσματα εξάγονται, συνήθως, με ελλιπή στοιχεία και μπορούν ακόμη και να αντιστραφούν από την προσθήκη νέων πληροφοριών. Αυτή η μη διατήρηση της μονοτονικότητας¹⁸⁷ δημιουργεί νέου τύπου προβλήματα που δεν εμφανίζονται στην παραγωγική λογική. Κάθε φορά που εξάγεται ένα συμπέρασμα πρέπει να ελέγχεται η εγκυρότητα των προηγούμενων συμπερασμάτων. Επίσης, οι συναγωγές που γίνονται με βάση ελλιπή στοιχεία οδηγούν, συνήθως, σε περισσότερα του ενός συμπεράσματα τα οποία δεν είναι συνεπή μεταξύ τους. Ένα σημαντικό ζήτημα, λοιπόν, που τίθεται, αν ακολουθήσει κανείς τη λογική των πλαισίων, είναι με ποιες διαδικασίες μπορεί να επιλεγεί το κατάλληλο συμπέρασμα¹⁸⁸.

Συμπερασματικά, μπορούμε να πούμε ότι ένα πλαίσιο είναι μια συλλογή χαρακτηριστικών, αντίστοιχων τιμών και κάποιων περιορισμών επί των τιμών (που επιβάλλονται με τη χρήση τελεστών). Η σημαντικότερη καινοτομία που εισάγουν τα πλαίσια είναι η χρήση προκαθορισμένων τιμών οι οποίες δηλώνονται ως τέτοιες με τη βοήθεια κάποιου τελεστή (π.χ. "συνήθως"). Με τις προκαθορισμένες τιμές επιχειρείται ουσιαστικά να συλληφθεί ο τρόπος με τον οποίον οι προσδοκίες προσδιορίζουν την ανθρώπινη κατανόηση.

4.2.2.2 Εννοιολογικές εξαρτήσεις (*conceptual dependencies*) και σενάρια (*scripts*)

Τα πλαίσια είναι ένας πολύ γενικός τρόπος αναπαράστασης της πραγματολογικής γνώσης στον οποίο δεν προσδιορίζεται εκ των προτέρων τι είδους αντικείμενα και σχέσεις μπορούν να περιγραφούν. Οι εννοιολογικές εξαρτήσεις και τα σενάρια αποτελούν ένα περισσότερο εξειδικευμένο τρόπο αναπαράστασης της γνώσης. Οι εννοιολογικές εξαρτήσεις αφορούν τη γνώση συμβάντων και τα σενάρια τη γνώση που ενέχεται σε ακολουθίες συμβάντων που περιγράφουν στερεοτυπικές καταστάσεις. Η γνώση αυτή περιγράφεται με τη βοήθεια ορισμένων εννοιολογικών αρχέτυπων που αφορούν πράξεις, όπως PTRANS = μεταφορά της φυσικής θέσης ενός αντικειμένου (π.χ. πηγαίνω), MTRANS = κάθε είδους κίνηση ή μετασχηματισμός νοητικών αναπαραστάσεων (π.χ.

¹⁸⁷ Ένα σύστημα είναι μονότονο αν η προσθήκη νέας γνώσης δεν μεταβάλλει την τιμή αληθείας της ήδη υπάρχουσας σε αυτό γνώσης.

¹⁸⁸ Για τον σκοπό αυτό αναπτύχθηκε αργότερα από τον McCarthy η "λογική περιγραφή" (*circumscription logic*). Ο McCarthy επιχειρεί με αυτή να εξηγήσει πώς είναι δυνατόν να κάνουμε συλλογισμούς στηριζόμενοι σε μια βάση γνώσης που περιέχει ασυνέπειες. Ο βασικός στόχος του McCarthy είναι να εξηγήσει την ανεκτικότητα που έχουν οι άνθρωποι στην αμφισημία. Όπως, όμως, επισημαίνει και ο Dreyfus (1996, σελ. 188), το σημαντικό με την ανεκτικότητα στις αμφισημίες "δεν είναι το πώς να συλλογισμάστε όταν υπάρχουν εξαιρέσεις, αλλά πώς το πλαίσιο (*context*) μας επιτρέπει να αγνοούμε τελείως τις περισσότερες εξαιρέσεις"

λέγω), PROPEL = εφαρμογή μιας φυσικής δύναμης σε ένα αντικείμενο (π.χ. σπρώχνω), MOVE = κίνηση ενός μέρους του σώματος από τον κάτοχό του (π.χ. κλωτσάω) κ.α.¹⁸⁹. Ο Schank προσπάθησε με αυτά τα αρχέτυπα να αναλύσει τη σημασία ρημάτων¹⁹⁰ και στη συνέχεια να κωδικοποιήσει στερεοτυπικές κοινωνικές δραστηριότητες ως ακολουθίες προκαθορισμένων πράξεων. Χαρακτηριστικό παράδειγμα είναι το σενάριο του εστιατορίου, όπου περιγράφονται οι συνήθεις ενέργειες στις οποίες προβαίνει κάποιος που εισέρχεται σε ένα εστιατόριο. Κάθε ενέργεια προσδιορίζεται από το τι έχει προηγηθεί και ταυτόχρονα δημιουργεί κάποιες συγκεκριμένες προσδοκίες: η παραγγελία του γλυκού προσδιορίζεται από την κατανάλωση του κυρίως πιάτου και δημιουργεί την προσδοκία για το σερβίρισμα του γλυκού. Τα σενάρια, υπ' αυτή την έννοια, ενσωματώνουν τον τρόπο που οι προσδοκίες καθοδηγούν την κατανόηση στην περίπτωση των ανθρώπων. Ωστόσο, θα πρέπει να παρατηρήσουμε ότι οι "προσδοκίες" στην περίπτωση των σεναρίων έχουν μια σημαντική διαφορά από τις ανθρώπινες προσδοκίες: πρόκειται για στερεοτυπικές προσδοκίες. Αντίθετα, οι ανθρώπινες προσδοκίες δεν είναι πάντα στερεοτυπικές. Ο άνθρωπος έχει την ικανότητα να διαμορφώνει τις προσδοκίες του βασιζόμενος και στις λεπτομέρειες της εκάστοτε συγκυρίας.

4.2.2.3 Προβλήματα που αντιμετώπισαν οι πρώτες τεχνικές αναπαράστασης της γνώσης

4.2.2.3.1 Το πρόβλημα της κοινότητας γνώσης

Το κυριότερο πρόβλημα που αντιμετώπισαν όλες οι τεχνικές αναπαράστασης της γνώσης είναι το πρόβλημα της κοινότητας γνώσης (common sense knowledge problem). Ας δούμε ένα παράδειγμα. Έστω ότι δίνεται σε ένα πρόγραμμα κατανόησης ιστοριών το ακόλουθο κείμενο: "Τη Δευτέρα πλημμύρισαν οι δρόμοι από την πολλή βροχή. Την επόμενη μέρα τα σχολεία παρέμειναν κλειστά". Με ποιο τρόπο μπορεί να εξαχθεί το συμπέρασμα ότι η πολλή βροχή είναι η αιτία τού ότι τα σχολεία παρέμειναν κλειστά; Κατ' αρχάς, το συμπέρασμα αυτό στηρίζεται σε ένα σύνολο από κοινότητες αλήθειες του τύπου "Τα σχολεία είναι κτίρια στα οποία μαζεύονται τα παιδιά τις εργάσιμες μέρες", "Τα παιδιά πηγαίνουν στα σχολεία οδικώς", "Το νερό εμποδίζει την κίνηση των αυτοκινήτων", "Τα παιδιά δεν αντέχουν τις κακουχίες", "Η προστασία της ζωής και της

¹⁸⁹ Βλ. Schank & Abelson 1977.

¹⁹⁰ Για παράδειγμα η πρόταση "Jane punched Mary" μεταφράστηκε ως εξής: "At some time past, Jane applied a force to the object Jane's fist, in the direction from-Jane-to-Mary; she did this by simply moving her fist in the direction from-Jane-to-Mary; her action of applying force to the fist caused Jane's fist to be in contact with Mary".

υγείας των παιδιών είναι σημαντικότερη από την μόρφωσή τους", "Η κολύμβηση ενέχει κινδύνους" κ.λπ. Το ζήτημα, λοιπόν, που τίθεται είναι ότι η κατανόηση στηρίζεται σε ένα τεράστιο σώμα κοινοτοπής γνώσης, η οποία δεν δηλώνεται ρητά επειδή θεωρείται αυτονόητη.

Μέχρις εδώ θα μπορούσε κανείς να αντιπείνει ότι ακριβώς επειδή η κατανόηση της φυσικής γλώσσας βασίζεται σε ένα τεράστιο σώμα γνώσης που προϋποτίθεται άρρητα, το πρόγραμμα αναπαράστασης της γνώσης στην ΤΝ έχει αποδυθεί σε μια τόσο λεπτομερή και εξαντλητική περιγραφή όλων όσων γνωρίζουμε σε διάφορα επιμέρους πεδία. Με άλλα λόγια, το πρόβλημα είναι υπαρκτό και αναγνωρισμένο από την κοινότητα της ΤΝ. Δεν αποτελεί, λοιπόν, επιχείρημα το να ισχυριστεί κανείς απλώς ότι η κατανόηση έστω και μιας πρότασης της φυσικής γλώσσας στηρίζεται σε ένα τεράστιο σώμα γνώσεων που δεν δηλώνεται ρητά στην πρόταση.

Ένας πιο προσεκτικός τρόπος ανάλυσης του προβλήματος μας οδηγεί σε τρία κρίσιμα ερωτήματα:

1. Μπορεί η κοινοτοπή γνώση να αναλυθεί σε ένα αριθμό ανεξάρτητων μεταξύ τους συνόλων γεγονότων κάθε ένα από τα οποία αποτελείται από πεπερασμένο αριθμό γεγονότων;
2. Είναι δυνατή η καταγραφή της κοινοτοπής γνώσης υπό μια μορφή που δεν θα εξαρτάται και πάλι από την κοινοτοπή γνώση;
3. Υπάρχει τρόπος εντοπισμού των συγκεκριμένων στοιχείων της βάσης γνώσης που είναι σημαντικά για την κατανόηση ενός κειμένου;

Το πραγματικό πρόβλημα με το πρόγραμμα αναπαράστασης της γνώσης στην ΤΝ είναι ότι στηρίζεται στην καταφατική απάντηση των τριών παραπάνω ερωτημάτων.

Ας δούμε με ένα παράδειγμα το πρώτο από αυτά τα ερωτήματα. Πρόκειται για ένα απόσπασμα από μια παιδική ιστορία, το οποίο ο Charniak επεχείρησε να κωδικοποιήσει μέσω πλαισίων (αναφέρεται στο Dreyfus 1992, σελ. 57):

"Σήμερα ο Jack έχει τα γενέθλιά του. Η Penny και η Janet πήγαν στο κατάστημα για να πάρουν δώρα. Η Janet αποφάσισε να πάρει έναν χαρταετό. 'Μην το κάνεις αυτό', είπε η Penny. 'Ο Jack έχει χαρταετό. Θα σε αναγκάσει να τον πάρεις πίσω'"

Το πρόβλημα που θα εξετάσουμε είναι πώς μπορεί να εντοπίσει το πρόγραμμα ότι το υπογραμμισμένο "τον" αναφέρεται στον χαρταετό που θα αγόραζε η Janet και όχι στον χαρταετό που ήδη έχει ο Jack. Είναι προφανές ότι το συντακτικό κριτήριο ότι το "τον" αναφέρεται στην τελευταία εμφάνιση της λέξης "χαρταετός" στο κείμενο δεν λειτουργεί.

Ο τρόπος που ο Charniak επεχείρησε να λύσει αυτό το πρόβλημα είναι με τη δημιουργία πολύ λεπτομερών πλαισίων που σχετίζονται με τα διάφορα στοιχεία που αναφέρονται σε τέτοιου είδους ιστορίες. Με άλλα λόγια, προσπάθησε να κωδικοποιήσει το σύνολο των γνώσεων που εμπλέκονται στην κατανόηση τέτοιου τύπου ιστοριών. Δεν μας ενδιαφέρει στην προκειμένη περίπτωση η ακριβής λύση που προτάθηκε, αλλά η

γενικότερη λογική που ακολουθήθηκε. Τι είδους γνώσεις θα έπρεπε να κωδικοποιηθούν; Ας εξετάσουμε ορισμένες από τις γνώσεις που προϋποθέτει η κατανόηση της εξεταζόμενης ιστορίας:

1. όταν κανείς έχει τα γενέθλιά του, αναμένει κάποια δώρα από τους φίλους του
2. αν κάποιος έχει ένα αντικείμενο Α, δεν συνηθίζεται να του κάνουνε δώρο ένα αντικείμενο του ίδιου τύπου
3. όταν ο Χ δίνει ένα δώρο στον Υ έχει ως στόχο να ευχαριστήσει τον Υ
4. όταν ο Υ δεχθεί από τον Χ ως δώρο ένα αντικείμενο ίδιο με κάποιο που ήδη έχει, ο Υ επιστρέφει στον Χ το δώρο κλπ.¹⁹¹

Οι παραπάνω γνώσεις έχουν την μορφή κανόνα. Ωστόσο, αυτοί οι κανόνες δεν είναι αυστηροί, έχουν πολλές εξαιρέσεις. Ας εξετάσουμε ενδεικτικά τον κανόνα 4. Με μια πρώτη ματιά γίνεται φανερό ότι δεν ισχύει αν το δώρο είναι μία σοκολάτα, ένα χιλιάρικο ή ένα δαχτυλίδι¹⁹². Πόσες τέτοιες εξαιρέσεις υπάρχουν; Είναι δυνατή η εκ των προτέρων καταγραφή όλων των δυνατών εξαιρέσεων; Είναι πεπερασμένος ο αριθμός των διαφορετικών ειδών αντικειμένων που μπορεί κανείς να δωρίσει;

Όμως, ακόμα και αν απαντήσουμε καταφατικά στο προηγούμενο ερώτημα, και δεχτούμε, επίσης, ότι με κάποιο τρόπο κατορθώνουμε να χωρίσουμε σε δύο υποσύνολα αυτά τα αντικείμενα - όπου στο ένα θα ανήκουν όλα εκείνα που μπορούν να δωριστούν μόνο αν ο αποδέκτης δεν τα έχει ήδη, ενώ στο δεύτερο όσα μπορούν να δωριστούν ακόμα και αν ο αποδέκτης διαθέτει τουλάχιστον ένα - ακόμα και σε αυτή την περίπτωση, τα πράγματα αποδεικνύονται πολύ πιο περίπλοκα. Το βασικό πρόβλημα, εν προκειμένω, είναι ότι προϋποθέτουμε ότι το εκάστοτε είδος αντικειμένων που μπορεί να δωριστεί, ανήκει αποκλειστικά είτε στο

¹⁹¹ Όμως δεν είναι μόνο τέτοιου τύπου γνώσεις στις οποίες στηρίζομαστε για να κατανοήσουμε την υπό εξέταση ιστορία. Για παράδειγμα, γνωρίζουμε τι σημαίνουν οι επιμέρους όροι που χρησιμοποιούνται: τι είναι γενέθλια, τι είναι δώρο, τι είναι κατάστημα, τι είναι χαρταετός κ.λπ. Είναι δυνατός ο ορισμός αυτών των όρων προκειμένου να "κατανοηθούν" από το πρόγραμμα; Για μια αρνητική απάντηση σε αυτό το ερώτημα βλ. Wittgenstein 1953/1977 (βλ. υποσημείωση 201) και Quine 1951.

Με κάποιο τρόπο, επίσης, καταλαβαίνουμε ότι η Penny και η Janet όταν πήγαν στο μαγαζί για να αγοράσουν δώρα, αυτά τα δώρα ήταν για τον Jack και ότι όταν η Janet αποφάσισε να αγοράσει έναν χαρταετό, τον αγόρασε με στόχο να τον κάνει δώρο (βλ. Dreyfus 1992, σελ 58). Αυτά τα στοιχεία δεν αναφέρονται ρητά στο κείμενο, και όμως εμείς τα καταλαβαίνουμε. Πώς γίνεται, άραγε, αυτό;

¹⁹² Υπάρχουν όμως και πολλές περιπτώσεις στις οποίες είναι εντελώς ακαθόριστο το εάν ισχύει ή όχι ο κανόνας. Για παράδειγμα, τι γίνεται αν το δώρο είναι ένα εισιτήριο στον τελικό του μπάσκετ; Το ότι έχει κανείς ήδη ένα εισιτήριο, τον κάνει να επιστρέψει το δώρο ή μήπως κρατάει το δεύτερο εισιτήριο για να το δώσει σε κάποιον άλλον;

πρώτο υποσύνολο είτε στο δεύτερο. Με άλλα λόγια, δεχόμαστε ότι το εάν ένα αντικείμενο μπορεί να αποτελέσει ή όχι δώρο καθορίζεται αποκλειστικά από το είδος του αντικειμένου. Όμως, ένας δεύτερος χαρταετός θα μπορούσε να είναι ένα πολύ αποδεκτό δώρο για έναν συλλέκτη χαρταετών. Με άλλα λόγια, το εάν ένα δώρο είναι αποδεκτό ή όχι εξαρτάται εκτός από το είδος του δώρου και από την κατηγορία του αποδέκτη. Είναι δυνατή η εκ των προτέρων καταγραφή για κάθε δυνατό δώρο όλων των δυνατών αποδεκτών του; Όμως, ακόμα και αν δεχτούμε αυτό το εξαιρετικά απίθανο ενδεχόμενο, πάλι θα βασιζόμασταν σε μια αμφίβολη προκειμένη: θα προϋποθέταμε, δηλαδή, ότι κάθε δυνατός παραλήπτης ενός συγκεκριμένου δώρου ενεργεί πάντα με τον ίδιο τρόπο. Για παράδειγμα, είναι σαν να λέμε ότι οι συλλέκτες χαρταετών δεν επιστρέφουν ποτέ τους χαρταετούς που δέχονται ως δώρα. Τι γίνεται, όμως, αν ο συλλέκτης δεχθεί για εκατοστή τρίτη φορά τον ίδιο χαρταετό ως δώρο. Ίσως να τον κρατήσει ίσως και όχι. Το τι θα πράξει, τελικά, είναι συνάρτηση της συγκεκριμένης κατάστασης στην οποία βρίσκεται. Αν, για παράδειγμα, δεν έχει άλλο χώρο στο σπίτι του, ενδέχεται να τον επιστρέψει. Αν πάλι αυτός που του κάνει το δώρο είναι κάποιο σημάιων στέλεχος της Μαφίας, τότε ακόμα και αν δεν έχει χώρο στο σπίτι του, ενδεχομένως να κρατήσει το δώρο για να μην διακινδυνεύσει τη ζωή του.

Βλέπουμε, λοιπόν, ότι ακόμα και αν εστιάσουμε την προσοχή μας σε έναν μόνο κανόνα οδηγούμαστε σε μια, εκ πρώτης όψεως, απειρία διαφορετικών συνθηκών που προσδιορίζουν τον τρόπο εφαρμογής του. Χωρίς τη γνώση αυτών των επιμέρους συνθηκών ο κανόνας είναι ουσιαστικά άχρηστος. Η δυνατότητα, όμως, ορθής χρήσης του κανόνα από ένα πρόγραμμα προϋποθέτει την εκ των προτέρων καταγραφή όλων των επιμέρους συνθηκών. Κατά συνέπεια, δεν μπορεί να δοθεί, εκ πρώτης τουλάχιστον όψεως, μια θετική απάντηση στο πρώτο από τα τρία υποερωτήματα με τα οποία αναλύσαμε το πρόβλημα της κοινότοπης γνώσης, γιατί οι συνθήκες που προσδιορίζουν τον τρόπο χρήσης έστω και ενός μόνο κανόνα φαίνεται να είναι άπειρες. Ωστόσο, θα πρέπει να επισημάνουμε ότι δεν παραθέσαμε κάποιο επαχείρημα που οδηγεί με βεβαιότητα στο συμπέρασμα ότι ο αριθμός των γεγονότων που συνιστούν την κοινότοπη γνώση (ή κάποιες αυτόνομες υποπεριοχές της) είναι άπειρος. Αυτό που δείξαμε είναι ότι ο αριθμός των γεγονότων και των κανόνων που συνιστούν την κοινότοπη γνώση είναι εξαιρετικά μεγάλος. Κατά συνέπεια, μπορούμε βάσιμα να υποθέσουμε ότι η αναπαράσταση της κοινότοπης γνώσης οδηγεί σε ένα πολύ σημαντικό πρόβλημα πολυπλοκότητας.

Ακόμα, όμως, και αν υποθέσουμε ότι ξεπερνάμε τα προβλήματα που θέτει το πρώτο από τα υποερωτήματα, απομένουν τα υπόλοιπα δύο. Το δεύτερο κρίσιμο ερώτημα, είναι το εάν είναι δυνατή η καταγραφή της κοινότοπης γνώσης υπό μια μορφή που δεν θα εξαρτάται και πάλι από την κοινότοπη γνώση; Το πρόβλημα αυτό αντιμετωπίζεται στην κλασική ΤΝ με τη χρήση τεχνητών γλωσσών. Οι τεχνητές γλώσσες είναι κατασκευασμένες κατά τέτοιο τρόπο ώστε να αποφεύγουν την αμφισβή

που παρουσιάζει η φυσική γλώσσα. Επίσης, οι εν λόγω γλώσσες στηρίζονται σε ένα πεπερασμένο αριθμό εννοιολογικών αρχετύπων με βάση τα οποία επιχειρούν να εκφράσουν τις έννοιες που εκφράζονται με τη φυσική γλώσσα. Κάθε εννοιολογικό αρχέτυπο είναι, εκ κατασκευής, ανεξάρτητο από όλα τα υπόλοιπα. Προκειμένου, λοιπόν, η καταγραφή της κοινότοπης γνώσης να μπορεί να γίνει κατά τέτοιον τρόπο ώστε να μην εξαρτάται και πάλι από την κοινότοπη γνώση, θα πρέπει να υπάρχει ένα πεπερασμένο σύνολο εννοιολογικών αρχετύπων με το οποίο να μπορεί να εκφραστεί κάθε τι που απαρτίζει την κοινότοπη γνώση. Η συγκεκριμένη προσέγγιση οδήγησε στην ανάπτυξη και επεξεργασία νέων οντολογιών για την ανάλυση της γνώσης. Το εγχείρημα αυτό έχει εξίσου οδηγηθεί σε ανυπέρβλητα προβλήματα.

Τέλος, ακόμα και αν απαντήσουμε καταφατικά στα δύο πρώτα υποερωτήματα, δηλαδή θεωρήσουμε, πρώτον, ότι η κοινότοπη γνώση μπορεί να αναλυθεί σε ένα αριθμό ανεξάρτητων μεταξύ τους συνόλων γεγονότων κάθε ένα από τα οποία αποτελείται από πεπερασμένο αριθμό γεγονότων και, δεύτερον, ότι αυτά τα γεγονότα μπορούν να καταγραφούν υπό μια μορφή που δεν θα εξαρτάται και πάλι από την κοινότοπη γνώση, ακόμα και σε αυτή την περίπτωση παραμένει το τρίτο ερώτημα: Με ποιο τρόπο μπορεί να εντοπιστεί ποια συγκεκριμένα στοιχεία της βάσης γνώσης είναι σημαντικά για την κατανόηση ενός κειμένου; Πρόκειται για το πρόβλημα της *σημαντικότητας* (relevance). Ο Haugeland (1985) αναλύει το συγκεκριμένο πρόβλημα σε μια σειρά επιμέρους προβλημάτων τα οποία θα αναφέρουμε εν συντομία.

Καταρχήν, ο συνήθης τρόπος αντιμετώπισης του προβλήματος της σημαντικότητας είναι με την αναζήτηση λέξεων-κλειδιών. Ωστόσο τα πράγματα δεν είναι τόσο απλά, πολύ συχνά απουσιάζουν οι καθοδηγητικές λέξεις. Ο Haugeland (1985, σελ. 201) αναφέρει ένα χαρακτηριστικό παράδειγμα:

A: Που είναι το μπρίκι;

B: Δεν ξέρω, αλλά ρίξε μια ματιά στο ντουλάπι.

Στην προκειμένη περίπτωση, λείπει η λέξη "μέσα" (πριν από τη φράση "στο ντουλάπι") που θα ενεργοποιήσει το σχετικό πλαίσιο (δηλαδή, το πλαίσιο "μέσα") ούτως ώστε να διασαφηνιστεί ότι η λέξη "μπρίκι" αναφέρεται στο σκεύος της κουζίνας και όχι στο πλοίο, διότι ένα πλοίο, λόγω μεγέθους, δεν μπορεί να βρίσκεται μέσα σε ένα ντουλάπι¹⁹³.

Ένα δεύτερο πρόβλημα, κατά τον Haugeland (1985 σελ. 202), είναι το πρόβλημα της απουσίας θέματος:

Άφησα το αδιάβροχο στο μπάνιο, επειδή ήταν βρεμένο.

Ποιο ήταν βρεμένο;

¹⁹³ Εκτός και αν πρόκειται για μινιατούρα πλοίου ή το ντουλάπι είναι τεραστίων διαστάσεων...

Η απάντηση σε αυτό το ερώτημα δεν φαίνεται να ανήκει ούτε στο στερεότυπο του "αδιόβροχου", ούτε στο στερεότυπο του "μπάνιου". Κατά συνέπεια, δεν φαίνεται να υπάρχει τρόπος με τον οποίον θα μπορούσε να προβλεφτεί εκ των προτέρων μια τέτοιου είδους "γνώση".

Τέλος, σύμφωνα με τον Haugeland (1985, σελ. 203), η κατάσταση χειροτερεύει δραματικά όταν διευρύνεται το προς ανάλυση κείμενο. Αρκεί να σκεφτούμε την παιδική ιστορία με τον χαρταετό που αναλύθηκε παραπάνω. Η εκτίμηση της συνολικής κατάστασης φαίνεται να προηγείται και να προσδιορίζει την κατανόηση των επιμέρους προτάσεων.

Θα μπορούσαμε, επίσης, να προσθέσουμε το πρόβλημα των προκαθορισμένων τιμών (defaults): πώς μπορεί να "αποφασίσει" το πρόγραμμα το εάν πρέπει να μεταχειριστεί την εξεταζόμενη κατάσταση ως τυπική και έτσι να υιοθετήσει την προκαθοριζόμενη τιμή του πλαισίου και όχι κάποια άλλη; Βέβαια, θα μπορούσε κανείς, εκπαιδευμένος στην λογική τέτοιων προγραμμάτων, να αντιτείνει ότι, εν προκειμένω, δεν τίθεται κανένα ζήτημα απόφασης: η προκαθορισμένη τιμή ενός χαρακτηριστικού χρησιμοποιείται πάντα, εκτός εάν έχει "κληρονομηθεί" ή δοθεί κάποια διαφορετική τιμή για το συγκεκριμένο χαρακτηριστικό. Μια τέτοια απάντηση, ωστόσο, δεν αντιμετωπίζει τον πυρήνα του προβλήματος. Το πότε αντιμετωπίζουμε κάτι ως τυπικό δεν είναι πάντοτε συνάρτηση απλώς και μόνο της προγενέστερης ιστορίας.

4.2.2.3.2 Το πρόβλημα του πλαισίου

Ένα ανάλογο πρόβλημα με το πρόβλημα της κοινότητας γνώσης έχει εμφανιστεί στην ρομποτική, με τη διαφορά ότι δεν αφορά την κοινότητα γνώσης αλλά ένα άλλο είδος πραγματολογικής γνώσης, την περιστασιακή ή προσωρινή γνώση. Το πρόβλημα μπορεί να διατυπωθεί ως εξής: ένα ρομπότ για να λειτουργήσει μέσα στον κόσμο βασίζεται στην γνώση που έχει για τον κόσμο. Η γνώση αυτή είναι μεταβαλλόμενη επειδή και ο κόσμος είναι μεταβαλλόμενος. Ένας τρόπος μεταβολής του κόσμου είναι από τις παρεμβάσεις του ρομπότ. Τουλάχιστον, λοιπόν, για αυτές τις αλλαγές θα πρέπει το ρομπότ να έχει κάποιον τρόπο παρακολούθησής τους για να τροποποιεί κατάλληλα τη βάση γνώσης του, ούτως ώστε αυτή να εκφράζει κάθε φορά την τρέχουσα κατάσταση του κόσμου. Το πρόβλημα που δημιουργείται είναι: πώς "καταλαβαίνει" το ρομπότ, δεδομένης μιας πράξης του, ποια από όλα τα γεγονότα που είναι αποθηκευμένα στη βάση γνώσης πρέπει να αλλάξει; Πρόκειται, ουσιαστικά, για μια συγκεκριμένη έκφανση του προβλήματος της σημαντικότητας.

Αυτό το πρόβλημα, στο πλαίσιο της κλασικής ρομποτικής, επιχειρείται να λυθεί με τη βοήθεια των λεγόμενων "πλαισιακών αξιωμάτων" (frame axioms). Τα αξιώματα αυτά προσδιορίζουν για κάθε πράξη συγκεκριμένες συνέπειες. Για παράδειγμα, τέτοια αξιώματα θα μπορούσαν να είναι τα εξής (βλ. Lozand 1999): "το πάτημα του διακόπτη μεταβάλλει την φωτεινότητα της λάμπας", "Η πώληση της λάμπας μεταβάλλει την ιδιοκτησία της" κλπ. Όπως, όμως, επισημαίνει ο Lozand (1999), εφόσον το ρομπότ μόνο μέσω συναγωγής μπορεί να εξαγάγει ποια

γεγονότα της βάσης γνώσης ισχύουν, χρειάζεται να διαθέτει και αξιώματα για το τι δεν αλλάζει, δεδομένης κάποιας αλλαγής. Τέτοια αξιώματα θα μπορούσαν να είναι, για παράδειγμα, τα ακόλουθα: "Το πάτημα του διακόπτη δεν μεταβάλλει την ιδιοκτησία της λάμπας", "Η πώληση της λάμπας δεν μεταβάλλει την φωτεινότητά της". Έτσι, αν θεωρήσουμε ότι η βάση γνώσης κωδικοποιεί n γεγονότα και ότι κάθε πράξη μεταβάλλει μόνο ένα από αυτά τα γεγονότα, το πρόβλημα που προκύπτει είναι ότι για κάθε πλαισιακό αξίωμα μεταβολής ενός γεγονότος χρειάζονται $n-1$ πλαισιακά αξιώματα που καθορίζουν την μη-μεταβολή των υπολοίπων $n-1$ γεγονότων της βάσης γνώσης. Γίνεται, λοιπόν, φανερό ότι οδηγούμαστε έτσι σε ένα πολύ μεγάλο αριθμό πλαισιακών αξιωμάτων. Αυτός είναι ο τρόπος που διατυπώθηκε το πρόβλημα του πλαισίου από τους επιστήμονες της ρομποτικής.

Ωστόσο, το πρόβλημα είναι στην πραγματικότητα πολύ δυσκολότερο για δύο τουλάχιστον λόγους:

1. Είναι πολύ αμφίβολο ότι οι δυνατές πράξεις που μπορεί να επιτελέσει ένα ρομπότ είναι δυνατόν να καθορισθούν εκ των προτέρων - ειδικότερα αν το ρομπότ δεν κινείται σε κάποιον προκατασκευασμένο κόσμο. Δεν είναι τυχαίο, εξάλλου, ότι τα ρομπότ που κατασκευάζονται με τη λογική της κλασικής ΤΝ παρουσιάζουν μια έντονη "ευθραυστότητα" (brittleness) στη συμπεριφορά τους μόλις αφεθούν να λειτουργήσουν στον πραγματικό κόσμο. Ένας βασικός λόγος για τον οποίο οι δυνατές πράξεις που μπορεί να επιτελέσει ένα ρομπότ είναι εξαιρετικά δύσκολο να καθορισθούν εκ των προτέρων είναι ότι για να εξατομικευθεί μια πράξη δεν αρκεί η γνώση της σωματικής κίνησης με την οποία επιτελείται. Η ίδια σωματική κίνηση επιτελούμενη σε δύο διαφορετικά περιβάλλοντα μπορεί να συνιστά διαφορετική πράξη. Για παράδειγμα, η ίδια κίνηση του χεριού μπορεί στην μία περίπτωση να κλείνει το φως και στην άλλη να το ανοίγει. Με άλλα λόγια, η ταυτότητα της πράξης εξαρτάται εν μέρει και από τον ίδιο τον κόσμο.
2. Είναι πολύ αμφίβολο ότι κάθε πράξη συνδέεται πάντα με την αλλαγή των ιδίων γεγονότων. Με άλλα λόγια, είναι πολύ αμφίβολο ότι τα γεγονότα που επηρεάζονται καθορίζονται αποκλειστικά από την ταυτότητα της πράξης και όχι από τις συνθήκες εντός των οποίων επιτελείται η πράξη. Έστω, για παράδειγμα, το ακόλουθο πλαισιακό αξίωμα: η μεταφορά ενός αντικείμενου από ένα σημείο x σε ένα σημείο y του επιπέδου έχει ως συνέπεια το αντικείμενο, μετά το πέρας της πράξης, να βρίσκεται στο σημείο y . Ωστόσο, αυτός ο κανόνας είναι φανερό ότι δεν ισχύει πάντα. Αν το μετακινούμενο αντικείμενο είναι συνδεδεμένο με ένα ελατήριο, τότε μετά το πέρας της πράξης το αντικείμενο δεν θα παραμείνει στο σημείο y . Το ίδιο και αν το αντικείμενο είναι μεταλλικό και στο σημείο x υπάρχει ένας μαγνήτης. Μόνο η φαντασία μας μπορεί να περιορίσει το εύρος των δυνατών εξαιρέσεων. Όμως, και οι εξαιρέσεις έχουν εξαιρέσεις: το αντικείμενο θα παραμείνει στο σημείο y αν το ελατήριο είναι πολύ λεπτό, ή πολύ

κοντό ή πολύ εύθραστο κ.λπ. ή αν ο μαγνήτης είναι πολύ ασθενής, ή η απόσταση xy πολύ μεγάλη, ή στο y υπάρχει ένας άλλος ισχυρότερος μαγνήτης κ.λπ.

Στην κλασική ΤΝ έχουν αναπτυχθεί αρκετές μέθοδοι για την αποφυγή του προβλήματος του πλαισίου (βλ. ενδεικτικά Haugeland 1985, σελ. 205-211, Rich & Knight 1991, σελ. 126-128). Ωστόσο, καμιά από τις προτεινόμενες λύσεις δεν κατορθώνει να αντιμετωπίσει το πρόβλημα της σημαντικότητας το οποίο είναι ουσιαστικά η πηγή του προβλήματος του πλαισίου.

Προφανώς, εμείς οι άνθρωποι δεν πάσχουμε από το πρόβλημα του πλαισίου. Με κάποιο τρόπο, κάθε φορά που προβαίνουμε σε μια ενέργεια, γνωρίζουμε άμεσα τι αλλάζει και τι δεν μεταβάλλεται από όλα όσα ισχυον πριν την πράξη μας. Για αυτό ακριβώς τον λόγο είναι δύσκολο να κατανοήσουμε το πρόβλημα του πλαισίου. Δεν είναι τυχαίο, εξάλλου, ότι το πρόβλημα αυτό δεν εντοπίστηκε από κάποιους γνωσιολόγους αλλά από τους επιστήμονες που ασχολήθηκαν με την κατασκευή ρομπότ.

Τόσο το πρόβλημα του πλαισίου, όσο και το πρόβλημα της κοινότητας γνώσης αποτελούν σημαντικότερα εμπόδια για την ανάπτυξη του προγράμματος αναπαράστασης της γνώσης στην ΤΝ. Ωστόσο, οι ερευνητές στην κοινότητα της κλασικής ΤΝ δεν ποιήθηκαν από αυτά τα εμπόδια και συνέχισαν στη ίδια λογική στοχεύοντας σε όλο και λεπτομερέστερες αναλύσεις της πραγματολογικής γνώσης. Χαρακτηριστικότερο παράδειγμα προς αυτή την κατεύθυνση αποτελούν οι Lenat και Feigenbaum.

4.2.2.4 CYC : Το πρόγραμμα κωδικοποίησης της κοινότητας γνώσης

Ο Lenat και ο Feigenbaum είναι σήμερα οι πιο σημαντικοί, ίσως, ερευνητές στο ζήτημα της κοινότητας γνώσης. Βρίσκονται επικεφαλής ενός δεκαετούς ερευνητικού προγράμματος που ξεκίνησε το 1984, προϋπολογισμού 50 εκ. δολαρίων, το οποίο έχει ως στόχο την καταγραφή της κοινότητας γνώσης. Το πρόγραμμα αυτό ονομάζεται CYC (από το encyclopedia) και έχει χαρακτηριστεί ως: "Το συμπλήρωμα μιας εγκυκλοπαίδειας: όλα όσα γνωρίζουμε χωρίς να χρειάζεται ποτέ να τα διατυπώσουμε"¹⁹⁴

Οι βασικές θέσεις που προσδιορίζουν την έρευνα στο CYC είναι σύμφωνα με τους Lenat και Feigenbaum οι ακόλουθες:

1. Η αρχή της γνώσης: Για να εκτελέσει ένα πρόγραμμα κάποιο σύνθετο έργο με αποδοτικό τρόπο θα πρέπει να γνωρίζει πάρα πολλά σχετικά με τον κόσμο μέσα στον οποίο λειτουργεί. Αν απουσιάζει μια τέτοια

¹⁹⁴ Smith 1991

- γνώση, τότε αυτό που απομένει είναι η αναζήτηση (search) και ο συλλογισμός, που όμως δεν είναι αρκετά.
2. Η υπόθεση της ευρύτητας: Για να συμπεριφέρεται ένας πράττων (agent) με νοήμονα τρόπο σε απροσδόκητες καταστάσεις πρέπει να είναι ικανός να προσφεύγει σε όλο και γενικότερες γνώσεις και να αναγνωρίζει επιφανειακές αναλογίες ανάμεσα στην συγκεκριμένη κατάσταση και την γενικότερη γνώση.
 3. Η Τεχνητή Νοημοσύνη ως εμπειρική έρευνα: Η πρώτη μαθηματικοποίηση ή επικέντρωση σε προβλήματα - παιχνίδια (toy-problems) εξαλείφει τις λεπτομέρειες της πραγματικότητας οι οποίες στην συνέχεια αποδεικνύονται πολύ σημαντικές. Γι' αυτό το λόγο θα πρέπει να δοκιμάζουμε τις ιδέες μας πειραματικά σε πραγματικά προβλήματα όπου υπάρχει η δυνατότητα διάψευσης.

Οι Lenat και Feigenbaum χρησιμοποιούν για την κωδικοποίηση της κοινότοπης γνώσης τις σημαντικότερες τεχνικές που αναπτύχθηκαν για την αναπαράσταση της γνώσης. Επίσης, έχουν επεξεργαστεί μια νέα οντολογία στην οποία περιγράφουν τους διάφορους τύπους οντοτήτων και σχέσεων υπό μια ανεξάρτητη-από-τα-συμφραζόμενα μορφή. Με βάση αυτή την οντολογία ανέπτυξαν μια ειδική γλώσσα την LCYC. Οι Lenat και Feigenbaum θεωρούν συνεπώς ότι η κοινότοπη γνώση μπορεί να γραφτεί προτασιακά υπό την μορφή γεγονότων και κανόνων. Υπολογίζουν μάλιστα ότι για την καταγραφή αυτής της γνώσης επαρκούν περίπου ένα εκατομμύριο πλαίσια.

Οι Lenat και Feigenbaum αναγνωρίζουν ότι η φυσική τάση σε κάθε ερευνητικό πρόγραμμα αναπαράστασης της γνώσης είναι να γίνεται όλο και πιο αργό καθώς προστίθεται σε αυτό νέα γνώση, ωστόσο πιστεύουν ότι το κλειδί για την επίλυση του εν λόγω προβλήματος είναι η συνετή προσθήκη μεταγνώσης για την μείωση του χώρου έρευνας¹⁹⁵. Με αυτό τον τρόπο ελπίζουν να λύσουν το πρόβλημα της σημαντικότητας (relevance), δηλαδή, το πώς ο υπολογιστής θα εντοπίζει άμεσα εκείνα τα στοιχεία της βάσης γνώσης που είναι σημαντικά για την εκάστοτε περίπτωση¹⁹⁶.

¹⁹⁵ Βλ. Lenat & Feigenbaum 1991, σελ. 187.

¹⁹⁶ "Συχνά μας ρωτούν πώς αναμένουμε να "δεικτοδοτήσουμε" (index) αποτελεσματικά -δηλαδή, να βρούμε σημαντικές επιμέρους συνταντίσεις - καθώς η βάση γνώσης γίνεται όλο και μεγαλύτερη ... Η απάντησή μας ... είναι εκ πρώτης όψεως αναπάντεχη: Αναμείναιτε μέχρι τα προγράμματά μας να αρχίσουν να βρίσκουν πολλές, ευρέως διασπαρμένες αναλογίες, αλλά με μη αποτελεσματικό τρόπο, δηλαδή μόνο μέσω μεγάλων αναζητήσεων. Κατόπιν ερευνηστε τι πρόσθετη γνώση χρησιμοποιούν οι άνθρωποι για να εξαλείψουν μεγάλα τμήματα του χώρου αναζήτησης σε αυτές τις περιπτώσεις. Κωδικοποιήστε τη γνώση που εξήχθει κατ' αυτό τον τρόπο και προσθέστε τη στο σύστημα." Lenat & Feigenbaum 1991, σελ. 221.

Η λύση που προτείνουν οι Lenat και Feigenbaum είναι αρκετά προβληματική και δεν παρουσιάζει τα αναμενόμενα αποτελέσματα. Οι κανόνες, όταν αναφέρονται σε μη δομημένες περιοχές γνώσης, παρουσιάζουν πάντα εξαιρέσεις και το πρόβλημα, προφανώς, δεν λύνεται με το να προσθέσουμε κανόνες για αυτές τις εξαιρέσεις, γιατί και αυτοί οι κανόνες παρουσιάζουν άλλες εξαιρέσεις κ.ο.κ.

Το πρόγραμμα CYC έχει συμπληρώσει τα δέκα χρόνια για τα οποία προβλεπόταν η χρηματοδότησή του και ήδη πέρασαν και τα δύο χρόνια παράτασης που ζήτησε ο Lenat, χωρίς τα αναμενόμενα αποτελέσματα.

4.3 Υπάρχει λύση στα προβλήματα του πλαισίου και της κοινοτόπης γνώσης;

Τα προβλήματα του πλαισίου και της κοινοτόπης γνώσης αποτελούν προβλήματα μόνο για την κλασική ΤΝ και όχι για τους ανθρώπους. Αυτό σημαίνει ότι οι άνθρωποι με κάποιο τρόπο τα επιλύουν. Με ποιο τρόπο κατορθώνουν να επιτελέσουν ένα τόσο δύσκολο έργο;

4.3.1 Βασικές παραδοχές στις οποίες στηρίζεται η κλασική ΤΝ

Η κλασική ΤΝ δεν έκανε τίποτα άλλο από το να ακολουθήσει και να αναπτύξει με συνέπεια μια ευρύτατα διαδεδομένη αντίληψη περί της ανθρώπινης γνώσης ότι, δηλαδή, το σύνολο της ανθρώπινης γνώσης είναι ένα άθροισμα συμβολικών αναπαραστάσεων που αφορούν αντικειμενικά και αυθυπόστατα γεγονότα μαζί με ένα σύνολο ανεξάρτητων από το πλαίσιο χρήσης κανόνων για τη σύνδεση των γεγονότων. Με το πρόγραμμα αναπαράστασης της γνώσης διερευνήθηκε για πρώτη φορά στις λεπτομέρειές της αυτή η αντίληψη που, μέχρι τότε, διατυπωνόταν απλάς προγραμματικά στο πλαίσιο φιλοσοφικών και ψυχολογικών θεωριών. Τρεις κυρίαρχες παραδοχές στις οποίες στηρίζεται η συγκεκριμένη αντίληψη είναι οι ακόλουθες:

1. υπάρχει ένας προδεδομένος αντικειμενικός κόσμος που απαρτίζεται από ένα σύνολο αυθυπόστατων γεγονότων
2. τα γεγονότα αυτά συνδέονται μεταξύ τους με συγκεκριμένους τρόπους που εξαρτώνται αποκλειστικά από την ταυτότητα των εκάστοτε συνδεδόμενων γεγονότων και όχι από το πλαίσιο ή την κατάσταση εντός της οποίας λαμβάνουν χώρα τα εν λόγω γεγονότα
3. ο άνθρωπος γνωρίζει τον κόσμο με το να σχηματίζει εσωτερικές αναπαραστάσεις για τα γεγονότα του κόσμου και τους τρόπους σύνδεσής τους

Η titάνια προσπάθεια στην οποία οδηγήθηκε το πρόγραμμα αναπαράστασης της γνώσης, προκειμένου να εντοπίσει και να καταγράψει τα γεγονότα που "απαρτίζουν" τον κόσμο καθώς και τους κανόνες σύνδεσης των γεγονότων, καθίσταται εντελώς αδικαιολόγητη χωρίς την ισχύ των παραπάνω τριών παραδοχών. Στην ενότητα 4.2.2.3 ασκήσαμε κριτική στην δεύτερη παραδοχή δείχνοντας ότι ένα από τα βασικότερα

προβλήματα που αντιμετώπισε το εγχείρημα της κλασικής ΤΝ, ήταν η αδυναμία να εντοπιστούν κανόνες σύνδεσης γεγονότων που να εξαρτώνται αποκλειστικά από την ταυτότητα των συνδεόμενων γεγονότων και όχι από το πλαίσιο ή την κατάσταση εντός της οποίας λαμβάνουν χώρα τα γεγονότα. Αν δεχτούμε ως βάσιμη αυτή την κριτική, τίθεται το ερώτημα πώς μπορεί να αναλυθεί η κατάσταση εντός της οποίας λαμβάνουν χώρα τα γεγονότα. Με βάση την αντίληψη που επικρατεί στην κλασική ΤΝ, ο μόνος τρόπος που μπορεί να γίνει κάτι τέτοιο είναι με αναφορά σε πρόσθετα γεγονότα γιατί, σύμφωνα με την παραδοχή (1), ο κόσμος δεν είναι παρά το άθροισμα γεγονότων. Έτσι, αν δεχθεί κανείς τη συγκεκριμένη αντίληψη που εξετάζουμε, η κατάσταση εντός της οποίας ένα γεγονός οδηγεί σε κάποιο άλλο μπορεί να περιγραφεί μόνο ως ένα σύνολο επιπλέον γεγονότων που συνδέονται με κάποιον συγκεκριμένο τρόπο. Τίθεται, λοιπόν, εκ νέου, το πρόβλημα τού να ορισθούν οι κανόνες που συνδέουν αυτά τα επιπλέον γεγονότα που απαρτίζουν μία κατάσταση. Είναι δυνατόν τα συγκεκριμένα γεγονότα να συνδέονται κατά τρόπους που εξαρτώνται αποκλειστικά από την ταυτότητα των ιδίων των γεγονότων και όχι από τη συνολικότερη κατάσταση εντός της οποίας λαμβάνουν χώρα; Η κριτική που ασκήσαμε μέχρι τώρα στο πρόγραμμα αναπαράστασης της γνώσης μας υποδεικνύει μια αρνητική απάντηση. Όμως, ακόμα και αν υποθέταμε ότι κάτι τέτοιο είναι δυνατό, ανακύπτει ένα πρόσθετο πρόβλημα: είναι πεπερασμένος ο αριθμός των καταστάσεων; Αν σκεφτούμε ότι κάθε κατάσταση, με τη λογική της κλασικής ΤΝ, ορίζεται ως ένα σύνολο γεγονότων μεταξύ των οποίων ισχύουν συγκεκριμένες σχέσεις, οδηγούμαστε, αν όχι σε έναν άπειρο, τουλάχιστον σε έναν υπερβολικά μεγάλο αριθμό καταστάσεων. Αυτό από μόνο του βέβαια, πέρα από προβλήματα περιπλοκότητας, δεν φαίνεται να αποτελεί κάποιο ανυπέρβλητο εμπόδιο. Το σημαντικότερο πρόβλημα δημιουργείται από την παραδοχή (3), σύμφωνα με την οποία, για να γνωρίζει κάποιος ένα σύνολο γεγονότων πρέπει να διαθέτει εσωτερικές αναπαραστάσεις αυτών των γεγονότων και των κανόνων σύνδεσής τους. Έτσι, γνώση καταστάσεων σημαίνει για την κλασική ΤΝ την εκ των προτέρων αναπαράσταση των καταστάσεων.

Όλα αυτά τα προβλήματα γεννιούνται από την προσπάθεια να κατανοηθούν οι καταστάσεις ως ένα σύνολο γεγονότων μεταξύ των οποίων ισχύουν αντικειμενικές σχέσεις. Αυτή είναι η μόνη οδός για οποιονδήποτε δέχεται την παραδοχή (1), ότι, δηλαδή, ο κόσμος απαρτίζεται αποκλειστικά από ένα σύνολο αυθυπόστατων γεγονότων. Είναι δυνατόν να αμφισβητήσουμε την παραδοχή (1), δηλαδή να κατανοήσουμε τον κόσμο ως κάτι διαφορετικό από ένα σύνολο αντικειμενικών και αυθυπόστατων γεγονότων; Πιο συγκεκριμένα, τι θα μπορούσαν να είναι οι καταστάσεις αν όχι ένα σύνολο αντικειμενικών και αυθυπόστατων γεγονότων τα οποία συνδέονται με προκαθορισμένους τρόπους;

4.3.2 Εναλλακτικές θεωρήσεις της έννοιας της κατάστασης και η αναγκαιότητα ύπαρξης μη-προτασιακής γνώσης

Μια εναλλακτική πρόταση είναι να θεωρήσουμε τις καταστάσεις όχι ως κάτι που αφορά τον αντικειμενικό κόσμο, αλλά ως κάτι που αφορά τον τρόπο με τον οποίο το υποκείμενο συνδέεται με τον κόσμο, ή, καλύτερα, *υπάρχει* μέσα στον κόσμο. Σε μια πρώτη προσέγγιση, αυτό θα μπορούσε να σημαίνει ότι οι πεποιθήσεις και οι επιθυμίες ενός ανθρώπου επηρεάζουν τον τρόπο που αντιλαμβάνεται τον κόσμο. Πρόκειται για την γνωστή θέση ότι η αντίληψη είναι θεωρητικά φορτισμένη. Σύμφωνα με αυτή την προσέγγιση, ο τρόπος που αντιλαμβανόμαστε τον κόσμο δομείται ουσιαστικά από την θεωρία που ήδη έχουμε για τον κόσμο. Θα ονομάσουμε αυτή τη θέση "θεωρητικό ολισμό"¹⁹⁷.

Ο θεωρητικός ολισμός θέτει υπό αμφισβήτηση τόσο την παραδοχή (1) όσο και την παραδοχή (2), διότι α) ο κόσμος τον οποίο γνωρίζει το υποκείμενο συγκροτείται, ως ένα βαθμό, από το ίδιο το υποκείμενο και β) ο τρόπος που συνδέονται δύο γεγονότα εξαρτάται από το σύνολο της θεωρίας του υποκειμένου. Ωστόσο, ο θεωρητικός ολισμός, υπό μια έννοια, διατηρεί την παραδοχή (3). Αυτό συμβαίνει γιατί η ήδη υπάρχουσα γνώση που έχει το υποκείμενο για τον κόσμο εκλαμβάνεται ως ένα σύνολο πεποιθήσεων και κανόνων, δηλαδή ως μια *θεωρία* του κόσμου. Έτσι, ακόμα και αυτό που ονομάσαμε κοινότοπη γνώση, σύμφωνα με τον θεωρητικό ολισμό, δεν μπορεί παρά να είναι ένα είδος θεωρίας περί του κόσμου, δηλαδή ένα σύνολο αναπαραστάσεων για τα γεγονότα του κόσμου και τους τρόπους σύνδεσής τους¹⁹⁸.

Η κριτική που ασκούν ο Dreyfus και ο Haugeland στην κλασική ΤΝ είναι ριζικότερη από αυτή του θεωρητικού ολισμού γιατί αμφισβητούν ρητά και την παραδοχή (3). Η θέση που υποστηρίζουν, σε μια πρώτη προσέγγιση, μπορεί να διατυπωθεί ως εξής: υπάρχει ένα είδος γνώσης που δεν μπορεί να εκφραστεί μέσω συμβολικών αναπαραστάσεων και κανόνων, δηλαδή ως μια θεωρία του κόσμου. Πρόκειται για μια γνώση που δεν αφορά τα αντικείμενα του κόσμου, αλλά τον τρόπο που το υποκείμενο σχετίζεται με τον κόσμο. Είναι μια *πρακτική* γνώση που αφορά τρόπους αντίληψης και πράξης. Γι' αυτό και ο Dreyfus (1980) ονομάζει τη θέση του "πρακτικό ολισμό"¹⁹⁹.

Παραδείγματα τέτοιου είδους πρακτικών γνώσεων είναι η γνώση που έχει κανείς όταν κάνει ποδήλατο ή όταν μιλά τη μητρική του γλώσσα ή, για να χρησιμοποιήσουμε ένα πιο εξωτικό παράδειγμα, όταν έχει συνηθίσει να κινείται μέσα στον κόσμο φορώντας γυαλιά αντεστραμμένης οράσεως. Η γνώση αυτού του τύπου δεν εκφράζεται στη γλώσσα με

¹⁹⁷ Βασικός εκφραστής της μπορεί να θεωρηθεί ο Quine 1951.

¹⁹⁸ Με την ορολογία του κεφαλαίου 3 (ενότητα 3.3.1.4), αν δεχθούμε το θεωρητικό ολισμό, ο άνθρωπος λειτουργεί ως ένας *ερμηνευτής* του κόσμου.

¹⁹⁹ Πρβλ. Baltas 1987.

προτασιακούς όρους (γνωρίζω ότι p), αλλά με όρους δεξιοτήτων (skills). Για παράδειγμα, λέμε ότι γνωρίζω πώς (know how) να κάνω ποδήλατο, πώς να παίζω σκάκι, πώς να κινούμαι στο χώρο φορώντας γυαλιά αντεστραμμένης οράσεως. Γι' αυτό το λόγο, η πρακτική γνώση ονομάζεται, συχνά, "know-how"²⁰⁰ – έναν όρο που θα μεταφράσουμε κάπως ελεύθερα "πρακτική γνώση δεξιοτήτων".

Η πρακτική γνώση δεξιοτήτων είναι πολύ διαφορετική από την προτασιακή γνώση. Το πόσο διαφορετική είναι μπορεί να φανεί με το ακόλουθο παράδειγμα: ένας άνθρωπος μπορεί να γνωρίζει πλήρως όλους τους κανόνες του ποδοσφαίρου χωρίς να γνωρίζει να παίζει ποδόσφαιρο, αλλά και αντίστροφα, μπορεί να παίζει καταπληκτικό ποδόσφαιρο χωρίς να μπορεί να διατυπώσει ρητά όλους τους κανόνες του ποδοσφαίρου. Το ίδιο μπορεί να ισχύει και για τη γνώση μιας γλώσσας: για παράδειγμα, είναι δυνατόν κάποιος να γνωρίζει τέλεια τη γραμματική, το συντακτικό και το λεξιλόγιο μιας ξένης γλώσσας και παρόλα αυτά να μην μιλάει καλά τη γλώσσα. Αντίθετα, κάποιος που έχει την ίδια γλώσσα ως μητρική, παρότι μιλάει άπειρα, πολύ συχνά δεν μπορεί να διατυπώσει ρητά τους γραμματικούς και συντακτικούς κανόνες στους οποίους υπακούει η γλώσσα του²⁰¹. Ή, για να πάρουμε ένα ακόμα απλούστερο παράδειγμα, η γνώση που έχω όταν ξέρω ποδήλατο υπάρχει ανεξάρτητα από το αν εγώ μπορώ να την περιγράψω προτασιακά στην κάθε της λεπτομέρεια και, σίγουρα, η προτασιακή περιγραφή αποτελεί μια γνώση εντελώς διαφορετικού τύπου από τη γνώση που προσιδιάζει στην συγκεκριμένη δεξιότητα. Αν κάποιος άλλος ακούσει/διαβάσει και αποστηθίσει την προτασιακή αυτή περιγραφή, η γνώση που αποκομίζει είναι πολύ διαφορετικού είδους από τη δική μου, διότι, απλά, δεν θα αποκτήσει την ικανότητα να κάνει ποδήλατο.

²⁰⁰ Ο Dewey (1922, σελ. 177-178) εισάγει τον όρο ως εξής: "Μπορούμε να πούμε ότι 'γνωρίζουμε πώς' [know-how] μέσω των συνθησιών μας... Περιπατάμε και διαβάζουμε δυνατά, επιβιβάζομαστε και αποβιβάζομαστε από αυτοκίνητα, γυρνάμε και ξεντυνόμαστε, και κάνουμε ένα σωρό άλλες χρήσιμες πράξεις χωρίς να τις σκεπτόμαστε. Γνωρίζουμε κάτι, δηλαδή, πώς να το κάνουμε... Εάν διαλέξουμε να το αποκαλούμε αυτό γνώση... τότε άλλα πράγματα που επίσης αποκαλούνται γνώση, γνώση των και περί των πραγμάτων, γνώση ότι τα πράγματα είναι κατά τον ένα ή άλλο τρόπο, γνώση που ενέχει αναστοχασμό [reflection] και συνειδητή αποτίμηση, όλα αυτά τα πράγματα είναι μια γνώση διαφορετικού είδους".

²⁰¹ Πρβλ. Wittgenstein 1984, σελ. 54: "Στην πράξη πολύ σπάνια χρησιμοποιούμε τη γλώσσα σαν ένα τέτοιου είδους λογισμό [δηλ. έναν λογισμό που αναπτύσσεται σύμφωνα με ακριβείς κανόνες]. Διότι όχι μόνο δεν σκεπτόμαστε τους κανόνες χρήσης - τους ορισμούς, κλπ.- την ώρα που χρησιμοποιούμε τη γλώσσα, αλλά κι όταν μας ζητούν να δώσουμε αυτούς τους κανόνες, τις περισσότερες φορές, δεν είμαστε σε θέση να το κάνουμε. Δεν είμαστε σε θέση να καθορίσουμε με σαφήνεια τα όρια των εννοιών που χρησιμοποιούμε: όχι επειδή αγνοούμε τον πραγματικό τους ορισμό, αλλά επειδή δεν υπάρχει πραγματικός 'ορισμός' τους."

Ωστόσο, θα μπορούσε κανείς να αντιτείνει ότι όλα αυτά αφορούν την οπτική του πρώτου προσώπου, δηλαδή το πώς ο ίδιος ο ειδικός βιώνει και εκφράζει τη γνώση του. Με άλλα λόγια, η ένσταση είναι ότι τα προαναφερθέντα παραδείγματα αποτελούν μια πολύ καλή περιγραφή της φαινομενολογίας της εμπειρίας που έχει ένας ειδικός όταν ασκεί τη δεξιότητά του, αλλά δεν συνιστούν εξήγηση του πώς καταφέρνει και έχει μια τέτοια δεξιότητα. Εξάλλου, η συγκεκριμένη περιγραφή δεν αποκλείει ότι, σε ένα ασυνείδητο επίπεδο, ο ειδικός στηρίζεται και αυτός σε κάποια θεωρία -ένα σύνολο πεποιθήσεων και κανόνων -που του υπαγορεύει πώς να πράττει.

Ωστόσο, αυτό δεν είναι αληθές. Γιατί τα παραδείγματα που αναφέραμε δείχνουν ότι η εξαντλητική γνώση της θεωρίας ενός τομέα δεν αρκεί για να καταστήσει κάποιον ειδικό σε αυτόν τον τομέα. Με άλλα λόγια, η γνώση του ειδικού δεν εξαντλείται στη γνώση κάποιας θεωρίας, διότι, μεταξύ άλλων, καμιά θεωρία δεν μπορεί να προκαθορίσει την *επιδέξια* εφαρμογή της σε κάθε ενδεχόμενη κατάσταση²⁰². Θα πρέπει να αναρωτηθούμε τι αφαιρείται και τι προστίθεται όταν ακολουθούμε *επιδέξια* έναν κανόνα; Όταν πρωτομαθαίνουμε έναν κανόνα, τον εφαρμόζουμε βασιζόμενοι στην αναγνώριση κάποιων αντικειμενικών χαρακτηριστικών. Για παράδειγμα, μαθαίνουμε να βάζουμε δευτέρα στο αυτοκίνητο μόλις το ταχύμετρο φτάσει στα 20 χλμ. Με τον καιρό, όσο

²⁰² Ο Kant, για παράδειγμα, διακρίνει με σαφήνεια τη γνώση των κανόνων από τη γνώση της εφαρμογής των κανόνων. Αυτά τα δύο είδη γνώσης βασίζονται αντίστοιχα από δύο διαφορετικές ικανότητες: τη διάνοια και την κρίση. Στην *Κριτική του Καθαρού Λόγου* επισημαίνει χαρακτηριστικά: "Γι' αυτό είναι δυνατό ένας γιατρός, ένας δικαστής ή ένας πολιτικός να έχει το κεφάλι του γεμάτο από ωραίους κανόνες της παθολογίας, της νομικής και της πολιτικής σε τέτοιο βαθμό ώστε να μπορεί να γίνει ένας εμβριθής δάσκαλος στο κεφάλαιο αυτό, και όμως εύκολα να σφάλει στην εφαρμογή τους είτε επειδή του λείπει η φυσική κριτική ικανότητα (χωρίς ωστόσο να του λείπει νους), ώστε να βλέπει καθαρά το γενικό στην αφηρημένη του μορφή αλλά να μην μπορεί να διακρίνει αν μια περίπτωση στην συγκεκριμένη της μορφή υπάγεται σε αυτό [το γενικό] είτε ακόμα επειδή δεν έχει εξασκηθεί αρκετά στην κρίση αυτή με παραδείγματα και πραγματικά περιστατικά. Αυτή είναι και η μόνη και μεγάλη ωφέλεια των παραδειγμάτων, ότι δηλ. οξύνουν την κριτική δύναμη" (A134/B173). Πρβλ Wittgenstein (1953/1977, σελ. 282), ο οποίος αναφερόμενος στο εάν και πώς μπορούμε να μάθουμε να διακρίνουμε αν κάποιος εκδηλώνει γνήσια ένα συναίσθημα, επισημαίνει: "Μπορεί κανείς να μάθει να γνωρίζει τους ανθρώπους; Ναι, μερικοί μπορούν να το μάθουν. Όχι όμως από μια σειρά παραδόσεων (courses), αλλά χάρη στην 'εμπειρία'. - Μπορεί ένας άνθρωπος να το διδάξει σ' άλλον; Ασφαλώς. Κάθε τόσο του δίνει τη σωστή ένδειξη (tip). Εδώ έτσι φαίνονται τα 'μαθαίνω' και 'διδάσκω' (This is what 'learning' and 'teaching' are like here). - Εκείνο που μαθαίνει κανείς δεν είναι μια τεχνική: μαθαίνουμε ορθές κρίσεις. Και υπάρχουν και κανόνες, αλλά δεν αποτελούν σύστημα, και μόνο αυτοί που έχουν πείρα τους εφαρμόζουν σωστά. Διαφορετικά απ' ό,τι συμβαίνει με τους κανόνες υπολογισμού".

εξοικειωνόμαστε με την εφαρμογή του κανόνα, μαθαίνουμε να αλλάζουμε ταχύτητα βασιζόμενοι σε χαρακτηριστικά που εξαρτώνται από το εκάστοτε πλαίσιο, όπως για παράδειγμα, τον ήχο της μηχανής, την κλίση του δρόμου, την ολισθηρότητα του δρόμου, την επιθυμία μας να φρενάρουμε κ.λπ. Τέλος, όταν πλέον καταστούμε ειδικοί στην οδήγηση, μπορούμε να αλλάζουμε ταχύτητα εντελώς ασυνείδητα²⁰³. Για παράδειγμα, ένας έμπειρος οδηγός που τυγχάνει να είναι και έμπειρος συζητητής της φιλοσοφίας μπορεί να συμμετέχει σε μια φιλοσοφική κουβέντα χωρίς αυτό να επηρεάζει καθόλου την ικανότητά του στην οδήγηση.

Βλέπουμε, λοιπόν, ότι όταν μάθουμε να ακολουθούμε επιδέξια έναν κανόνα, τότε αυτός δεν μεταπίπτει απλά από το συνειδητό στο ασυνείδητο χωρίς καμιά ουσιαστική αλλαγή στη δομή του. Χαρακτηριστικό εδώ είναι το σχόλιο του Dreyfus (1992, σελ. xiii) στην επισήμανση του Feigenbaum ότι μόλις μάθουμε τον κανόνα δεσμάτος των παπουτσιών απλά ακολουθούμε αυτόν τον ίδιο κανόνα ασυνείδητα: "Αυτό θα ήταν σαν να ισχυριζόμασταν ότι εφόσον χρειαστήκαμε βοηθητικές ρόδες όταν μαθαίναμε ποδήλατο, πρέπει και τώρα να χρησιμοποιούμε αόρατες βοηθητικές ρόδες κάθε φορά που κάνουμε ποδήλατο"²⁰⁴. Αυτό που μάλλον συμβαίνει είναι ότι η δομή του κανόνα τροποποιείται ριζικά και καθίσταται, όπως θα έλεγε ο Dreyfus, ένα μυϊκό gestalt²⁰⁵. Έτσι,

²⁰³ Για μια αναλυτικότερη περιγραφή των σταδίων εκμάθησης ενός κανόνα μέχρι την επιδέξια εφαρμογή του, βλέπε Dreyfus & Dreyfus 1986. Οι αδελφοί Dreyfus παρουσιάζουν ένα φαινομενολογικό μοντέλο για το πώς αναπτύσσει ένας ενήλικας νέες δεξιότητες, δηλαδή το πώς από αρχάριος καθίσταται σταδιακά ειδικός. Το συγκεκριμένο μοντέλο μπορεί να μας βοηθήσει να κατανοήσουμε καλύτερα το πρόβλημα της κοινότητας γνώσης - αν σκεφτούμε ότι όλοι είμαστε ειδικοί του κοινού νου. Επιστημολογικά είναι ευκολότερο να μελετηθεί το πώς και τι γνωρίζει ένας ειδικός σε κάποιο επιμέρους πεδίο α, γιατί μόνο λίγοι από εμάς κατορθώνουν να αποκτήσουν μια δεξιότητα στο συγκεκριμένο πεδίο. Έτσι, μπορούμε να εντοπίσουμε και να εκτιμήσουμε καλύτερα όλα όσα γνωρίζει ο ειδικός, γιατί οι γνώσεις του δεν αποτελούν αυτονόητα για 'μας. Αντίθετα, δεδομένου ότι όλοι (ή σχεδόν όλοι) είμαστε ειδικοί του κοινού νου είναι πολύ δύσκολο να μελετήσουμε όλα όσα γνωρίζουμε αλλά ποτέ δεν θεματοποιούμε, ακριβώς επειδή αποτελούν κοινό τόπο για όλους μας.

²⁰⁴ Ακόμα πιο χαρακτηριστικό είναι το παράδειγμα της κολύμβησης: επειδή χρειαστήκαμε "μπρατσάκια" όταν μαθαίναμε κολύμβηση, χρησιμοποιούμε και τώρα αόρατα "μπρατσάκια" όταν κολυμπάμε.

²⁰⁵ Γενικά, κατά την απόκτηση μιας δεξιότητας - για παράδειγμα, την εκμάθηση οδήγησης, χορού ή της προφοράς μιας ξένης γλώσσας - στην αρχή πρέπει να ακολουθούμε τους κανόνες σιγά, αδέξια και συνειδητά. Αλλά αργότερα έρχεται κάποια στιγμή όπου τελικά μπορούμε να επιτελέσουμε το έργο αυτόματα. Σε αυτό το σημείο δεν φαίνεται να μεταθέτουμε απλώς αυτούς του ίδιους άκαμπτους κανόνες στο ασυνείδητο· αλλά μάλλον φαίνεται να αποκτούμε το μυϊκό gestalt

μετασχηματίζεται σε μια ολιστική γνώση που δεν είναι περαιτέρω αναλύσιμη για τον ίδιο τον κάτοχο της γνώσης²⁰⁶. Για παράδειγμα, όταν βάζουμε δευτέρα, όντας έμπειροι οδηγοί, δεν ακολουθούμε τον κανόνα "όταν το ταχύμετρο δείξει 20 χλμ/ώρα, βάλε δευτέρα", αλλά ανταποκρινόμαστε στις εκάστοτε διαθεσιμότητες²⁰⁷ (affordances) και αντιστάσεις του δρόμου. Δεν υπάρχει καμιά απόσταση μεταξύ του υποτιθέμενου κανόνα και του τρόπου εφαρμογής του, δηλαδή κάποιου είδους ερμηνεία. Ο έμπειρος οδηγός δεν λειτουργεί ως ένας ερμηνευτής, αλλά ως ένας *επιδέξιος χρήστης*²⁰⁸. Παράλληλα, το λεβιέ των ταχυτήτων παύει να βρίσκεται στο επίκεντρο της προσοχής και καθίσταται ένα εργαλείο που χρησιμοποιείται "τυφλά". Γενικότερα, η δραστηριότητα της αλλαγής ταχύτητας ενσωματώνεται με έναν επιδέξιο τρόπο, δηλαδή με ομαλότητα και χάρη, στη συνολικότερη δραστηριότητα της οδήγησης. Αν δεν παρουσιαστεί κάποιο απρόοπτο, ο οδηγός βιώνει τη δραστηριότητα της οδήγησης ως ενιαία και αδιαίρετη²⁰⁹. Ο αρχικός κανόνας ("όταν το ταχύμετρο δείξει 20 χλμ/ώρα, βάλε δευτέρα") δεν έχει απλά ξεχαστεί και ακολουθείται ασυνείδητα, αλλά δεν χρησιμοποιείται καθόλου. Ο αρχικός κανόνας μοιάζει με τις βοηθητικές ρόδες που πετιούνται μόλις κανείς μάθει ποδήλατο²¹⁰. Αυτό που απομένει είναι μια δεξιότητα που δεν μπορεί να διατυπωθεί υπό μορφή προτασιακού κανόνα όχι για κανέναν

που δίνει στη συμπεριφορά μας μια νέα ευκαμψία και ομαλότητα" Dreyfus 1992, σελ. 248-249.

²⁰⁶ Ο μυστηριακός τόνος αυτής της διαπίστωσης θα αρθεί, ως ένα βαθμό, στο έκτο κεφάλαιο όπου θα εξετάσουμε τις κατανεμημένες αναπαραστάσεις.

²⁰⁷ Πρόκειται για έναν τεχνικό όρο του Gibson (1996, σελ. 127): "Οι *διαθεσιμότητες* (affordances) του περιβάλλοντος είναι αυτό που το περιβάλλον προσφέρει στο ζώο, που προμηθεύει ή εφοδιάζει, για καλό ή για κακό. Το ρήμα *to afford* υπάρχει στο λεξικό, ενώ το ουσιαστικό *affordance* δεν υπάρχει. Το επινόησα. Με αυτό εννοώ κάτι που αναφέρεται τόσο στο περιβάλλον όσο και στο ζώο με ένα τρόπο που κανείς από τους υπάρχοντες όρους δεν το κάνει. Ο όρος υπονοεί την συμπληρωματικότητα του ζώου και του περιβάλλοντος".

²⁰⁸ Πρβλ. υποσημείωση 198 και την έννοια του χρήστη αναπαραστάσεων όπως την εισαγάγαμε στο τρίτο κεφάλαιο κατ' ανδιαστολή προς την έννοια του ερμηνευτή αναπαραστάσεων. Σύμφωνα με όσα είπαμε παραπάνω, η σχέση που έχουμε με τη μητρική μας γλώσσα είναι σχέση *χρήσης* της γλώσσας, ενώ όταν μαθαίνουμε μια νέα γλώσσα τότε η σχέση που έχουμε με αυτή είναι σχέση *ερμηνείας*. Προκειμένου να φτάσουμε να έχουμε σχέση ερμηνείας με μια γλώσσα πρέπει προηγουμένως να έχουμε σχέση *χρήσης* με μια άλλη γλώσσα (ήτοι τη μητρική γλώσσα). Η ερμηνεία προϋποθέτει ήδη ότι βρισκόμαστε στο χώρο του νοήματος, δηλαδή ότι χρησιμοποιούμε *επιδέξια* μια γλώσσα. Υπ' αυτή την έννοια, όταν μαθαίνουμε τη μητρική μας γλώσσα δεν λειτουργούμε ούτε ως χρήστες ούτε ως ερμηνευτές.

²⁰⁹ Επίσης, αν ο οδηγός εστιάσει την προσοχή του στο πώς ακριβώς και πότε αλλάζει ταχύτητα, τότε η συμπεριφορά του γίνεται λιγότερο επιδέξια.

²¹⁰ Ή με τις σκαλωσιές, για να θυμηθούμε ένα οικοδομικό παράδειγμα με φιλοσοφικές προεκτάσεις.

μυστηριακό λόγο, αλλά απλά λόγω της δομής των προτασιακών κανόνων. Ένας προτασιακός κανόνας έχει, συνήθως, τη δομή "εάν X, τότε Y". Το X είναι το σύνολο των συνθηκών που πρέπει να πληρούνται προκειμένου να εκτελεστεί το Y. Οι εν λόγω συνθήκες μπορούν να εκφραστούν προτασιακά μόνο αν είναι ανεξάρτητες από το εκάστοτε πλαίσιο χρήσης, την εκάστοτε κατάσταση. Για παράδειγμα, η συνθήκη "το ταχύμετρο δείχνει 20 χλμ/ώρα" δεν εξαρτάται από τις τυχαιότητες της εκάστοτε κατάστασης, όπως την κλίση του δρόμου, τον καιρό, τη διάθεση του οδηγού κ.λπ. Η εκάστοτε κατάσταση εντός της οποίας καλούμαστε να εφαρμόσουμε έναν κανόνα δεν είναι δυνατόν να προβλεφτεί εκ των προτέρων ούτε να διατυπωθεί εξαντλητικά με αντικειμενικούς όρους. Γι' αυτό και δεν είναι δυνατόν να διατυπωθούν προτασιακοί κανόνες που να προκαθορίζουν τον *απόδο* τρόπο εφαρμογής τους στην εκάστοτε συγκυρία. Όπως φάνηκε και από την εξέταση του προβλήματος του πλαισίου και του προβλήματος της κοινότητας γνώσης, ο άνθρωπος έχει την ικανότητα να καταλαβαίνει άμεσα τι είναι σημαντικό στην εκάστοτε κατάσταση και να προσανατολίζεται και να ενεργεί κατάλληλα χωρίς να βασίζεται σε προτασιακούς κανόνες. Αν δεν είχε αυτή την ικανότητα, δεν θα μπορούσε ποτέ να δρα *επίδο* μέσα στον κόσμο, δηλαδή να ανταποκρίνεται με χάρη στις τυχαιότητες που του παρουσιάζονται.

Όπως είδαμε, στην περίπτωση του έμπειρου οδηγού που αλλάζει ταχύτητα δεν λαμβάνει χώρα κάποιου είδους ερμηνεία, δηλαδή δεν υπάρχει απόσταση μεταξύ του κανόνα προς ακολούθηση και του τρόπου εφαρμογής του. Αυτό το γεγονός, ωστόσο, δεν σημαίνει ότι η αλλαγή ταχύτητας έχει το χαρακτήρα εξαρτημένου αντανakλαστικού. Η αλλαγή ταχύτητας συνοδεύεται από κάποιου είδους εμπειρία, υπό την έννοια ότι ο οδηγός έχει μια συναίσθηση της ικανότητας του πράττειν (sense of agency). Με άλλα λόγια, έχει συναίσθηση ότι η αλλαγή ταχύτητας αποτελεί δικιά του δραστηριότητα που ως τέτοια συναρμόζει με τους τρέχοντες στόχους του και περιορίζεται από τις γενικότερες δεσμεύσεις του. Γι' αυτό και ο έμπειρος οδηγός μπορεί να διακρίνει το εάν δεν επιτελεί *επίδο* την αλλαγή ταχύτητας, δηλαδή μπορεί να κάνει κανονιστικές διακρίσεις.

Η πρακτική γνώση δεξιοτήτων είναι ουσιαστικά ένα είδος γνώσης που δεν μπορεί να διακριθεί από τον χρήση της και να περιγραφεί μέσω αναπαραστάσεων που θα ίσχυαν ανεξάρτητα από αυτόν. Γι' αυτό και το know-how είναι τόσο δύσκολο να μεταφερθεί από τον έναν άνθρωπο στον άλλο. Σε αντίθεση με την προτασιακή γνώση που μεταφέρεται με το γραπτό ή τον προφορικό λόγο, το know-how μπορεί να μεταφερθεί μόνο μέσα από μια επίπονη διαδικασία ανάπτυξης κάποιας σωματικής ή πνευματικής δεξιότητας. Μια τέτοια διαδικασία προϋποθέτει χιλιάδες επαναλήψεις μέσα από την αντιμετώπιση χιλιάδων επιμέρους περιπτώσεων. Γι' αυτό και η χρήση παραδειγμάτων και το παιχνίδι είναι οι

βασικοί τρόποι με τους οποίους μπορούμε να αναπτύξουμε μια σωματική ή πνευματική δεξιότητα²¹¹. Το know-how είναι μια ουσιαδώς ενσώματη (embodied) γνώση. Επίσης, δεδομένου ότι το know-how είναι μια γνώση που έχει να κάνει με τις ικανότητες δράσης και αντίληψης του υποκειμένου, δεν αφορά άμεσα κάποιες οντότητες του εξωτερικού κόσμου, αλλά τον τρόπο που το υποκείμενο σχετίζεται πρωταρχικά με το σώμα του²¹² και μέσω αυτού με τον φυσικό και κοινωνικό κόσμο²¹³. Γι' αυτό και η συγκεκριμένη γνώση δεν ανα-παριστά κάτι από τον εξωτερικό κόσμο και ούτε μπορεί να αναλυθεί με όρους αναφοράς, αλλά αναλύεται με όρους επιδέξιας δραστηριότητας (skillful activity)²¹⁴. Όπως επισημαίνει ο Cussins (2000), ο τρόπος που μου παρουσιάζεται ο κόσμος όταν βασιζομαι στην πρακτική γνώση δεξιοτήτων είναι ως ένα σύνολο διαθεσιμοτήτων και αντιστάσεων, ως ένα σύνολο ιχνών (trails) που με καθοδηγούν στη δραστηριότητά μου. Η ανάλυση του Cussins, ωστόσο, θέτει μεγάλα ζητήματα που υπερβαίνουν το πλαίσιο της παρούσας διατριβής και γι' αυτό δεν θα την εξετάσουμε περαιτέρω²¹⁵.

Εφόσον για να εφαρμοστεί επιδέξια μια θεωρία και οι κανόνες της στην εκάστοτε κατάσταση απαιτείται η γνώση της εφαρμογής των κανόνων, δηλαδή μια πρακτική γνώση δεξιοτήτων, η τελευταία αναδεικνύεται, ουσιαστικά, αναγκαία προϋπόθεση για την επιδέξια χρήση της προτασιακής γνώσης. Αν η μόνη γνώση που διαθέταμε ήταν προτασιακή, τότε στην καλύτερη περίπτωση θα συμπεριφερόμασταν ως

²¹¹ Βλ. υποσημ. 202.

²¹² Στο όγδοο κεφάλαιο (ενότητα 8.2) θα αναφερθούμε αναλυτικότερα σε μια έκφανση αυτής της σχέσης εξετάζοντας την συναίσθηση της ικανότητας του πράττειν (sense of agency).

²¹³ Βλ. Dreyfus 1996, σελ. 183. Βλ. επίσης Merleau-Ponty (1962, σελ. 146): "Το σώμα είναι το γενικό μας μέσο για να έχουμε έναν κόσμο. Μερικές φορές περιορίζεται στις πράξεις που είναι αναγκαίες για τη διατήρηση της ζωής, και συστοιχά θέτει γύρω μας έναν βιολογικό κόσμο· άλλες φορές, επεξεργαζόμενο αυτές τις πρωταρχικές πράξεις και κινούμενο από το κυριολεκτικό στο μεταφορικό τους νόημα, εκδηλώνει μέσω αυτών ένα πυρήνα νέας σημασίας: αυτό είναι αληθές για κινητικές συνήθειες όπως ο χορός. Μερικές φορές, τέλος, το νόημα που στοχεύεται δεν μπορεί να επιτευχθεί με τα φυσικά μέσα του σώματος· πρέπει τότε να κατασκευάσει για τον εαυτό του ένα όργανο, και να προβάλλει ως εκ τούτου γύρω του έναν πολιτισμικό κόσμο".

²¹⁴ Βλ. Cussins 1990, Cussins 1992 και Cussins 2000.

²¹⁵ Βλ., επίσης, Haugeland 1995. Ο Dreyfus (2001), βασιζόμενος στον Merleau-Ponty (1962), επιχειρεί να αναλύσει αυτό το διαφορετικό είδος αποβλεπτικότητας - που το ονομάζει "επιδέξια ανταπόκριση" (skillful coping) - αναφερόμενος σε μια δραστηριότητα κατά την οποία το υποκείμενο άγεται προς μια κατάσταση ισορροπίας (equilibrium) χωρίς να επιδιώκει την επίτευξη ενός προκαθορισμένου στόχου. Ο Dreyfus θεωρεί ότι αυτό που ονομάζει "επιδέξια ανταπόκριση" συγγενεύει με ό,τι ο Wittgenstein ονομάζει "εύρεση του δρόμου μου μέσα στον κόσμο".

αδέξια ρομπότ που έχουν αφεθεί να λειτουργήσουν μέσα στον πραγματικό κόσμο. Εκφραζόμενοι με ένα λιγότερο οικείο λεξιλόγιο θα μπορούσαμε να πούμε ότι η θεματική κατανόηση του κόσμου, δηλαδή η θεωρητική γνώση του κόσμου – η γνώση αντικειμενικών γεγονότων- στηρίζεται σε μια μη θεματική, μη προτασιακή κατανόηση, σε μια προκατανόηση²¹⁶. Για αναλυτικούς και μόνο λόγους, θα μπορούσαμε να διακρίνουμε αυτήν την προκατανόηση σε επιμέρους είδη. Δύο σημαντικότερα τέτοια είδη προκατανόησης (ή μη-προτασιακής γνώσης) είναι η προκατανόηση του σώματός μας και η προκατανόηση των κοινωνικών σχέσεων.

4.3.3 Μη-προτασιακή γνώση που αφορά τη δομή και τη λειτουργία του σώματος

Υπάρχει ένα τεράστιο σώμα "γνώσεων" που το διαθέτουμε ακριβώς επειδή έχουμε τη συγκεκριμένη δομή σώματος με τις συγκεκριμένες λειτουργίες. Έτσι, κάθε στιγμή έχουμε μια προκατανόηση του σώματός μας και των δυνατοτήτων του. Επειδή, βέβαια, αυτή η προκατανόηση πάντα ήδη προϋποτίθεται σε ό,τι κάνουμε, τείνουμε να μην την προσέχουμε καν. Ένας τρόπος να αναδειχθεί το πόσο η προκατανόηση του σώματός μας επηρεάζει τον τρόπο που έχουμε δομήσει τον κόσμο μας, είναι να προσπαθήσουμε προς στιγμήν να σκεφτούμε πώς θα κατανοούσαμε τον φυσικό και κοινωνικό κόσμο αν είχαμε ένα ριζικά διαφορετικό είδος σώματος. Για παράδειγμα, πώς θα ήταν ο κόσμος αν το σώμα μας ήταν υγρό ή αέριο; Θα υπήρχαν χαρταετοί ή καρέκλες σε έναν τέτοιο κόσμο και πώς θα ήταν τα εστιατόρια; Έχοντας ανθρώπινο σώμα γνωρίζουμε με έναν άρρητο τρόπο πάρα πολλά πράγματα για το πώς να αλληλεπιδρούμε με τον κόσμο και τους άλλους. Αυτή την άρρητη γνώση την χρησιμοποιούμε αδιαλείπτως για να κατανοούμε τις πράξεις των συνανθρώπων μας. Για παράδειγμα, αν διαβάσουμε στην εφημερίδα ότι κάποιος αθλητής βραβεύτηκε γιατί κατάφερε να αναρριχηθεί χωρίς τη βοήθεια εργαλείων σε μια λεία κατακόρυφη επιφάνεια ύψους δέκα μέτρων, θα καταλάβουμε αμέσως γιατί πήρε το βραβείο. Όλοι μας γνωρίζουμε πόσο δύσκολο είναι να αναρριχηθεί κανείς σε μια λεία κατακόρυφη επιφάνεια. Αν είχαμε σώμα κατσαρίδας, μια τέτοια πράξη, σίγουρα, δεν θα ήταν τόσο δύσκολη. Όλα αυτά που γνωρίζουμε περί του σώματός μας και των δυνατοτήτων του δεν χρεώστηκε ποτέ να τα διδαχτούμε στο σχολείο γιατί τα ξέρουμε άμεσα επειδή έχουμε το σώμα που έχουμε.

Γενικεύοντας, λοιπόν, μπορούμε να πούμε ότι ο τρόπος που κατανοούμε τον κόσμο εντός του οποίου ζούμε συγκροτείται σε έναν σημαντικό βαθμό από τον τρόπο που κατανοούμε το σώμα μας. Αν σκεφτούμε πόσο ριζικά διαφορετικό είδος σώματος και, κυρίως, πόσο ριζικά διαφορετικούς τρόπους "αίσθησης" του σώματός τους έχουν οι

²¹⁶ Βλ. Heidegger 1927.

υπολογιστές, αναδεικνύεται με προφάνεια το πόσο απίθανο είναι να μπορέσουν κάποτε να κατανοήσουν τον κόσμο όπως τον κατανοούν οι άνθρωποι.

4.3.4 Μη-προτασιακή γνώση που προέρχεται από την ένταξη σε ένα κοινωνικό περιβάλλον

Εκτός όμως από τη σιωπηρή (tacit) γνώση της σωματικότητάς μας, υπάρχει μια εξίσου σημαντική σιωπηρή γνώση που σχετίζεται με την ικανότητά μας να είμαστε κοινωνικά όντα. Ένα παράδειγμα τέτοιας "γνώσης" αποτελεί η ασυναίσθητη ρύθμιση της απόστασης από την οποία κανείς μιλά με άλλους ανθρώπους ανάλογα με τη σχέση που έχει μαζί τους²¹⁷. Η "γνώση" αυτή είναι πολιτισμικά καθοριζόμενη γιατί οι αρμόζουσες κάθε φορά αποστάσεις ποικίλουν μεταξύ διαφορετικών κοινοτήτων. Αρκεί να συγκρίνουμε την απόσταση από την οποία είθισται να συνομιλούν δύο Ταϊλανδοί με εκείνη από την οποία συνομιλούν δύο Σουηδοί. Σε κάθε περίπτωση πάντως, η γνώση της αρμόζουσας κάθε φορά απόστασης δεν είναι κάτι που μαθαίνεται από τα βιβλία, ούτε χρειάζεται να διδάσκεται υπό την μορφή κανόνων. Η γνώση της αρμόζουσας απόστασης ομιλίας αναπτύσσεται σιγά-σιγά καθώς το παιδί μεγαλώνοντας σε μια συγκεκριμένη κοινωνία παρατηρεί χιλιάδες επιμέρους περιπτώσεις τις οποίες διαρκώς μιμείται στο παιχνίδι ή στην πραγματική ζωή.

Επίσης, ένα πολύ μεγάλο σώμα γνώσεων μπορεί να χρησιμοποιείται χωρίς καν να το κατέχουμε!²¹⁸ Τέτοια γνώση είναι η γνώση που έχει "αποθηκευτεί" στα εργαλεία και τους θεσμούς που έχουμε αναπτύξει ως κοινωνικά όντα. Για παράδειγμα, η περίπλοκη συμπεριφορά ενός ανθρώπου μέσα σε μια τράπεζα δεν οφείλεται τόσο στην ύπαρξη μιας σύνθετης αναπαράστασης των τραπεζικών σχέσεων εντός του, όσο στην πολυπλοκότητα του ίδιου του τραπεζικού συστήματος -μια πολυπλοκότητα η οποία έχει "εγγραφεί" και συσσωρευτεί στα ίδια τα εργαλεία που έχουν αναπτυχθεί, σταδιακά ανά τους αιώνες, για την υποστήριξη αυτού του θεσμού²¹⁹. Το παράδειγμα αυτό δείχνει, λοιπόν, ότι ένα μεγάλο μέρος της γνώσης δεν είναι αναγκαίο να αποθηκεύεται εντός του οργανισμού. Το ίδιο το εξωτερικό περιβάλλον και ειδικότερα τα εργαλεία μπορούν πολύ αποτελεσματικά να χρησιμοποιηθούν για αυτόν τον σκοπό. Για την ακρίβεια, η εν λόγω γνώση δεν βρίσκεται

²¹⁷ Βλ. Dreyfus 1991, σελ. 18.

²¹⁸ Βλ. Putnam 1975, Haugeland 1995, Clark & Chalmers 1998.

²¹⁹ Ένα αντίστοιχο παράδειγμα που αναδεικνύει τη μη προτασιακή γνώση που είναι "αποθηκευμένη" στο φυσικό περιβάλλον αναφέρεται από τον Simon (1969, σελ. 63-64): η πολύπλοκη διαδρομή που ακολουθεί ένα μυρμήγκι πάνω στην αμμουδιά δεν οφείλεται στην ύπαρξη μιας πολύπλοκης αναπαράστασης της αμμουδιάς εντός του εντόμου αλλά στη ζεύξη ενός απλού οργανισμού με ένα πολύπλοκο περιβάλλον.

αποθηκευμένη ούτε στα εργαλεία από μόνα τους. Πώς θα μπορούσε, άλλωστε, ένα εργαλείο να έχει οποιαδήποτε γνώση. Αυτό που ισχύει μάλλον είναι ότι η συγκεκριμένη γνώση "ενεργοποιείται" κάθε φορά που χρησιμοποιείται το εργαλείο. Έτσι, ο άνθρωπος αν γνωρίζει αλλά πώς να χρησιμοποιεί ένα εργαλείο - δηλαδή κατέχει το απαραίτητο know-how - αποκτά ταυτόχρονα και την ικανότητα να χρησιμοποιεί μια γνώση που δεν κατέχει. Υπ' αυτή την έννοια, τα εργαλεία, όπως και οι εξωτερικές αναπαραστάσεις, φέρουν ένα είδος περιεχομένου, με τη διαφορά ότι σε αντίθεση με τις δεύτερες το εν λόγω περιεχόμενο δεν είναι αναπαραστασιακό -όπου αυτό, σε μια πρώτη προσέγγιση, σημαίνει ότι δεν καθίσταται φανερό²²⁰. Όπως επισημαίνει και ο Haugeland (1995, σελ. 233) αναφερόμενος σε ένα πασίγνωστο παράδειγμα του Heidegger: "Ένα σφυρί, για παράδειγμα, είναι σημαντικό (significant) πέρα από το ίδιο από την άποψη τού τι εξυπηρετεί: για το κάρφωμα των ξύλων, μέσω της επιδέξιας χρήσης του κατά ένα συγκεκριμένο τρόπο, προκειμένου να χτισθεί κάτι κ.ο.κ. Τα καρφιά, το ξύλο, το έργο και ο ξυλουργός είναι όμοια εμπλεγμένοι σε αυτό το "πλέγμα σημαντικότητας" (web of significance), με τους αντίστοιχους τρόπους τους. Αυτά είναι τα αντικείμενα με νόημα (meaningful objects) που απαρτίζουν τον ίδιο τον κόσμο και κανένα από αυτά δεν είναι μια αναπαράσταση"²²¹.

4.4 Επίλογος

Τα προβλήματα του πλαισίου και της κοινότοπης γνώσης δεν εγείρονται καν ως προβλήματα στους ανθρώπους, γι' αυτό και δεν τίθεται ζήτημα επίλυσής τους. Οι άνθρωποι γνωρίζουν τον κόσμο όντας πάντα σε μια κατάσταση. Κατ' αυτό τον τρόπο, γνωρίζουν άμεσα κάθε φορά το τι είναι σημαντικό και τι αξίζει προσοχής. Οι άνθρωποι βρίσκονται ήδη πάντα σε μια κατάσταση γιατί, μεταξύ άλλων, ήδη πάντα έχουν ένα σώμα που εμπλέκεται κάθε φορά με ένα συγκεκριμένο τρόπο στον κόσμο, ήδη πάντα βρίσκονται σε κάποια διάθεση και ήδη πάντα έχουν μια προκατανόηση των κοινωνικών σχέσεων και, ειδικότερα, της δικής τους κοινωνικής θέσης στην εκάστοτε συγκυρία.

Στο κεφάλαιο αυτό δείξαμε ότι ακόμα κι αν δεχθούμε προσωρινά ότι οι συμβολικές αναπαραστάσεις επαρκούν για να εκφράσουν την προτασιακή γνώση, υπάρχει ένα άλλο είδος γνώσης που δεν είναι προτασιακό και δεν μπορεί να συλληφθεί με τη βοήθεια συμβολικών αναπαραστάσεων και κανόνων χειρισμού αυτών.

²²⁰ Το συγκεκριμένο ζήτημα, ωστόσο, υπερβαίνει το πλαίσιο της διατριβής και γι' αυτό δεν θα ασχοληθούμε περαιτέρω μαζί του. Για μια πρώτη ανάλυση του ζητήματος βλ. Heidegger 1927, §17.

²²¹ Θα επανέλθουμε σε αυτό το ζήτημα στο τελευταίο κεφάλαιο της διατριβής για να δείξουμε πώς μπορεί να χρησιμοποιηθεί σε μια μη αναπαραστασιακή προσέγγιση των νοητικών φαινομένων.

Στο έκτο κεφάλαιο θα διερευνήσουμε αν η πρακτική γνώση δεξιοτήτων μπορεί να συλληφθεί με ένα νέο είδος αναπαραστάσεων, τις επονομαζόμενες "κατανεμημένες αναπαραστάσεις". Πριν από αυτό, ωστόσο, για λόγους ιστορικής προτεραιότητας θα ασχοληθούμε με ένα άλλο είδος αναπαραστάσεων που προτείνεται από ορισμένους γνωστικούς επιστήμονες. Πρόκειται για τις εικονιστικές αναπαραστάσεις.

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ

Νοητικά Είδηλα

5.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο εξετάσαμε μια βασική κριτική κατά της ΑΥΘΝ η οποία έχει ως στόχο να δείξει τα όρια αυτού του εξηγητικού πλαισίου. Πιο συγκεκριμένα, είδαμε ότι υπάρχουν μη-προτασιακά είδη γνώσης που σχετίζονται με την κοινότητα και την περιστασιακή γνώση τα οποία δεν μπορούν να εκφραστούν με συμβολικές αναπαραστάσεις. Έτσι, δείξαμε ότι η ΑΥΘΝ δεν επαρκεί για την εξήγηση των συγκεκριμένων ειδών μη-προτασιακής γνώσης.

Στο παρόν κεφάλαιο θα εξετάσουμε μια απόπειρα να εξηγηθεί με τη βοήθεια αναπαραστάσεων ένα άλλο είδος γνώσης που ενέχει μη-προτασιακά στοιχεία, ήτοι τη γνώση που κατακτούμε μέσω της φαντασίας και μέσω της αντίληψης. Στη θέση των νοητικών συμβόλων αυτή τη φορά προτείνεται ένα νέο είδος αναπαραστάσεων: τα νοητικά είδωλα (mental images)²²². Η θεωρία περί νοητικών ειδώλων, παρότι προτάθηκε ως αντίπαλη προς τη θεωρία περί νοητικών συμβόλων, τελικά παραμένει μέσα στο γενικότερο πλαίσιο που θέτει η ΑΥΘΝ και απλώς αποτελεί μια προσπάθεια επέκτασης του συγκεκριμένου εξηγητικού πλαισίου. Αυτό ισχύει γιατί η θεωρία περί νοητικών ειδώλων δέχεται τη συστατική παραδοχή της ΑΥΘΝ ότι δηλαδή η εξήγηση των νοητικών φαινομένων που παράγονται στη φαντασία δεν είναι τίποτε άλλο παρά κατοχή και χειρισμός αναπαραστάσεων εικονιστικής μορφής.

Από μεθοδολογική άποψη, σε αυτό το κεφάλαιο δεν θα θέσουμε υπό αμφισβήτηση το ότι τα νοητικά είδωλα αποτελούν όντως ένα είδος αναπαραστάσεων, δηλαδή έχουν αναπαραστασιακό περιεχόμενο. Με άλλα λόγια, δεν θα εξετάσουμε το ερώτημα "δυνάμει τίνος πράγματος αποκτά μια εσωτερική κατάσταση του εγκεφάλου αναπαραστασιακό περιεχόμενο;". Αυτό το ερώτημα θα εξεταστεί στο έβδομο κεφάλαιο. Στο παρόν κεφάλαιο θα υποθέσουμε προσωρινά ότι με κάποιο τρόπο τα νοητικά είδωλα αποκτούν περιεχόμενο και θα εξετάσουμε από κει και πέρα τι προβλήματα δημιουργούνται.

²²² Δεν υιοθετούμε τον όρο "νοητική εικόνα" ως μετάφραση του "mental image" γιατί κατ' αυτόν τον τρόπο (α) χάνεται η διάκριση μεταξύ image και picture; (ο όρος "image" αναφέρεται σε ένα προϊόν της φαντασίας, ενώ ο όρος "picture" σε μια εξωτερική εικόνα) και (β) προαποφασίζεται χωρίς έρευνα ότι τα mental images είναι μια μορφή εικόνων. Επιλέξαμε την απόδοση "νοητικό είδωλο" και όχι τον αρχαιοελληνικό όρο "φάντασμα" που είναι ακριβώς ό,τι οι Λατίνοι μετέφρασαν ως "image", λόγω της πολύ διαφορετικής σημασίας που έχει αποκτήσει ο όρος "φάντασμα" στα νέα Ελληνικά. Ο όρος "είδωλο", όπως θα δούμε παρακάτω σε ένα σχετικό παράθεμα, χρησιμοποιείται ήδη από τον Πλάτωνα με παρεμφερή τρόπο και ταυτόχρονα έχει και στα νέα Ελληνικά τη σημασία, μεταξύ άλλων, του image.

5.2 Η υπόθεση περί ύπαρξης νοητικών ειδώλων

5.2.1 Προϊστορία της υπόθεσης περί ύπαρξης νοητικών ειδώλων

Η θέση ότι ο νους χρησιμοποιεί νοητικά είδωλα για να αναπαραστήσει τον κόσμο έχει μια μακρά ιστορία στη φιλοσοφία. Μια πρώτη εκδοχή της συναντάμε στο *Θεαίτητο* του Πλάτωνα, ο οποίος προτείνει ότι η μνήμη θα μπορούσε να είναι κάτι ανάλογο με μια πλάκα κεριού πάνω στην οποία αποτυπώνονται οι αισθήσεις και οι σκέψεις. Οποδήποτε αποτυπώνεται το διατηρούμε στη μνήμη και το γνωρίζουμε για όσο χρόνο υφίσταται εκεί το αποτύπωμα: "Και ὁ μὲν ἂν ἐκμαγῆ, μνημονεύειν τε καὶ ἐπίστασθαι, ἕως ἂν ἐνή τὸ εἶδωλον αὐτοῦ" (Πλάτων 1975a, 191 c-d). Ο Αριστοτέλης περιγράφει αυτά τα αποτυπώματα της μνήμης ως ένα είδος εικόνων (*De Memoria* 450 a-b). Ενώ, στο *Περὶ Ψυχῆς* (431a 15-20) δηλώνει ότι η ψυχή ποτέ δεν νοεῖ χωρίς νοητικά είδωλα [φαντάσματα]. Στη νεώτερη φιλοσοφία και κυρίως στον Βρετανικό Εμπειρισμό η έννοια του νοητικού ειδώλου αποκτά κεντρική σημασία. Σύμφωνα με τον Tye (1991, σελ. 5): "αντίθετα προς τον Καρτέσιο και τον Hobbes, οι βρετανοί εμπειριστές υποστήριζαν ότι κάθε είδους σκέψη συνίσταται στον χειρισμό είτε απλών νοητικών ειδώλων που πηγάζουν από την αισθητηριακή εμπειρία είτε σύνθετων νοητικών ειδώλων που κατασκευάζονται από αυτά τα απλά νοητικά είδωλα". Σταδιακά, ωστόσο, η έννοια των νοητικών ειδώλων άρχισε να θεωρείται όλο και πιο προβληματική μέχρι που στα μέσα του 20ου αιώνα έφτασε να αντιμετωπίζεται ως εντελώς λανθασμένη. Μετά την κριτική, κυρίως, των Ryle (1949), Wittgenstein (1953/1977) και Austin (1962) τα υποτιθέμενα είδωλα που δίδονται στο "μάτι του νου" απορρίφθηκαν ως εντελώς ανυπόστατα²²³.

Στο χώρο της ψυχολογίας, η στάση των ερευνητών όσον αφορά την έννοια των νοητικών ειδώλων ακολούθησε μια αντίστοιχη πορεία. Από την κεντρική θέση που κατείχε μέσα στην ενδοσκοπική ψυχολογία του Wundt, αλλά και στη ψυχολογία του James, φτάνουμε στην πλήρη απόρριψή της κατά την περίοδο επικράτησης του συμπεριφορισμού²²⁴ με το επιχείρημα ότι τα νοητικά είδωλα δεν μπορούν να αποτελέσουν αντικείμενο μιας επιστημονικής ψυχολογίας²²⁵.

Τόσο στην περίπτωση της φιλοσοφίας όσο και σ' αυτή της ψυχολογίας η έννοια των νοητικών ειδώλων ήταν άρρηκτα δεμένη με την έννοια της ενδοσκοπικής. Η αποδοχή ή η απόρριψη της μιας συνεπάγεται την αποδοχή ή η απόρριψη της άλλης. Δεν υπάρχουν νοητικά είδωλα

²²³ Βλ. κεφάλαιο 8, ενότητα 8.2.2.

²²⁴ Αυτό συνέβη ήδη από το 1913 ύστερα από τη δημοσίευση ενός άρθρου από τον Watson κατά των ειδώλων της φαντασίας. Ο Watson (1913b) ταυτίζει τα νοητικά είδωλα με τις 'κεντρικά εγείρμενες αισθήσεις' (centrally aroused sensations).

²²⁵ Βλ. κεφάλαιο 1, ενότητα 1.2.1.

5.1 Εισαγωγή

Στο προηγούμενο κεφάλαιο εξετάσαμε μια βασική κριτική κατά της ΑΥΘΝ η οποία έχει ως στόχο να δείξει τα όρια αυτού του εξηγητικού πλαισίου. Πιο συγκεκριμένα, είδαμε ότι υπάρχουν μη-προτασιακά είδη γνώσης που σχετίζονται με την κοινότητα και την περιστασιακή γνώση τα οποία δεν μπορούν να εκφραστούν με συμβολικές αναπαραστάσεις. Έτσι, δείξαμε ότι η ΑΥΘΝ δεν επαρκεί για την εξήγηση των συγκεκριμένων ειδών μη-προτασιακής γνώσης.

Στο παρόν κεφάλαιο θα εξετάσουμε μια απόπειρα να εξηγηθεί με τη βοήθεια αναπαραστάσεων ένα άλλο είδος γνώσης που ενέχει μη-προτασιακά στοιχεία, ήτοι τη γνώση που κατακτούμε μέσω της φαντασίας και μέσω της αντίληψης. Στη θέση των νοητικών συμβόλων αυτή τη φορά προτείνεται ένα νέο είδος αναπαραστάσεων: τα νοητικά είδωλα (*mental images*)²²². Η θεωρία περί νοητικών ειδώλων, παρότι προτάθηκε ως αντίπαλη προς τη θεωρία περί νοητικών συμβόλων, τελικά παραμένει μέσα στο γενικότερο πλαίσιο που θέτει η ΑΥΘΝ και απλώς αποτελεί μια προσπάθεια επέκτασης του συγκεκριμένου εξηγητικού πλαισίου. Αυτό ισχύει γιατί η θεωρία περί νοητικών ειδώλων δέχεται τη συστατική παραδοχή της ΑΥΘΝ ότι δηλαδή η εξήγηση των νοητικών φαινομένων που παράγονται στη φαντασία δεν είναι τίποτε άλλο παρά κατοχή και χειρισμός αναπαραστάσεων εικονιστικής μορφής.

Από μεθοδολογική άποψη, σε αυτό το κεφάλαιο δεν θα θέσουμε υπό αμφισβήτηση το ότι τα νοητικά είδωλα αποτελούν όντως ένα είδος αναπαραστάσεων, δηλαδή έχουν αναπαραστασιακό περιεχόμενο. Με άλλα λόγια, δεν θα εξετάσουμε το ερώτημα "δυνάμει τίνος πράγματος αποκτά μια εσωτερική κατάσταση του εγκεφάλου αναπαραστασιακό περιεχόμενο;". Αυτό το ερώτημα θα εξεταστεί στο έβδομο κεφάλαιο. Στο παρόν κεφάλαιο θα υποθέσουμε προσωρινά ότι με κάποιο τρόπο τα νοητικά είδωλα αποκτούν περιεχόμενο και θα εξετάσουμε από κει και πέρα τι προβλήματα δημιουργούνται.

²²² Δεν υιοθετούμε τον όρο "νοητική εικόνα" ως μετάφραση του "mental image" γιατί κατ' αυτόν τον τρόπο (α) χάνεται η διάκριση μεταξύ *image* και *picture*; (ο όρος "image" αναφέρεται σε ένα προϊόν της φαντασίας, ενώ ο όρος "picture" σε μια εξωτερική εικόνα) και (β) προαποφασίζεται χωρίς έρευνα ότι τα *mental images* είναι μια μορφή εικόνων. Επιλέξαμε την απόδοση "νοητικό είδωλο" και όχι τον αρχαιοελληνικό όρο "φάντασμα" που είναι ακριβώς ό,τι οι Λατίνοι μετέφρασαν ως "image", λόγω της πολύ διαφορετικής σημασίας που έχει αποκτήσει ο όρος "φάντασμα" στα νέα Ελληνικά. Ο όρος "είδωλο", όπως θα δούμε παρακάτω σε ένα σχετικό παράθεμα, χρησιμοποιείται ήδη από τον Πλάτωνα με παρεμφερή τρόπο και ταυτόχρονα έχει και στα νέα Ελληνικά τη σημασία, μεταξύ άλλων, του *image*.

5.2 Η υπόθεση περί ύπαρξης νοητικών ειδώλων

5.2.1 Προϊστορία της υπόθεσης περί ύπαρξης νοητικών ειδώλων

Η θέση ότι ο νους χρησιμοποιεί νοητικά είδωλα για να αναπαραστήσει τον κόσμο έχει μια μακρά ιστορία στη φιλοσοφία. Μια πρώτη εκδοχή της συναντάμε στο *Θεαίητο* του Πλάτωνα, ο οποίος προτείνει ότι η μνήμη θα μπορούσε να είναι κάτι ανάλογο με μια πλάκα κεριού πάνω στην οποία αποτυπώνονται οι αισθήσεις και οι σκέψεις. Οποδήποτε αποτυπώνεται το διατηρούμε στη μνήμη και το γνωρίζουμε για όσο χρόνο υφίσταται εκεί το αποτύπωμα: "Και δ. μὲν ἂν ἐκμαγῆ, μνημονεύειν τε καὶ ἐπίστασθαι, ἕως ἂν ἐνή τὸ εἰδωλον αὐτοῦ" (Πλάτων 1975a, 191 c-d). Ο Αριστοτέλης περιγράφει αυτά τα αποτυπώματα της μνήμης ως ένα είδος εικόνων (*De Memoria* 450 a-b). Ενώ, στο *Περί Ψυχῆς* (431a 15-20) δηλώνει ότι η ψυχή ποτέ δεν νοεῖ χωρίς νοητικά είδωλα [φαντάσματα]. Στη νεώτερη φιλοσοφία και κυρίως στον Βρετανικό Εμπειρισμό η έννοια του νοητικού ειδώλου αποκτά κεντρική σημασία. Σύμφωνα με τον Tyne (1991, σελ. 5): "αντίθετα προς τον Καρτέσιο και τον Hobbes, οι βρετανοί εμπειριστές υποστήριζαν ότι κάθε είδους σκέψη συνίσταται στον χειρισμό είτε απλών νοητικών ειδώλων που πηγάζουν από την αισθητηριακή εμπειρία είτε σύνθετων νοητικών ειδώλων που κατασκευάζονται από αυτά τα απλά νοητικά είδωλα". Σταδιακά, ωστόσο, η έννοια των νοητικών ειδώλων άρχισε να θεωρείται όλο και πιο προβληματική μέχρι που στα μέσα του 20ου αιώνα έφτασε να αντιμετωπίζεται ως εντελώς λανθασμένη. Μετά την κριτική, κυρίως, των Ryle (1949), Wittgenstein (1953/1977) και Austin (1962) τα υποτιθέμενα είδωλα που δίδονται στο "μάτι του νου" απορρίφθηκαν ως εντελώς ανυπόστατα²²³.

Στο χώρο της ψυχολογίας, η στάση των ερευνητών όσον αφορά την έννοια των νοητικών ειδώλων ακολούθησε μια αντίστοιχη πορεία. Από την κεντρική θέση που κατείχε μέσα στην ενδοσκοπική ψυχολογία του Wundt, αλλά και στη ψυχολογία του James, φτάνουμε στην πλήρη απόρριψή της κατά την περίοδο επικράτησης του συμπεριφορισμού²²⁴ με το επιχείρημα ότι τα νοητικά είδωλα δεν μπορούν να αποτελέσουν αντικείμενο μιας επιστημονικής ψυχολογίας²²⁵.

Τόσο στην περίπτωση της φιλοσοφίας όσο και σ' αυτή της ψυχολογίας η έννοια των νοητικών ειδώλων ήταν άρρηκτα δεμένη με την έννοια της ενδοσκόπησης. Η αποδοχή ή η απόρριψη της μιας συνεπάγεται την αποδοχή ή η απόρριψη της άλλης. Δεν υπάρχουν νοητικά είδωλα

²²³ Βλ. κεφάλαιο 8, ενότητα 8.2.2.

²²⁴ Αυτό συνέβη ήδη από το 1913 ύστερα από τη δημοσίευση ενός άρθρου από τον Watson κατά των ειδώλων της φαντασίας. Ο Watson (1913b) ταυτίζει τα νοητικά είδωλα με τις 'κεντρικά εγχειρόμενες αισθήσεις' (centrally aroused sensations).

²²⁵ Βλ. κεφάλαιο 1, ενότητα 1.2.1.

χωρίς να υπάρχει κάποιος που να τα βλέπει. Υπ' αυτή την έννοια, τα νοητικά είδωλα νοούνται ως οντότητες που παρουσιάζονται στο φαινομενολογικό επίπεδο, δηλαδή από την οπτική του πρώτου προσώπου.

5.2.2 Η διατύπωση της υπόθεσης μέσα στο πλαίσιο της γνωσιακής επιστήμης

Στα τέλη της δεκαετίας του '60 η εγκατάλειψη του συμπεριφορισμού και η άνοδος της γνωσιακής ψυχολογίας συνέβαλε στην αναζωπύρωση του ενδιαφέροντος για τα νοητικά είδωλα. Ορισμένοι ψυχολόγοι, αμφισβητώντας την καθολική εφαρμογή της υπόθεσης περί γλώσσας της σκέψης, υποστήριξαν ότι οι νοητικές λειτουργίες που επιτελούμε με τη φαντασία δεν βασίζονται σε συμβολικές αναπαραστάσεις αλλά σε εικονιστικές²²⁶. Η θέση ότι υπάρχουν νοητικές αναπαραστάσεις που δεν είναι συμβολικές, δεν αποτελεί ένα περιφερειακό ζήτημα που κατά συνέπεια έχει μικρές μόνο επιπτώσεις στην αναπαραστασιακή - υπολογιστική θεωρία του νου (ΑΥΘΝ): Η φαντασία παίζει κεντρικό ρόλο στις γνωσιακές λειτουργίες και έτσι οι αναπαραστάσεις που αυτή χρησιμοποιεί εμπλέκονται σε πολλά είδη λειτουργιών. Επίσης, το πρόβλημα επιτείνεται αν ισχύει η άποψη που προωθούν οι υποστηρικτές των νοητικών ειδώλων σύμφωνα με την οποία η λειτουργία της φαντασίας είναι ανάλογη με εκείνη της αντίληψης. Μια τέτοια άποψη συνεπάγεται ότι η αντίληψη βασίζεται και αυτή σε εικονιστικού τύπου αναπαραστάσεις. Δεν είναι, λοιπόν, τυχαίο το ότι οι σύγχρονοι υπέρμαχοι των νοητικών ειδώλων συγκέντρωσαν τα πυρά των σημαντικότερων υποστηρικτών της 'ορθόδοξης' γνωσιακής επιστήμης. Τόσο ο Fodor όσο, κυρίως, ο Pylyshyn και ο Dennett ενεπλάκησαν ενεργά στη διαμάχη επιχειρώντας να δείξουν τα αδιέξοδα στα οποία οδηγεί η υπόθεση περί ύπαρξης νοητικών ειδώλων²²⁷.

Το βασικό ερώτημα που πρέπει να εξετάσουμε όσον αφορά τη συγκεκριμένη υπόθεση είναι αν η εκ νέου υποστήριξή της μέσα στο πλαίσιο της γνωσιακής επιστήμης φέρνει κάτι καινούργιο ή αποτελεί απλώς μια επαναδιατύπωση κάποιων ιδεών που έχουν υποστεί συντριπτική κριτική από τη σύγχρονη φιλοσοφία. Όπως θα δούμε, η απάντηση σε αυτό το ερώτημα δεν είναι μονοσήμαντη: παρότι οι νέοι υποστηρικτές των νοητικών ειδώλων διανοίγουν ως ένα βαθμό κάποια προοπτική η οποία δεν ήταν παλιότερα διαθέσιμη, τελικά δεν καταφέρνουν να αποφύγουν σημαντικά λάθη του παρελθόντος.

Μια πρώτη διαφοροποίηση που παρατηρείται στη σύγχρονη προσέγγιση των νοητικών ειδώλων αφορά το εύρος των μελετώμενων φαινομένων: Το πεδίο της έρευνας εντοπίζεται σε αυτό που οι γνωσιακοί ψυχολόγοι αποκαλούν "imagery", έναν όρο που θα αποδώσουμε με τον

²²⁶ Βλ. Paivio 1971, Shepard & Metzler 1971/1992.

²²⁷ Βλ. Block 1981, Dennett 1969/1981, Pylyshyn 1981, Pylyshyn 2001.

όρο "χρήση νοητικών ειδώλων" καθώς αναφέρεται στη δημιουργία και το μετασχηματισμό νοητικών ειδώλων στη φαντασία. Έτσι, τουλάχιστον σε μια πρώτη φάση, η προτεινόμενη θεωρία δεν αφορά τη φαντασία εν γένει: Η λεγόμενη δημιουργική φαντασία, για παράδειγμα, δεν συνοδεύεται πάντα από νοητικά είδωλα. Ταυτόχρονα, όμως, υπάρχει και μια *διεύρυνση* του πεδίου έρευνας σε άλλες γνωσιακές ικανότητες όπως η μνήμη και η αντίληψη. Για παράδειγμα, πολλά από τα πειράματα των υποστηρικτών των νοητικών ειδώλων αφορούν είτε την ανάκληση από τη μνήμη εικόνων ή σκηνών είτε τη δημιουργία νοητικών ειδώλων που υπερτίθενται της αντίληψης, δηλαδή προβάλλονται στον αντιληπτό κόσμο²²⁸.

Εκτός από τη μεταβολή του πεδίου έρευνας, οι γνωσιακοί επιστήμονες που υποστηρίζουν την ύπαρξη νοητικών ειδώλων μεταβάλλουν και το *επίπεδο* στο οποίο διεξάγουν την έρευνά τους. Εκείνο που τους ενδιαφέρει πρωταρχικά δεν είναι ο τρόπος που τα νοητικά είδωλα δίδονται στη συνείδηση -στην ενδοσκόπηση-, αλλά οι διαδικασίες που *υπόκεινται* αυτής της εμφάνισης και την καθιστούν δυνατή. Με άλλα λόγια, δεν μελετούν τα νοητικά είδωλα στο φαινομενολογικό επίπεδο αλλά στο λειτουργικό και το νευροφυσιολογικό επίπεδο. Χαρακτηριστικά ο Kosslyn - ο σημαντικότερος υποστηρικτής των νοητικών ειδώλων- επισημαίνει:

"Η μακρόχρονη διαμάχη περί της χρήσης νοητικών ειδώλων [mental imagery] αφορά τη μορφή των αναπαραστάσεων που υπόκεινται της συνειδητής εμπειρίας του 'βλέπει με τα μάτια του νου', και όχι τη φύση των οπτικών περιεχομένων που αποθηκεύονται στη μνήμη ή το γεγονός ότι οι άνθρωποι έχουν τέτοιες εμπειρίες" (Kosslyn 1984, σελ. 104)

Έτσι, το πρόβλημα των νοητικών ειδώλων στη σύγχρονη εκδοχή του φαίνεται να αποσυνδέεται από το πρόβλημα της ενδοσκόπησης. Η σύγχρονη διαμάχη δεν αφορά το εάν οι αναπαραστάσεις που εμφανίζονται στην ενδοσκόπηση έχουν ή όχι εικονιστική μορφή. Τόσο οι υποστηρικτές όσο και οι πολέμιοι των νοητικών ειδώλων δέχονται ότι η *εμπειρία* νοητικών ειδώλων αποτελεί αναμφισβήτητο γεγονός. Όπως επισημαίνει και ο Kosslyn (1994, σελ. 6): "Το ζήτημα δεν ήταν το εάν οι άνθρωποι είχαν την εμπειρία οπτικών νοητικών ειδώλων. Όλοι συμφωνούσαν ότι έχουν"²²⁹. Εκείνο που τίθεται υπό συζήτηση είναι το εάν οι υποκείμενοι μηχανισμοί χρησιμοποιούν αναπαραστάσεις των οποίων η μορφή δεν είναι

²²⁸ Βλ. ενότητα 5.2.4.

²²⁹ Βλ., επίσης, Kosslyn 1990, σελ. 84: "Κανείς δεν αρνείται ότι οι άνθρωποι έχουν την εμπειρία του 'βλέπει με τα μάτια του νου', εκεί που υπάρχει διαφάνια είναι όσον αφορά το τι αποκαλύπτει αυτή η εμπειρία για τον τρόπο που ο εγκέφαλος αποθηκεύει τις σχετικές πληροφορίες"

συμβολική αλλά εικονιστική. Θα πρέπει, με άλλα λόγια, να κάνουμε μια σαφή διάκριση ανάμεσα στα νοητικά είδωλα που παρουσιάζονται στο φαινομενολογικό επίπεδο και τις αναπαραστάσεις που υπόκεινται αυτής της εμπειρίας και, κατά συνέπεια, δεν δίδονται στην ενδοσκόπηση. Η σύγχρονη διατύπωση της υπόθεσης περί ύπαρξης νοητικών ειδώλων αφορά ρητά την μορφή των νοητικών αναπαραστάσεων που υπόκεινται της εμπειρίας ειδώλων στη φαντασία. Σε ό,τι ακολουθεί θα χρησιμοποιούμε τον όρο "νοητικά είδωλα" για να αναφερόμαστε στις νοητικές αναπαραστάσεις εικονιστικής μορφής που υπόκεινται των εμπειριών που έχουμε από την οπτική του πρώτου προσώπου όταν μας παρουσιάζονται είδωλα στη φαντασία, ενώ για τις σχετικές εμπειρίες θα χρησιμοποιούμε τον όρο "είδωλα της φαντασίας". Έτσι λοιπόν, η αναφορά σε νοητικά είδωλα θα πρέπει να νοηθεί όπως ακριβώς και η αναφορά στις φοντοριανές συμβολικές νοητικές αναπαραστάσεις: δεν αφορούν την εμπειρία αλλά τους μηχανισμούς που υπόκεινται αυτής.

Βέβαια, η διάκριση που προτείνουμε εδώ ανάμεσα σε νοητικά είδωλα και σε είδωλα της φαντασίας δεν πρέπει να κατανοηθεί οντολογικά αλλά γνωσιολογικά. Δεν πρόκειται για δύο διακριτές οντότητες αλλά για δύο διαφορετικούς τρόπους με τους οποίους μπορούν να μας δοθούν οι συγκεκριμένες αναπαραστάσεις: ως νοητικά είδωλα από την οπτική του τρίτου προσώπου και ως είδωλα της φαντασίας από την οπτική του πρώτου προσώπου. Ο Kosslyn, ο οποίος χρησιμοποιεί μόνο έναν όρο - αυτόν του 'mental image' - για να αναφερθεί στο εν λόγω φαινόμενο, στο τελευταίο του βιβλίο (Kosslyn 1994), αναγνωρίζει την εγγενή αμφισημία του όρου:

"Αλλά αυτός ο όρος (ενν. mental images) είναι εξαιρετικά αμφίσημος, και το μεγαλύτερο ενδιαφέρον στη ψυχολογία δόθηκε σε μια μόνο όψη της χρήσης νοητικών ειδώλων -στο ρόλο της στην επεξεργασία πληροφοριών, όχι στη φαινομενολογία της ή στο ρόλο της στην συναισθηματική ζωή. Σε αυτό το βιβλίο θα εστιάσουμε στη φύση των εσωτερικών συμβάντων που υπόκεινται της εμπειρίας του 'βλέπει με τα μάτια του νου', δεν θα εξετάσουμε τις ποιότητες της ίδιας της εμπειρίας. Ο όρος 'image' θα αναφέρεται στην εσωτερική αναπαράσταση που χρησιμοποιείται στην επεξεργασία πληροφοριών, όχι στην ίδια την εμπειρία. Η εμπειρία της χρήσης ειδώλων στη φαντασία είναι ένα σημάδι ότι τα υποκείμενα εγκεφαλικά συμβάντα λαμβάνουν χώρα, και κατά συνέπεια παίζει έναν ανεκτίμητο ρόλο στην έρευνα" (Kosslyn 1994, σελ. 3)

Η θέση ότι τα νοητικά είδωλα και τα είδωλα της φαντασίας αποτελούν από οντολογική άποψη ένα και μοναδικό φαινόμενο παρουσιάζει αρκετά προβλήματα όπως θα δούμε στο όγδοο κεφάλαιο

(ενότητα 8.3)²³⁰. Επίσης, ανοικτό παραμένει το ζήτημα για το εάν η χρήση νοητικών ειδώλων συνεπάγεται κατ' ανάγκη την εμφάνιση ειδώλων στη φαντασία και το αντίστροφο. Θα ήταν πολύ περιοριστικό το να υποθέσουμε ότι κάθε εγκεφαλική λειτουργία που κάνει χρήση νοητικών ειδώλων 'παράγει' πάντα είδωλα στη φαντασία ή, για να το πούμε κάπως πιο ποιητικά, αφήνει κάποιο ίχνος στη συνείδηση.

Ας εξετάσουμε όμως τώρα την ορθότητα των επιχειρημάτων που προτείνονται για την υποστήριξη της υπόθεσης περί ύπαρξης νοητικών ειδώλων. Θα περιοριστούμε κυρίως στη θεωρία και τα επιχειρήματα που αναπτύσσει ο Kosslyn γιατί αυτός είναι ο κυριότερος και πιο συστηματικός υποστηρικτής των νοητικών ειδώλων. Πριν προχωρήσουμε, ωστόσο, είναι σκόπιμο να επισημάνουμε ένα πρόβλημα που δημιουργείται από τον όρο που επιλέγει ο Kosslyn για να χαρακτηρίσει τις αναπαραστάσεις της ΑΥΘΝ. Ο Kosslyn ονομάζει αυτές τις αναπαραστάσεις "προτασιακές" και τις αντιπαράθετει προς τις απεικονιστικές [depictive] ή εικονιστικές [pictorial] αναπαραστάσεις που ο ίδιος υποστηρίζει. Ωστόσο, αυτή η επιλογή δημιουργεί μια σύγχυση διότι ο όρος "προτασιακός", όπως χρησιμοποιείται στη φιλοσοφία, απ' όπου και ξεκίνησε (βλ. κεφάλαιο 2, ενότητα 2.3.2 και κεφάλαιο 3, ενότητα 3.3.1.2), αναφέρεται σε ένα είδος περιεχομένου. Από την άλλη μεριά, η διαμάχη του Kosslyn με τους υποστηρικτές της κλασικής γνωσιακής επιστήμης αφορά ρητά την αναπαραστασιακή μορφή. Κατά συνέπεια, είναι λανθασμένο το να αποδίδουμε στην αναπαραστασιακή μορφή έναν χαρακτηρισμό που προσιδιάζει στο αναπαραστασιακό περιεχόμενο. Γι' αυτό, προκειμένου να διακρίνουμε τις απεικονιστικές αναπαραστάσεις από τις αναπαραστάσεις που δέχεται η κλασική γνωσιακή επιστήμη θα αναφερόμαστε στις τελευταίες με τον όρο "συμβολικές αναπαραστάσεις" που υιοθετήσαμε στο δεύτερο κεφάλαιο.

5.2.3 Φαινομενολογικά τεκμήρια

Η φαινομενολογία της εμπειρίας ειδώλων στη φαντασία παρέχει τα πρωταρχικά τεκμήρια που οδήγησαν στην υποστήριξη της υπόθεσης περί νοητικών ειδώλων. Ωστόσο, ταυτόχρονα, η πεποίθηση ότι αυτά τα τεκμήρια "συλλέγονται" μέσω της ενδοσκόπησης, τα έχει καταστήσει αμφισβητήσιμα στα μάτια των υποστηρικτών και κυρίως των επικριτών των νοητικών ειδώλων. Γι' αυτό το λόγο, η επίκληση των φαινομενολογικών τεκμηρίων δεν είναι συστηματική, ούτε γίνεται σε πρώτο επίπεδο. Παρόλα αυτά θεωρούμε ότι τα συγκεκριμένα τεκμήρια παίζουν καθοριστικό ρόλο στο ρητορικό επίπεδο γιατί όλοι μας έχουμε την

²³⁰ Βλ., επίσης, Block 1983, σελ. 588.

εμπειρία ειδώλων στη φαντασία ή, τουλάχιστον, έχουμε συνηθίσει να περιγράφουμε με αυτόν τον τρόπο τα περιεχόμενα της φαντασίας²³¹.

5.2.3.1 Το επιχείρημα με βάση τη φαινομενολογία των ειδώλων της φαντασίας

Ένα πρώτο επιχείρημα υπέρ της ύπαρξης νοητικών ειδώλων βασίζεται στη φαινομενολογία της φαντασίας και γενικότερα της χρήσης ειδώλων. Για την ακρίβεια δεν πρόκειται για επιχείρημα αλλά για ένα, εκ πρώτης όψεως, "εύλογο" συμπέρασμα: Εφόσον κατά τη χρήση ειδώλων στη φαντασία έχουμε την εμπειρία κάποιων εικόνων που παρουσιάζονται στον νου, τι πιο φυσικό να συμπεράνουμε ότι οι αναπαραστάσεις που υποκείνται αυτής της εμπειρίας έχουν και αυτές εικονιστική μορφή.

Υπάρχουν τουλάχιστον τρεις παραδοχές στις οποίες βασίζεται η (όποια) ευλογοφάνεια του παραπάνω συμπεράσματος:

1. η φαινομενολογία της χρήσης ειδώλων στη φαντασία είναι όμοια με τη φαινομενολογία της αντίληψης (εξωτερικών) εικόνων
2. τα νοητικά φαινόμενα ανάγονται στην ύπαρξη και τον χειρισμό ασυνείδητων αναπαραστάσεων στο λειτουργικό και νευροφυσιολογικό επίπεδο.
3. Το περιεχόμενο των νοητικών φαινομένων -όπως αυτό δίδεται στο φαινομενολογικό επίπεδο- κατοπτρίζεται στην μορφή των υποκείμενων αναπαραστάσεων του λειτουργικού - νευροφυσιολογικού επιπέδου.

Θεωρούμε ότι και οι τρεις παραπάνω παραδοχές είναι προβληματικές. Όσον αφορά την πρώτη παραδοχή θα δείξουμε στο όγδοο κεφάλαιο (ενότητα 8.2.2) ότι είναι λανθασμένη: η φαινομενολογία της εμφάνισης ειδώλων του Χ στη φαντασία ομοιάζει περισσότερο με τη φαινομενολογία της αντίληψης του ίδιου του Χ παρά με την φαινομενολογία της αντίληψης μιας (εξωτερικής) εικόνας του Χ. Ωστόσο, στο παρόν κεφάλαιο δεν θα κάνουμε πρωθύστερη χρήση αυτής της κριτικής. Θα δεχθούμε, προσωρινά, ως ορθή την παραπάνω αφετηριακή παραδοχή των υποστηρικτών αλλά και των περισσότερων από τους επικριτές των νοητικών ειδώλων και θα εξετάσουμε τα προβλήματα που προκύπτουν από 'κει και πέρα.

Η δεύτερη παραδοχή αποτελεί, όπως είδαμε και στο δεύτερο κεφάλαιο, τη συστατική υπόθεση της ΑΥΘΝ με τη διαφορά ότι τώρα στη θέση των συμβολικών αναπαραστάσεων προτείνονται οι εικονιστικές αναπαραστάσεις. Ο συστατικός ρόλος που παίζει η συγκεκριμένη παραδοχή στη συγκρότηση των σύγχρονων αναπαραστασιακών θεωριών περί νοητικών φαινομένων καθιστά σαφές ότι η απόρριψή της οδηγεί στην

²³¹ Στο όγδοο κεφάλαιο (ενότητα 8.2.2) θα ασχοληθούμε διεξοδικά με το συγκεκριμένο ζήτημα.

απόρριψη και των αναπαραστασιακών θεωριών. Ωστόσο, προκειμένου να δικαιολογηθεί μια τέτοια στάση απαιτείται μια συστηματική κριτική της επάρκειας της έννοιας της νοητικής αναπαράστασης στην εξήγηση των νοητικών φαινομένων. Η κριτική αυτή, η οποία αποτελεί τον κεντρικό στόχο ολόκληρης της διατριβής, θα επιτελεστεί σταδιακά. Ήδη στο προηγούμενο κεφάλαιο αμφισβητήσαμε, σε ένα πρώτο στάδιο, την επάρκεια της ΑΥΘΝ όσον αφορά το εύρος των εξηγούμενων φαινομένων. Συγκεκριμένα, δείξαμε ότι υπάρχουν είδη γνώσης- ήτοι, γνωστικού περιεχομένου- που δεν είναι δυνατόν να εκφραστούν με συμβολικές αναπαραστάσεις. Αυτά τα είδη γνώσης τα ονομάσαμε "μη-προτασιακή γνώση". Στο παρόν και στο επόμενο κεφάλαιο θα ασκήσουμε κριτική στις απόπειρες εξήγησης ειδών γνώσης που ενέχουν μη-προτασιακά στοιχεία με αναδρομή σε μη-συμβολικές αναπαραστάσεις. Μετά την ολοκλήρωση αυτών των κριτικών θα είμαστε σε θέση στα κεφάλαια 7 και 8 να αμφισβητήσουμε πλέον άμεσα την παραδοχή 2, δείχνοντας ότι τα νοητικά φαινόμενα δεν μπορούν να εξηγηθούν με βάση την έννοια της νοητικής αναπαράστασης.

Τέλος, όσον αφορά την τρίτη παραδοχή αποτελεί και αυτή μια βασική παραδοχή της ΑΥΘΝ. Όπως είδαμε στο δεύτερο κεφάλαιο, το περιεχόμενο των νοητικών αναπαραστάσεων (και κατ' επέκταση των νοητικών φαινομένων) θεωρείται ότι κατοπτρίζεται στη μορφή των νοητικών αναπαραστάσεων (στο λειτουργικό επίπεδο). Για την ακρίβεια ο Fodor μιλάει, όπως ήδη διαπιστώσαμε, για ένα είδος ισομορφισμού: "Η δομή των γλωσσικών προτάσεων είναι κατά τον ακόλουθο τρόπο *ισομορφική* προς τη δομή των προτάσεων που εκφράζουν: *εάν μια γλωσσική πρόταση S εκφράζει την πρόταση P, τότε συντακτικά συστατικά της S εκφράζουν τα συστατικά της P*" (Fodor & Lepore 1994, σελ. 147). Ωστόσο, στο παρόν κεφάλαιο, η παραδοχή 3 παρουσιάζει, τουλάχιστον, δύο διαφοροποιήσεις:

α) η αναφορά στο περιεχόμενο των νοητικών φαινομένων, όπως εξειδικεύεται στην περίπτωση της χρήσης ειδώλων στη φαντασία, αφορά τον τρόπο που αυτό παρουσιάζεται από την οπτική του πρώτου προσώπου, ενώ στην περίπτωση της ΑΥΘΝ αφορά τον τρόπο που το περιεχόμενο δίδεται από την οπτική του τρίτου προσώπου.

β) δεν γίνεται σαφής διάκριση μεταξύ λειτουργικού και νευροφυσιολογικού επιπέδου

Η πρώτη διαφοροποίηση περιπλέκει την κατάσταση όσον αφορά το τι ακριβώς υποστηρίζουν οι υπέρμαχοι των νοητικών ειδώλων. Ενώ η κλασική ΑΥΘΝ έχοντας απορρίψει την οπτική του πρώτου προσώπου και την ενδοσκόπηση, κατάφερε να μιλά για τρία επίπεδα περιγραφής -το σημασιολογικό, το συντακτικό/λειτουργικό και το νευροφυσιολογικό/φυσικό, οι υποστηρικτές των νοητικών ειδώλων προσθέτουν κι ένα τέταρτο επίπεδο, το φαινομενολογικό. Τα ερωτήματα που δημιουργεί αυτή η προσθήκη είναι (α) σε τι ακριβώς έγκειται το περιεχόμενο των νοητικών φαινομένων που δίδεται στο φαινομενολογικό επίπεδο και (β) εάν αυτό το περιεχόμενο διαφοροποιείται από το

περιεχόμενο που δίδεται από την οπτική του τρίτου προσώπου. Τα ερωτήματα αυτά θα εξεταστούν στο όγδοο κεφάλαιο (ενότητα 8.2).

Η δεύτερη διαφοροποίηση, όπως θα δούμε στη συνέχεια, σχετίζεται με τη επιμονή του Kosslyn (1994) να υποστηρίζει τόσο ότι τα νοητικά είδωλα είναι ουσιαστικά "λειτουργικά είδωλα" [functional images] που δημιουργούνται σε έναν "λειτουργικό χώρο", όσο και ότι η νευροφυσιολογική έρευνα συμβάλλει καθοριστικά στην επιβεβαίωση της υπόθεσης ότι ο εγκέφαλος χρησιμοποιεί εικονιστικές αναπαραστάσεις με την ανακάλυψη της ύπαρξης εικονιστικών αναπαραστάσεων πάνω στον εγκεφαλικό ιστό.

Εκείνο, πάντως, που πρέπει να επισημάνουμε είναι ότι ακόμα κι αν δεχτούμε ότι στη φαντασία εμφανίζονται εικόνες και ότι αυτής της εμπειρίας υπόκεινται νοητικές αναπαραστάσεις, δεν συνεπάγεται από τις συγκεκριμένες παραδοχές ότι οι υποκείμενες νοητικές αναπαραστάσεις είναι εικονιστικής μορφής. Τίποτε δεν απαγορεύει το να υπόκεινται των ειδώλων της φαντασίας κάποιες συμβολικές αναπαραστάσεις²³².

Αν, πάλι, δεχτούμε ότι οι αναπαραστάσεις που υπόκεινται των ειδώλων της φαντασίας είναι εικονιστικής μορφής, οδηγούμαστε σε ένα εξωφρενικό, εκ πρώτης όψεως, συμπέρασμα. Αν δεχτούμε ότι ισχύει η υπόθεση περί ύπαρξης νοητικών ειδώλων, τότε θα πρέπει να δεχτούμε ότι όταν φαντάζομαι ένα ροζ ποντίκι τότε κάποιο τμήμα του εγκεφαλικού φλοιού όχι μόνο εμφανίζει το σχήμα του ποντικιού αλλά "βάφεται" και ροζ. Ο Kosslyn για να αποφύγει μια τέτοια εύκολη κριτική (ή, καλύτερα, γελοιοποίηση) της θεωρίας του, ξεκαθαρίζει ότι όταν μιλά περί νοητικών ειδώλων δεν εννοεί ότι υφίστανται κυριολεκτικά ως εικόνες κάπου στον εγκέφαλο, αλλά ως λειτουργικές εικόνες. Με άλλα λόγια, τα νοητικά είδωλα υπάρχουν σε ένα μέσο που λειτουργεί ως χώρος. Ο Kosslyn ονομάζει αυτό το μέσο "οπτική ενδιάμεση μνήμη" (visual buffer). Φυσικά, το πρόβλημα δεν επιλύεται με την εισαγωγή νέων όρων. Εκείνο που προέχει στην προκειμένη περίπτωση είναι η εξήγηση του πώς υπάρχει και δρα μια λειτουργική εικόνα. Θα επανέλθουμε σε αυτό το ζήτημα στην ενότητα 5.2.4.

5.2.3.2 Το επιχείρημα με βάση την ομοιότητα της φαινομενολογίας της φαντασίας με εκείνη της αντίληψης

Ένα δεύτερο επιχείρημα υπέρ της ύπαρξης νοητικών ειδώλων στηρίζεται στην ομοιότητα της φαντασίας με την αντίληψη. Αυτή η ομοιότητα υποστηρίζεται σε δύο επίπεδα: στο φαινομενολογικό και στο λειτουργικό/νευροφυσιολογικό. Στην παρούσα ενότητα θα εξετάσουμε την ομοιότητα στο πρώτο επίπεδο. Η ομοιότητα στο δεύτερο επίπεδο θα εξεταστεί στην ενότητα 5.2.5.1.

²³² Βλ. Block 1983.

Το επιχείρημα μπορεί να διατυπωθεί ως εξής: η εμπειρία που έχουμε όταν φανταζόμαστε το X είναι όμοια με την εμπειρία που έχουμε όταν αντιλαμβανόμαστε το X, κατά συνέπεια οι νοητικές αναπαραστάσεις που υπόκεινται των δύο αυτών ειδών εμπειρίας είναι του ίδιου τύπου, δηλαδή πρόκειται για εικονιστικές αναπαραστάσεις. Δεν χρειάζεται ιδιαίτερη προσπάθεια για να γίνει σαφές ότι το συγκεκριμένο επιχείρημα δεν ευσταθεί. Καταρχάς, ισχύει η παρατήρηση που έγινε στη προηγούμενη ενότητα: δεν μπορούμε να συμπεράνουμε το είδος των (υποτιθέμενων) υποκείμενων αναπαραστάσεων στηριζόμενοι στην φαινομενολογία της εμπειρίας που αυτές υποστηρίζουν. Όμως το σημαντικότερο πρόβλημα στην προκειμένη περίπτωση είναι ότι το επιχείρημα βασίζεται σε μια αστήρικτη παραδοχή, ότι, δηλαδή, οι νοητικές αναπαραστάσεις που υπόκεινται της αντίληψης είναι εικονιστικής μορφής.

Ο Kosslyn, στα πρώτα χρόνια ανάπτυξης της θεωρίας του, επιχείρησε να υποστηρίξει μια συγγενική θέση, ότι δηλαδή η χρήση ειδώλων στη φαντασία και η αντίληψη στηρίζονται στους ίδιους εγκεφαλικούς μηχανισμούς. Για την απόδειξη αυτής της θέσης επικαλέστηκε κυρίως συμπεριφοριστικά τεκμήρια. Ένα πρώτο τεκμήριο είναι η παρεμβόδιση (interference) της χρήσης ειδώλων στη φαντασία από την αντίληψη. Πιο συγκεκριμένα, "...τα νοητικά είδωλα παρεμποδίζουν την αντίληψη όταν είναι πολύ παρεμφερή με τα προς ανίχνευση ερεθίσματα" (Kosslyn 1994, σελ. 328). Ένα δεύτερο τεκμήριο είναι η διευκόλυνση της χρήσης ειδώλων στη φαντασία από την αντίληψη, αν ο σχηματισμός που παράγεται στη φαντασία είναι όμοιος με αυτόν που γίνεται αντιληπτός²³³. Τέλος, η διαπίστωση ιδιαιτεροτήτων του ίδιου τύπου στη λειτουργία της χρήσης ειδώλων στη φαντασία και της αντίληψης υποδεικνύει και αυτή, σύμφωνα με τον Kosslyn, την ύπαρξη κοινού μηχανισμού που υποστηρίζει τις δύο ικανότητες²³⁴.

Δεν θα υπεισέλθουμε στις λεπτομέρειες των πειραμάτων με τα οποία ορισμένοι ψυχολόγοι έφθασαν στα παραπάνω συμπεράσματα, ούτε, πολύ περισσότερο, θα επιχειρήσουμε να αμφισβητήσουμε αυτά τα συμπεράσματα²³⁵. Ο λόγος για τον οποίον κρατούμε τη συγκεκριμένη στάση είναι ότι ακόμα κι αν δεχθεί κανείς τη θέση ότι η χρήση ειδώλων στη φαντασία και η αντίληψη βασίζονται στους ίδιους εγκεφαλικούς μηχανισμούς τότε δεν προκύπτει από αυτή την παραδοχή όσον αφορά το είδος των αναπαραστάσεων που χρησιμοποιούν οι εν λόγω μηχανισμοί.

²³³ "... εάν ο σχηματισμός που παράγεται στη φαντασία είναι όμοιος με τον αντιληπτό σχηματισμό, τότε η χρήση νοητικών ειδώλων μπορεί να διευκολύνει την οπτική κωδικοποίηση και όχι να την παρεμποδίσει" (Kosslyn 1994, σελ. 328)

²³⁴ "... εάν η χρήση ειδώλων στη φαντασία και η αντίληψη βασίζονται αμφότερες στον ίδιο μηχανισμό κωδικοποίησης, τότε ιδιοτροπίες στην λειτουργία αυτών των μηχανισμών θα πρέπει συχνά να εμφανίζονται και στις δύο περιπτώσεις." (Kosslyn 1994, σελ. 328)

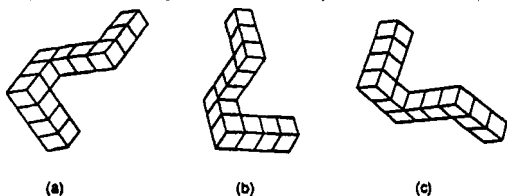
²³⁵ Για επιχείρηματα προς αυτή την κατεύθυνση βλ. Pylyshyn 2001.

Για παράδειγμα, θα μπορούσαν να χρησιμοποιούν συμβολικές αναπαραστάσεις, εικονιστικές αναπαραστάσεις ή να μην στηρίζονται καν σε αναπαραστάσεις. Εδώ λοιπόν, απαιτείται ένα επιπλέον βήμα, απαιτείται να δειχθεί ότι οι εγκεφαλικοί μηχανισμοί στους οποίους στηρίζεται η αντίληψη χρησιμοποιούν αναπαραστάσεις εικονιστικής μορφής. Όπως θα δούμε στην ενότητα 5.2.5.2, ο Kosslyn στο τελευταίο του βιβλίο επιχειρεί να υποστηρίξει ευθέως αυτή τη θέση ανατρέχοντας στα νευροφυσιολογικά δεδομένα.

5.2.4 Συμπεριφοριστικά τεκμήρια

Τα πιο γνωστά τεκμήρια που οδήγησαν στην αναζωπύρωση του ενδιαφέροντος για τα νοητικά είδωλα είναι χωρίς αμφιβολία τα πειράματα νοητικής περιστροφής (mental rotation) των Shepard και Metzler (Shepard & Metzler 1971/1992) καθώς και τα πειράματα νοητικής σάρωσης (mental scanning) του Kosslyn (1978). Αναμφίβολα η εποπτικότητα αυτών των πειραμάτων έπαιξε ένα σημαντικό ρόλο στην ευρύτατη γνωστοποίηση και μελέτη τους.

Τόσο τα πειράματα νοητικής περιστροφής όσο και τα πειράματα νοητικής σάρωσης στηρίζονταν σε μια κοινή μεθοδολογική γραμμή: τα εξεταζόμενα υποκείμενα υποχρεώνονταν να εκτελέσουν κάποιο νοητικό έργο που απαιτούσε τη χρήση της φαντασίας και ο πειραματιστής χρονομετρούσε την απόδοσή τους. Τα συμπεράσματα εξάγονταν από την ανάλυση των χρονικών διαφορών που προέκυπταν κατά την εκτέλεση διαφορετικών νοητικών έργων. Έτσι, κατά το πείραμα νοητικής περιστροφής παρουσιάστηκαν στα υποκείμενα σχήματα όπως τα παρακάτω και τους ζητήθηκε να προσδιορίσουν αν το δεύτερο και το τρίτο σχήμα απεικονίζουν το ίδιο τρισδιάστατο αντικείμενο με εκείνο που

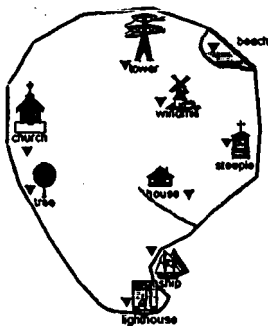


απεικονίζει το πρώτο σχήμα.

Το συμπέρασμα του πειράματος ήταν ότι ο χρόνος εκτέλεσης καθενός από τα δύο έργα είναι ανάλογος της γωνίας που

πρέπει να περιστραφεί νοητικά το κάθε σχήμα προκειμένου να συμπέσει με το πρώτο σχήμα.

Το πείραμα της νοητικής σάρωσης είχε ως εξής: δόθηκε στα υποκείμενα ο χάρτης ενός νησιού πάνω στον οποίο σημειωνόταν με εικονίδια διάφορες τοποθεσίες και τους ζητήθηκε να μελετήσουν τον χάρτη ούτως ώστε να είναι σε θέση να τον ζωγραφίσουν από μνήμης. Κατόπιν ζητήθηκε από τα εξεταζόμενα υποκείμενα να φανταστούν τον χάρτη και να εστιάσουν την προσοχή τους σε μια συγκεκριμένη τοποθεσία. Στη συνέχεια, τους ζητήθηκε να εστιάσουν διαδοχικά σε όλες τις υπόλοιπες τοποθεσίες. Το αποτέλεσμα του πειράματος ήταν ότι ο



χρόνος που χρειάστηκε για να εστιαστεί η προσοχή από μια τοποθεσία σε μια άλλη ήταν ανάλογος της απόστασης μεταξύ των τοποθεσιών.

Το αποτέλεσμα αυτό όπως και εκείνο του προηγούμενου πειράματος επιβεβαιώνει, σύμφωνα με τον Kosslyn, την υπόθεση ότι οι αναπαραστάσεις που υπόκεινται της εμφάνισης ειδώλων στη φαντασία είναι εικονιστικές και όχι συμβολικές, διότι εάν τα νοητικά είδωλα "απεικονίζουν τις πληροφορίες, τότε θα πρέπει να διαρκεί

περισσότερο χρόνο ο εντοπισμός των αναπαραστάσεων τμημάτων που βρίσκονται μακρύτερα από το σημείο εστίασης" (Kosslyn 1990, σελ. 89). Με άλλα λόγια, τα πειραματικά αποτελέσματα επιβεβαιώνουν την υπόθεση ότι οι αναπαραστάσεις που υπόκεινται της εμφάνισης ειδώλων στη φαντασία είναι εικονιστικές και όχι συμβολικές, διότι η απόσταση είναι ένα εγγενές χαρακτηριστικό των εικονιστικών αναπαραστάσεων και όχι των συμβολικών. "Εγγενές χαρακτηριστικό" σημαίνει ότι οι διάφορες φυσικές αποστάσεις μεταξύ τμημάτων της αναπαραστασιακής μορφής (δηλαδή, μεταξύ τμημάτων του αναπαραστασιακού φορέα) μιας εικονιστικής αναπαράστασης αναπαριστούν αποστάσεις μεταξύ αντίστοιχων τμημάτων του αναπαριστώμενου. Για αυτό λοιπόν το λόγο, ο Kosslyn (1990, σελ. 87), ισχυρίζεται ότι "εάν απεικονιστικές αναπαραστάσεις υπόκεινται της εμπειρίας του 'έχειν ένα είδωλο στη φαντασία', τότε η χωρική φύση των αναπαραστάσεων θα πρέπει να επηρεάσει το πώς γίνεται η επεξεργασία των ειδώλων". Αντίθετα, αν οι αναπαραστάσεις που υπόκεινται της εμφάνισης ειδώλων στη φαντασία είναι συμβολικές, τότε αποστάσεις μεταξύ αναπαριστώμενων αντικειμένων δεν αναπαριστώνται με τις αποστάσεις μεταξύ των αντίστοιχων αναπαραστασιακών φορέων, με άλλα λόγια οι αποστάσεις μεταξύ τμημάτων του φορέα μιας συμβολικής αναπαράστασης δεν έχουν αναπαραστασιακή σημασία. Γι' αυτό, αν τα είδωλα της φαντασίας στηρίζονταν σε συμβολικές αναπαραστάσεις, οι αποστάσεις μεταξύ των τοποθεσιών στο πείραμα του Kosslyn θα αναπαριστώνταν και αυτές συμβολικά και κατά συνέπεια ο χρόνος μετάβασης από τη μια τοποθεσία στην άλλη δεν θα μπορούσε να διαφέρει.

Ομολογουμένως, η παραπάνω συλλογιστική φαίνεται αρκετά ανορθόδοξη. Μία πρώτη γενική παρατήρηση είναι ότι και πάλι προϋποτίθεται η παραδοχή 2 που επισημάναμε στην ενότητα 5.2.3.1. Προϋποτίθεται, δηλαδή, ότι τα νοητικά φαινόμενα ανάγονται στην ύπαρξη και τον χειρισμό ασυνείδητων αναπαραστάσεων στο λειτουργικό και νευροφυσιολογικό επίπεδο. Το ερώτημα στο οποίο έρχονται να δώσουν απάντηση τα προαναφερθέντα πειράματα είναι το εξής: με δεδομένο ότι η εμφάνιση ειδώλων στη φαντασία στηρίζεται στην ύπαρξη και την επεξεργασία ασυνείδητων νοητικών αναπαραστάσεων, τι είδους

αναπαραστάσεις πρέπει να είναι αυτές; Κατά συνέπεια, είναι σαφές ότι τα πειράματα νοητικής περιστροφής και νοητικής σάρωσης δεν αποδεικνύουν την αναγκαιότητα ύπαρξης νοητικών αναπαραστάσεων, αλλά απλώς στοχεύουν στο να ενισχύσουν τη θέση ότι, *αν υπάρχουν τέτοιες αναπαραστάσεις, θα πρέπει να είναι εικονιστικής μορφής*.

Ένα δεύτερο ζήτημα που πρέπει να εξετάσουμε είναι το εάν τα τεκμήρια στα οποία βασίζονται τα προαναφερθέντα πειράματα είναι όντως συμπεριφοριστικά. Ο Kosslyn αφήνει να εννοηθεί ότι με το συγκεκριμένο τύπο πειραμάτων η μελέτη των ειδώλων της φαντασίας έπαψε να βασίζεται στην ενδοσκοπήση και έγινε επιστημονική²³⁶. Εφόσον, δηλαδή, η μετρούμενη μεταβλητή είναι ο χρόνος και, ειδικότερα, οι χρονικές διαφορές (ή, οι χρόνοι αντίδρασης²³⁷) πρόκειται για ένα αντικειμενικό μέγεθος που διαπιστώνεται με βάση τη συμπεριφορά των υποκειμένων.

Όμως τα πράγματα δεν είναι τόσο απλά, διότι η συμπεριφορά των υποκειμένων και, συγκεκριμένα, η πράξη με την οποία δηλώνεται το τέλος ενός έργου βασίζεται σε μη διυποκειμενικά διαπιστώσιμα τεκμήρια, δηλαδή σε τεκμήρια που δίδονται ενδοσκοπικά. Μόνο το ίδιο το υποκείμενο μπορεί να γνωρίζει πότε τελειώνει μια νοητική σάρωση. Το ζήτημα που θέτουμε εδώ δεν έχει, φυσικά, να κάνει με την ενδεχόμενη ανειλικρίνεια των εξεταζόμενων υποκειμένων, αλλά με το γεγονός ότι τα παραπάνω πειράματα, σε τελική ανάλυση, δεν βασίζονται σε συμπεριφοριστικά τεκμήρια αλλά στην ενδοσκοπήση²³⁸.

Η έννοια της ενδοσκοπήσης, βέβαια, παρουσιάζει πολλά προβλήματα (που θα εξεταστούν αναλυτικά στο όγδοο κεφάλαιο), όμως υπάρχει κάτι στο οποίο οι περισσότεροι θα μπορούσαν να συμφωνήσουν: μέσω της ενδοσκοπήσης έχουμε πρόσβαση σε νοητικά περιεχόμενα. Αν, μάλιστα, δεχτούμε την παραδοχή 2, μπορούμε να ισχυριστούμε ότι μέσω της ενδοσκοπήσης έχουμε πρόσβαση στο *περιεχόμενο* των νοητικών αναπαραστάσεων και όχι στη μορφή τους. Ο λόγος που κάνουμε αυτή την επισήμανση είναι ότι ανοίγει το δρόμο για μια εναλλακτική εξήγηση των χρονικών διαφορών που παρατηρούνται στα προαναφερθέντα πειράματα. Είναι ενδεχόμενο οι παρατηρούμενες χρονικές διαφορές να μην οφείλονται στο είδος της υποκειμένης αναπαραστασιακής μορφής αλλά στο είδος του αναπαριστώμενου *περιεχομένου*²³⁹. Για παράδειγμα, ο Pylyshyn (1981) έκανε το ακόλουθο πείραμα: παρουσίασε στα εξεταζόμενα υποκείμενα ένα χάρτη με φωτισμένες τις επιμέρους περιοχές του. Τα υποκείμενα είχαν τη

²³⁶ Δεν είναι τυχαίο, εξάλλου, ότι τα συγκεκριμένα πειράματα σήμαναν την επανένταξη της μελέτης των νοητικών ειδώλων στην ψυχολογική έρευνα.

²³⁷ Shepard & Metzler 1971/1992, σελ. 217.

²³⁸ Ενδεχομένως, αυτό το μεθοδολογικό πρόβλημα μπορεί να ξεπεραστεί με την ταυτόχρονη χρήση νευροφυσιολογικών τεκμηρίων. Θα εξετάσουμε τη χρησιμότητα των συγκεκριμένων τεκμηρίων στη ενότητα 5.2.5.

²³⁹ Βλ. Pylyshyn 2001.

δυνατότητα να εξετάζουν την κάθε περιοχή ανοίγοντας το κατάλληλο φως. Προφανώς στην προκειμένη περίπτωση ο χρόνος που απαιτείται για να εστιάσουμε την προσοχή από τη μια περιοχή του χάρτη σε μια οποιαδήποτε άλλη είναι πολύ μικρός και σχεδόν ο ίδιος. Ο Pylyshyn στη συνέχεια υπέδειξε στα εξεταζόμενα υποκείμενα να φανταστούν τον χάρτη και να μεταβούν από μια περιοχή σε μια άλλη φωτίζοντας διαδοχικά τις δύο περιοχές. Το αποτέλεσμα του πειράματος ήταν να μην παρατηρηθεί η αναλογική σχέση μεταξύ του χρόνου μετάβασης και της διανυόμενης απόστασης.

Το πείραμα μοιάζει τετριμμένο, γιατί ο πειραματιστής υπαγορεύει στο υποκείμενο τον τρόπο μετάβασης από το ένα σημείο του χάρτη στο άλλο και το υποκείμενο απλώς επαναλαμβάνει νοητικά αυτήν την ενέργεια. Αλλά αυτό ακριβώς θέλει να τονίσει ο Pylyshyn. Αν ο πειραματιστής απαιτήσει η μετάβαση να γίνει μέσω σάρωσης του ενδιάμεσου χώρου τότε ο χρόνος μετάβασης είναι ανάλογος της καλυπτόμενης απόστασης. Αν πάλι απαιτήσει η μετάβαση να γίνει με το άνοιγμα ενός προβολέα, η μετάβαση είναι ακαριαία. Το υποκείμενο ακολουθεί τις εντολές του πειραματιστή κάνοντας χρήση της κοινότητας γνώσης που διαθέτει για τον κόσμο. Αποτελεί κοινότητα γνώση, για παράδειγμα, ότι όταν κινούμαστε με σταθερή ταχύτητα η κάλυψη μεγαλύτερης απόστασης απαιτεί περισσότερο χρόνο. Το υποκείμενο γνωρίζει, λοιπόν, από την πείρα του, ότι αν κοιτούσε ένα σημείο του χάρτη και έπρεπε να "σύρει" το βλέμμα του μέχρι ένα άλλο σημείο θα απαιτούνταν χρόνος ανάλογος της απόστασης των δύο σημείων. Γι' αυτό όταν φαντάζεται το ίδιο πράγμα προσπαθεί να μιμηθεί το τι θα συνέβαινε στην περίπτωση της κανονικής αντίληψης.

Η επισήμανση αυτή μπορεί να γίνει ακόμα πιο φανερή αν εξετάσουμε το πείραμα της νοητικής περιστροφής. Αν φανταστούμε ότι το προς περιστροφή αντικείμενο είναι πολύ βαρύ, τότε αυτό δεν μπορούμε να το περιστρέψουμε με ομαλή ταχύτητα: κατά την έναρξη της περιστροφής απαιτείται κάποιος χρόνος μέχρι να φτάσει σε μια σταθερή ταχύτητα, το ίδιο και για να σταματήσει η κίνηση²⁴⁰. Όμως το βάρος δεν είναι ένα εγγενές χαρακτηριστικό των εικόνων: η φωτογραφία μιας ράβδου χρυσού δεν περιστρέφεται δυσκολότερα από τη φωτογραφία ενός φτερού! Και από αυτό το παράδειγμα γίνεται φανερό ότι οι διαφορές στους παρατηρούμενους χρόνους μετακίνησης ή περιστροφής εξαρτώνται από το είδος του νοητικού περιεχομένου και όχι από το είδος της (υποτιθέμενης) υποκείμενης αναπαριστασιακής μορφής.

Όμως, ακόμα κι αν δεχόμασταν το επιχείρημα του Kosslyn ότι οι παρατηρούμενες χρονικές διαφορές οφείλονται στο είδος της αναπαριστασιακής μορφής που υπόκειται της εμπειρίας ειδώλων στη φαντασία πάλι δημιουργούνται προβλήματα. Το πρώτο πρόβλημα είναι τι

²⁴⁰ Haugeland 1985, σελ. 227.

ακριβώς εννοεί; Πρόκειται για κυριολεκτικές εικόνες στον εγκεφαλικό φλοιό που σαρώνονται και περιστρέφονται ή πρόκειται για λειτουργικές εικόνες που υφίστανται σε έναν λειτουργικό χώρο; Όπως είδαμε στην ενότητα 5.2.3.1, ο Kosslyn, σε ένα πρώτο επίπεδο, υποστηρίζει τη δεύτερη εκδοχή. Τα πειράματα νοητικής σάρωσης και νοητικής περιστροφής δείχνουν ότι τα νοητικά είδωλα "χρησιμοποιούν ένα 'λειτουργικό χώρο' για να αναπαραστήσουν τον πραγματικό χώρο" (Kosslyn 1994, σελ. 7).

Αν δεχτούμε αυτή την εκδοχή, τίθεται ένα δεύτερο ερώτημα: υπό ποια έννοια ένας λειτουργικός χώρος μπορεί να έχει εγγενώς τις ιδιότητες του φυσικού χώρου; Όπως προαναφέραμε, "εγγενώς" εδώ σημαίνει ότι πρόκειται για ιδιότητες του αναπαραστασιακού φορέα, πέρα και ανεξάρτητα από το αναπαραστασιακό περιεχόμενο. Για παράδειγμα, εγγενής ιδιότητα της συμβολικής αναπαράστασης "πράσινος χάρακας 30 εκατοστών" (ως έκτυπου) είναι ότι έχει μαύρο χρώμα (παρότι αναπαριστά έναν πράσινο χάρακα) ή ότι έχει μήκος 5cm (παρότι αναπαριστά έναν χάρακα 30 cm). Όμως τι εγγενείς ιδιότητες μπορεί να έχει μια λειτουργική εικόνα; Προκειμένου να απαντήσουμε σε αυτό το ερώτημα πρέπει να εξετάσουμε υπό ποια μορφή μπορεί να υπάρχει μια λειτουργική εικόνα. Ο Kosslyn επιχειρεί να συγκεκριμενοποιήσει το τι είναι μια λειτουργική εικόνα με τη βοήθεια της οπτικής ενδιάμεσης μνήμης. Η οπτική ενδιάμεση μνήμη αποτελεί το μέσο εντός του οποίου εμφανίζονται οι λειτουργικές εικόνες. Ο Kosslyn (1980), σε μια πρώτη προσέγγιση, προτείνει να φανταστούμε αυτό το μέσο ως μια οθόνη καθοδικών ακτίνων που αποτελείται από ένα μεγάλο αριθμό μικροσκοπικών τετραγωνιδίων ή κυττάρων. Ο Tye (1991, σελ. 34) ερμηνεύοντας τον Kosslyn λέει ότι "ο σχηματισμός που εμφανίζεται από την τοποθέτηση κουκίδων σε αυτά τα κύτταρα είναι εικονιστικός, επειδή έχει χωρικά χαρακτηριστικά που αντιστοιχούν στα χωρικά χαρακτηριστικά του αναπαριστώμενου αντικειμένου. Συγκεκριμένα, κουκίδες στη μήτρα-πλέγμα [matrix] αναπαριστούν σημεία της επιφάνειας του αντικειμένου, ενώ σχετικές αποστάσεις και γεωμετρικές σχέσεις μεταξύ των κουκίδων συμφωνούν προς τις ίδιες σχέσεις μεταξύ σημείων του αντικειμένου". Η μήτρα που σχηματίζεται κατ' αυτόν τον τρόπο δεν πρέπει να νοηθεί ότι εκτείνεται κατ' ανάγκη σε μια ενιαία περιοχή. Για παράδειγμα, δύο κύτταρα που απεικονίζουν δύο γειτονικά σημεία του αντικειμένου μπορούν να βρίσκονται σε διαφορετικές περιοχές του εγκεφάλου, ενώ δύο κύτταρα που απεικονίζουν δύο απομακρυσμένα μεταξύ τους σημεία του αντικειμένου μπορούν να βρίσκονται το ένα δίπλα στο άλλο στον εγκέφαλο. Υπ' αυτή την έννοια, λοιπόν, ο Kosslyn ισχυρίζεται ότι τα νοητικά είδωλα δεν χρειάζεται να είναι κυριολεκτικές εικόνες που εκτείνονται σε μια ενιαία επιφάνεια, αλλά λειτουργικές εικόνες, δηλαδή

σηματισμοί που λειτουργούν ως εάν να ήταν εικόνες μιας ενιαίας επιφάνειας²⁴¹.

Το πρόβλημα που τίθεται μετά απ' όλα αυτά είναι ότι οι λειτουργικές εικόνες δεν φαίνεται τελικά να έχουν κάποιες *εγγενείς χωρικές ιδιότητες που αντιστοιχούν στις χωρικές ιδιότητες του εικονιζόμενου αντικειμένου*. Θα μπορούσε, ωστόσο, κανείς να αντιτείνει ότι, αν φανταστούμε τις λειτουργικές εικόνες ως παραμορφωμένες εικόνες, δύο κύτταρα μπορούν να είναι απομακρυσμένα μεταξύ τους και παρόλα αυτά να αναπαριστούν δύο γειτονικά σημεία του αντικειμένου. Για παράδειγμα, αν τυπώσουμε μια εικόνα πάνω σε ένα λάστιχο και κατόπιν το τεντώσουμε προς μια διεύθυνση, τότε ορισμένα σημεία που αναπαριστούν γειτονικά σημεία του αντικειμένου μπορούν να απομακρυνθούν πολύ μεταξύ τους, ενώ άλλα που αναπαριστούν απομακρυσμένα σημεία μπορούν να έρθουν πολύ κοντά. Όμως, μια κατ' αυτόν τον τρόπο κατασκευασμένη εικόνα δεν είναι μια *λειτουργική εικόνα* αλλά απλώς μια *παραμορφωμένη* εικόνα, δηλαδή μια εικόνα που διατηρεί τη τοπολογία της αρχικής εικόνας. Πιο συγκεκριμένα, δεν αποτελεί λειτουργική εικόνα γιατί η παραμορφωμένη εικόνα εκτείνεται σε μια *ενιαία επιφάνεια* και οι σχετικές θέσεις των σημείων της παραμορφωμένης εικόνας διατηρούν την ίδια *διάταξη* με τις σχετικές θέσεις των αντίστοιχων σημείων του αντικειμένου.

Τι θα μπορούσε, όμως, να αποτελεί λειτουργική εικόνα; Ο ίδιος ο Kosslyn (1994, σελ. 5), αναφέρει το ακόλουθο παράδειγμα:

"Μια απεικονιστική αναπαράσταση είναι ένας τύπος εικόνας, ο οποίος καθορίζει τις θέσεις και τις τιμές συστημάτων [configurations] σημείων σε ένα χώρο. Για παράδειγμα, το σχέδιο μιας μπάλας πάνω σ' ένα κουτί θα ήταν μια απεικονιστική αναπαράσταση. Ο χώρος εντός του οποίου εμφανίζονται τα σημεία δεν χρειάζεται να είναι φυσικός, όπως πάνω σ' αυτή τη σελίδα, αλλά μπορεί να είναι όπως ένας πίνακας [array] στον υπολογιστή ο οποίος καθορίζει χωρικές σχέσεις καθαρά λειτουργικά. Δηλαδή, οι φυσικές θέσεις μέσα στον υπολογιστή για κάθε σημείο σε έναν πίνακα δεν είναι οι ίδιες διευθετημένες σε πίνακα· είναι μόνο δυνάμει του τρόπου ανάγνωσης και επεξεργασίας αυτών των πληροφοριών που φτάνει να λειτουργούν ως εάν να ήταν διευθετημένες σε ένα πίνακα"

Αν, όμως, τελικά ο τρόπος που είναι διευθετημένα τα κύτταρα μιας λειτουργικής εικόνας δεν μπορεί να αντιστοιχηθεί στον τρόπο που είναι διευθετημένα τα σημεία του αναπαριστώμενου, τότε η απόσταση μεταξύ δύο σημείων του αναπαριστώμενου δεν μπορεί να εξαχθεί από τη *φυσική απόσταση* μεταξύ των αντίστοιχων κυττάρων. Ούτε μπορεί να

²⁴¹ "Αυτές οι περιοχές συνιστούν μια λειτουργικά χαρακτηριζόμενη δομή· δεν χρειάζεται να είναι ανατομικά γειτονικές" (Kosslyn 1994, σελ. 86).

δικαιολογηθεί η παρατηρούμενη αναλογία μεταξύ χρόνου μετάβασης και διανυόμενης απόστασης μεταξύ σημείων του αναπαριστώμενου με αναδρομή στις αντίστοιχες φυσικές αποστάσεις των αντίστοιχων κυττάρων²⁴². Γι' αυτό, όπως επισημαίνει ο Pyllyshyn (2001): "Η χωρικότητα μιας κήτρας, ή οποιουδήποτε άλλου τύπου λειτουργικού χώρου, πρέπει να καθοριστεί συμβατικά ή να υποθεθεί πέρα και πάνω από οποιαδήποτε εγγενή ιδιότητα της μορφής της αναπαράστασης". Κατά συνέπεια, ένας τέτοιος εξωγενής περιορισμός μπορεί να εφαρμοστεί σε οποιαδήποτε αναπαραστασιακή μορφή. Άρα, ακόμα κι αν δεχόμασταν ότι υπάρχουν λειτουργικές εικόνες, η υπόθεση αυτή δεν εξηγεί καλύτερα από τις συμβολικές αναπαραστάσεις τις χρονικές διαφορές που παρατηρούνται κατά τη σάρωση διαφορετικών αποστάσεων μεταξύ τμημάτων των εικόνων που παρουσιάζονται στη φαντασία²⁴³.

Ο μόνος τρόπος να εξηγηθούν οι παρατηρούμενες διαφορές στο χρόνο σάρωσης με αναδρομή στις εγγενείς ιδιότητες των νοητικών ειδώλων είναι να δεχτούμε ότι τα εν λόγω είδωλα δεν είναι λειτουργικές εικόνες αλλά κυριολεκτικές εικόνες - με την έννοια ότι σχηματίζονται κυριολεκτικά σε τμήματα της εγκεφαλικής επιφάνειας. Αυτή η ιδέα, όπως είδαμε, παρουσιάζει άλλου τύπου προβλήματα. Θα επανέλθουμε, ωστόσο, στο συγκεκριμένο ζήτημα στην ενότητα 5.2.5.2.

5.2.5 Νευροφυσιολογικά τεκμήρια

Στο τελευταίο του βιβλίο, ο Kosslyn (1994) διακηρύσσει το τέλος της διαμάχης για τα νοητικά είδωλα και τη νίκη των υποστηρικτών αυτού του είδους αναπαράστασης. Ο καθοριστικός λόγος που οδηγεί τον Kosslyn σε αυτή τη διακήρυξη είναι η εξέταση των τεκμηρίων που προέρχονται από τη νευροφυσιολογική έρευνα. Τα νευροφυσιολογικά τεκμήρια, εν προκειμένω, χρησιμοποιούνται για την υποστήριξη των ακόλουθων δύο θέσεων:

- A1: οι ίδιες περιοχές του εγκεφάλου χρησιμοποιούνται για την αντίληψη και για τη χρήση ειδώλων στη φαντασία²⁴⁴
- A2: στις περιοχές του εγκεφάλου που βασίζεται η αντίληψη σχηματίζονται εικόνες, δηλαδή χάρτες που διατηρούν τη τοπολογία του αμφιβληστροειδή (retinotopic maps)²⁴⁵.

²⁴² Είναι ενδιαφέρον ότι και ο ίδιος ο Kosslyn σε κάποιο σημείο του τελευταίου του βιβλίου φτάνει σε μια παρεμφερή διαπίστωση: "...η οπτική ενδιάμεση μνήμη δεν είναι ούτε ομογενής ούτε ισότροπος και παρόλα αυτά ο χρόνος σάρωσης ενός αντικείμενου που φανταζόμαστε αυξάνει γραμμικά με την απόσταση που σαρώνεται πάνω στο αντικείμενο" (Kosslyn 1994, σελ. 344).

²⁴³ Βλ. Pyllyshyn 2001.

²⁴⁴ Βλ., ενδεικτικά, Kosslyn 1990, σελ. 78: "Νευροφυσιολογικά δεδομένα υποστηρίζουν ... την ιδέα ότι η επιθεώρηση νοητικών ειδώλων πραγματοποιείται από τους ίδιους μηχανισμούς που χρησιμοποιούνται για την αντίληπτική αναγνώριση".

Είναι προφανές ότι οι δύο παραπάνω θέσεις δεν είναι ισοδύναμες: η πρώτη δεν συνεπάγεται τη δεύτερη ούτε και το αντίστροφο. Θα ξεκινήσουμε με την εξέταση της θέσης Α1.

5.2.5.1 Χρήση κοινών εγκεφαλικών μηχανισμών κατά την αντίληψη και τη χρήση νοητικών ειδώλων

Η αλματώδης πρόοδος στις τεχνικές απεικόνισης του εγκεφάλου είναι, ίσως, ο κυριότερος λόγος για τη συσσώρευση μιας πληθώρας ευρημάτων που υποστηρίζουν τη θέση Α1. Από τη χρήση των συγκεκριμένων τεχνικών προκύπτει ότι οι ίδιες εγκεφαλικές περιοχές ενεργοποιούνται τόσο κατά την αντίληψη όσο και κατά τη χρήση ειδώλων στη φαντασία. Επίσης, νευροφυσιολογικά τεκμήρια υπέρ της θέσης Α1 έχουν συσσωρευτεί και από την εξέταση των ικανοτήτων υποκειμένων με νευροφυσιολογικές βλάβες στον εγκέφαλο. Χαρακτηριστική είναι η περίπτωση ενός ασθενούς με περιορισμένο οπτικό πεδίο (tunnel vision) εξ' αιτίας εγχείρησης μονόπλευρης ινιακής λοβοτομής ο οποίος εμφάνισε περιορισμό και στο εύρος του πεδίου των ειδώλων στη φαντασία (tunnel imagery) (βλ. Farah, Soso & Dasheiff 1992)²⁴⁶.

Όλα αυτά τα ευρήματα παρουσιάζουν από μόνα τους πολύ μεγάλο θεωρητικό ενδιαφέρον και έχουν προκαλέσει ήδη μια αρκετά ζωηρή διαμάχη για τον εάν όντως αποδεικνύουν την ορθότητα της παραδοχής Α1²⁴⁷. Ωστόσο, εδώ δεν θα επεισέλθουμε στο συγκεκριμένο ζήτημα γιατί ακόμα κι αν δεχθούμε την ισχύ της Α1, τίποτε δεν συνεπάγεται όσον αφορά το είδος των αναπαραστάσεων που υποτίθεται ότι χρησιμοποιούνται από τους σχετικούς εγκεφαλικούς μηχανισμούς. Η απλή διαπίστωση ότι μια ή περισσότερες περιοχές του εγκεφάλου ενεργοποιούνται τόσο κατά την αντίληψη όσο και κατά την χρήση ειδώλων στη φαντασία σίγουρα δεν αποκαλύπτει τίποτε για τον τρόπο που λειτουργούν αυτές οι περιοχές. Η λειτουργία τους θα μπορούσε κάλλιστα να βασίζεται είτε σε εικονιστικές αναπαραστάσεις, είτε σε συμβολικές αναπαραστάσεις, ή να μην βασίζεται καν σε αναπαραστάσεις²⁴⁸.

²⁴⁵ Βλ., ενδεικτικά, Kosslyn 1994, σελ. 67: "...περίπου 15 από τις οπτικές περιοχές ... είναι χωρικά οργανωμένες ... αυτές οι περιοχές σχεδόν διατηρούν τη τοπική γεωμετρία του αμφιβληστροειδή".

²⁴⁶ Για μια σύνοψη των σημαντικότερων νευροφυσιολογικών ευρημάτων βλ. Kosslyn 1994, σελ. 58-60.

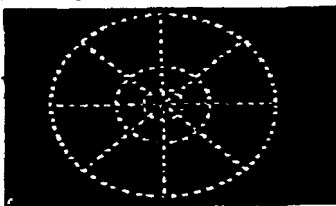
²⁴⁷ Υπάρχουν αρκετοί που αμφισβητούν τη δυνατότητα συναγωγής της Α1 από τα προαναφερθέντα τεκμήρια. Βλ., χαρακτηριστικά, Pylyshyn 2001.

²⁴⁸ Επίσης, κατά τη φαντασία χρησιμοποιούνται όχι μόνο αντιληπτικοί μηχανισμοί αλλά και κινητικοί. Τίθεται λοιπόν το ερώτημα τι θα μπορούσαμε να συμπεράνουμε από αυτό το γεγονός για τις υποτιθέμενες νοητικές αναπαραστάσεις που χρησιμοποιούνται; Ποιες θα μπορούσαν να είναι οι κατάλληλες αναπαραστάσεις για το κινητικό σύστημα; Αλλά το ίδιο πρόβλημα τίθεται και αν φύγουμε από το οπτικό σύστημα και εξετάσουμε τα συστήματα άλλων

5.2.5.2 Η ανακάλυψη εγκεφαλικών χαρτών που διατηρούν την τοπολογία του αμφιβληστροειδή

Ας έρθουμε τώρα στην εξέταση της θέσης A2, σύμφωνα με την οποία κατά την αντίληψη σχηματίζονται εικόνες και πιο συγκεκριμένα χάρτες που διατηρούν τη τοπολογία του αμφιβληστροειδή στις περιοχές του εγκεφάλου στις οποίες βασίζεται αυτή η λειτουργία. Ο Kosslyn και αρκετοί άλλοι ερευνητές θεωρούν ότι η συγκεκριμένη θέση έχει πλέον αποδειχθεί από τη σύγχρονη νευροφυσιολογική έρευνα. Πολλά από τα σχετικά νευροφυσιολογικά ευρήματα παρουσιάζονται από τον Kosslyn στο τελευταίο του βιβλίο. Θα αναφέρουμε, ενδεικτικά, το χαρακτηριστικότερο από αυτά. Πρόκειται για ένα εύρημα των Tootell, Silverman, Switkes & De Valois (1982) οι οποίοι, θυσιάζοντας έναν πίθηκο στο βωμό της επιστήμης, ανακάλυψαν ότι στην περιοχή V1 του εγκεφάλου του²⁴⁹ σχηματίζεται ένας χάρτης που διατηρεί την τοπολογία του αμφιβληστροειδή. Πιο συγκεκριμένα, το πείραμα έγινε ως εξής: Καταρχάς, ο πίθηκος εκπαιδεύτηκε να κοιτά στο κέντρο ενός φωτεινού σχηματισμού με σταθερή και χαρακτηριστική γεωμετρία. Όταν πλέον έμαθε να εκτελεί αυτό το έργο και ενώ κοιτούσε το σχηματισμό τού έγινε ένεση με διάλυμα ραδιενεργού ζάχαρης. Αυτό έγινε γιατί οι περιοχές του εγκεφάλου που ενεργοποιούνται περισσότερο, αιματώνονται και περισσότερο και, κατά συνέπεια, αποκτούν μεγαλύτερη ραδιενέργεια. Στη συνέχεια, το ζώο θανατώθηκε και μετρήθηκε η ραδιενέργεια στον εγκέφαλό του απ' όπου βρέθηκε ότι στην περιοχή V1 είχε δημιουργηθεί ο ίδιος σχηματισμός ελαφρώς παραμορφωμένος.

Φωτεινός σχηματισμός που βλέπει ο πίθηκος



Σχηματισμός στην περιοχή V1 του εγκεφάλου του πιθήκου



τροπικότητας όπως το απτικό ή το οσφρητικό. Πώς θα μπορούσε να βοηθεί σε αυτά τα συστήματα η εικονιστική αναπαράσταση τους;

²⁴⁹ Από ανεξάρτητες έρευνες είναι γνωστό ότι η συγκεκριμένη περιοχή εμπλέκεται στα πρώιμα στάδια της αντίληψης.

Μετά το εύρημα των Tootell, Silverman, Switkes & De Vairois, το ερώτημα που τίθεται είναι αν υφίσταται πλέον κάποιος λόγος να αμφισβητήσει κανείς την ισχύ της θέσης A2. Εδώ θα πρέπει να είμαστε προσεκτικοί. Το προαναφερθέν εύρημα δείχνει ότι σε μια περιοχή του εγκεφάλου σχηματίζονται χάρτες που διατηρούν την τοπολογία του αμφιβληστροειδή. Ωστόσο, αυτοί οι χάρτες δεν ισοδυναμούν με εικόνες όπως ισχυρίζεται ο Kosslyn (1994, σελ. 13): "Οι οπτικές περιοχές που είναι χαρτογραφημένες κατά τρόπον ώστε να διατηρούν την τοπολογία του αμφιβληστροειδή ... αναπαριστούν τις πληροφορίες απεικονιστικά με την πιο κυριολεκτική έννοια: [εδώ] δεν υπάρχει καμία ανάγκη να μιλήσουμε για έναν αφηρημένο λειτουργικό χώρο όπως αυτόν που ορίστηκε με τον πίνακα μέσα στον υπολογιστή". Καταρχάς, μια εικόνα, εκτός από την τοπολογία ενός σχήματος μεταφέρει πολλές άλλες πληροφορίες, όπως για παράδειγμα τις χρωματικές αποχρώσεις και τη φωτεινότητα του σχήματος. Επίσης, όταν αντιλαμβανόμαστε τον κόσμο, τον αντιλαμβανόμαστε ως τρισδιάστατο. Τίθεται, λοιπόν, το ερώτημα πώς μπορεί να εξηγηθεί αυτό το φαινομενολογικό δεδομένο με τη βοήθεια των εγκεφαλικών χαρτών;

Ακόμα, όμως κι αν παραβλέψουμε τα παραπάνω προβλήματα και δεχθούμε τη θέση ότι στο εγκέφαλο σχηματίζονται κυριολεκτικά εικόνες, το ερώτημα που τίθεται είναι τι ακριβώς εξηγεί αυτή η ανακάλυψη όσον αφορά τη λειτουργία του μηχανισμού με τον οποίον αντιλαμβανόμαστε τον κόσμο; Υπό ποια έννοια, δηλαδή, μαθαίνουμε κάτι για το πώς πραγματοποιείται η αντίληψη.

Μια εικόνα προκειμένου να κατανοηθεί χρειάζεται να γίνει αντιληπτή. Οι εξωτερικές εικόνες γίνονται αντιληπτές με τη βοήθεια των ματιών μας. Με ποια όργανα, όμως, γίνονται αντιληπτές οι εικόνες που σχηματίζονται στον εγκεφαλικό φλοιό; Έχει νόημα να μιλάμε για το "μάτι του νου" ή, ακριβέστερα, το "μάτι του εγκεφάλου"²⁵⁰; Με το να ισχυριζόμαστε ότι κατά την αντίληψη αναπαράγεται εντός του κρανίου η πραγματικότητα υπό μορφή εικόνων δεν εξηγούμε τίποτα, απλώς μεταθέτουμε το πρόβλημα στο επίπεδο της αντίληψης των εσωτερικών αυτών εικόνων.

Όσα είπαμε μέχρις εδώ στην παρούσα ενότητα αφορούν την αντίληψη. Ο Kosslyn, ωστόσο, κάνει ένα βήμα παραπάνω και ισχυρίζεται ότι και κατά την χρήση ειδώλων στη φαντασία σχηματίζονται εικόνες στον εγκεφαλικό ιστό. Μάλιστα, στο τελευταίο του βιβλίο, εξελίσσοντας τη θεωρία του, τροποποιεί την έννοια της οπτικής ενδιάμεσης μνήμης ως εξής:

"Η οπτική ενδιάμεση μνήμη στην προηγούμενη εκδοχή της θεωρίας μοντελοποιήθηκε ως πίνακας μέσα σε υπολογιστή, ο οποίος ήταν ανισότροπος αλλά ομογενής. Η τρέχουσα αντίληψη είναι ότι η

²⁵⁰ "Η για το "αυτί του εγκεφάλου" και "τη μύτη του εγκεφάλου" αν πρόκειται για είδωλα της ακοής και της όσφρησης.

οπτική ενδιάμεση μνήμη αντιστοιχεί σε ένα σύνολο τοπογραφικά χαρτογραφημένων οπτικών περιοχών του εγκεφάλου οι οποίες είναι ανισότροπες και μη-ομογενείς" (Kosslyn 1994, σελ. 388).

Για να ισχύει η θέση ότι κατά τη χρήση ειδώλων στη φαντασία σχηματίζονται εικόνες στον εγκεφαλικό ιστό θα πρέπει προφανώς να ισχύουν τόσο η Α1 όσο και η Α2. Με άλλα λόγια, πρέπει να ισχύει ότι κατά τη χρήση ειδώλων στη φαντασία χρησιμοποιούνται οι ίδιοι εγκεφαλικοί μηχανισμοί με εκείνους που χρησιμοποιούνται κατά την αντίληψη και ότι οι εγκεφαλικοί μηχανισμοί της αντίληψης βασίζονται στη δημιουργία και επεξεργασία κυριολεκτικών εικονιστικών αναπαραστάσεων. Εκτός όμως από αυτό θα πρέπει να ισχύει ότι οι εν λόγω μηχανισμοί "λειτουργούν" με τον ίδιο τρόπο τόσο κατά την αντίληψη όσο και κατά τη χρήση ειδώλων στη φαντασία.

5.3 Τι συνιστά εικόνα;

Απ' όσα αναφέραμε, λοιπόν, στις ενότητες 5.2.3 - 5.2.5 καθίσταται φανερό ότι η υπόθεση περί ύπαρξης νοητικών ειδώλων δημιουργεί ανυπερβλήτες δυσκολίες και κάθε άλλο παρά βοηθά στην εξήγηση της λειτουργίας της αντίληψης και της φαντασίας. Ωστόσο, πριν κλείσουμε αυτό το κεφάλαιο θα ασχοληθούμε και με ένα γενικότερο ερώτημα που αφορά τη θεωρία περί νοητικών ειδώλων αλλά δεν περιορίζεται σε αυτήν. Πρόκειται για το ερώτημα πότε μια αναπαράσταση χαρακτηρίζεται εικονιστική και σε τι διαφέρει από τις συμβολικές αναπαραστάσεις; Αυτό το ζήτημα, προφανώς, αφορά τα θεμέλια της διαμάχης για τα νοητικά είδωλα: Είτε υποστηρίζει κανείς ότι οι νοητικές αναπαραστάσεις που υπόκεινται της εμπειρίας ειδώλων στη φαντασία είναι εικονιστικές είτε ότι είναι συμβολικές, πρέπει με κάποιον τρόπο να γνωρίζει σε τι διαφέρουν αυτά τα δύο είδη αναπαραστάσεων. Ωστόσο, το συγκεκριμένο ζήτημα κάθε άλλο παρά έχει τύχει ενδελεχούς μελέτης μέσα στο πλαίσιο της διαμάχης για τα νοητικά είδωλα. Οι δύο αντιμαχόμενες πλευρές, λίγο ως πολύ, υιοθετούν την κοινότοπη αντίληψη περί εικόνων και συμβόλων: μια εικόνα ομοιάζει με ό,τι αναπαριστά, ενώ ένα (ατομικό) σύμβολο σχετίζεται αυθαίρετα με το αναπαριστώμενο.

5.3.1 Η κυρίαρχη άποψη

Ας εξετάσουμε, όμως, πιο αναλυτικά μέσα από τα λεγόμενα του Kosslyn πώς ακριβώς αρθρώνεται η προαναφερθείσα αντίληψη. Όσον αφορά τις εικονιστικές αναπαραστάσεις ο Kosslyn υποστηρίζει ότι "μεταφέρουν νόημα μέσω της ομοιότητάς τους προς το αντικείμενο, με τμήματα της αναπαράστασης να αντιστοιχούν σε τμήματα του αντικειμένου" (Kosslyn 1994, σελ. 5). Επίσης, σε ένα άλλο βιβλίο του επισημαίνει ότι "το πρωταρχικό χαρακτηριστικό των αναπαραστάσεων αυτής της μορφής είναι ότι κάθε τμήμα της αναπαράστασης πρέπει να αντιστοιχεί σε ένα τμήμα του αντικειμένου έτσι ώστε οι σχετικές αποστάσεις μεταξύ δύο τμημάτων του αντικειμένου να διατηρούνται στις

αποστάσεις μεταξύ των αντίστοιχων τμημάτων της αναπαράστασης" (Kosslyn 1981, σελ. 50). Αντίθετα, οι συμβολικές αναπαραστάσεις είναι "νοητικές γλωσσικές προτάσεις [sentences]" που αποτελούνται από τμήματα τα οποία "μεταφέρουν νόημα μέσω αυθαιρέτα αποδιδόμενων συσχετισμών" (Kosslyn 1994, σελ. 5). Έτσι, οι εικονιστικές αναπαραστάσεις διαφέρουν από τις συμβολικές κατά το ότι "η συσχέτιση ανάμεσα σε μια αναπαράσταση και σε αυτό που αναπαριστά δεν είναι αυθαίρετη" οι απεικονίσεις 'ομοιάζουν' με το αναπαριστώμενο αντικείμενο ή αντικείμενα. Δηλαδή, (α) κάθε τμήμα της αναπαράστασης πρέπει να αντιστοιχεί σε ένα ορατό τμήμα του αντικειμένου ή των αντικειμένων, και (β) οι αναπαριστώμενες 'αποστάσεις' μεταξύ τμημάτων της αναπαράστασης πρέπει να κατοπτρίζουν τις αποστάσεις μεταξύ των αντίστοιχων τμημάτων του πραγματικού αντικειμένου (όπως εμφανίζονται από μια συγκεκριμένη οπτική γωνία)" (Kosslyn 1990, σελ. 86).

Από τα παραπάνω παραθέματα βλέπουμε, λοιπόν, ότι και ο Kosslyn αποδέχεται την αντίληψη ότι η κύρια διαφορά μεταξύ εικονιστικών και συμβολικών αναπαραστάσεων έγκειται στον τρόπο που το αναπαραστασιακό περιεχόμενο αντιστοιχίζεται στην αναπαραστασιακή μορφή. Η υιοθέτηση αυτής της κοινότοπης αντίληψης οδηγεί σε μια θέση που, εκ πρώτης όψεως, φαίνεται ως άμεση συνέπεια: η διαφορά μεταξύ εικονιστικών και συμβολικών αναπαραστάσεων πρέπει να αναζητηθεί στις διαφορές των αντίστοιχων αναπαραστασιακών μορφών. Δεν είναι τυχαίο, εξάλλου, το ότι και ο ίδιος ο Kosslyn θεωρεί ότι η διαμάχη για τα νοητικά είδωλα αφορά την αναπαραστασιακή μορφή και όχι το αναπαραστασιακό περιεχόμενο. Θυμίζουμε ένα χαρακτηριστικό απόσπασμα που προαναφέραμε:

"Η μακρόχρονη διαμάχη περί της χρήσης νοητικών ειδώλων [mental imagery] αφορά τη μορφή των αναπαραστάσεων που υπόκεινται της συνειδητής εμπειρίας του 'βλέπεις με τα μάτια του νου', και όχι ... το γεγονός ότι οι άνθρωποι έχουν τέτοιες εμπειρίες" (Kosslyn 1984, σελ. 104 –η έμφαση προστέθηκε)

Ωστόσο, η θέση ότι η διαφορά των εικονιστικών αναπαραστάσεων από τις συμβολικές πρέπει να αναζητηθεί στη διαφορά των αντίστοιχων αναπαραστασιακών μορφών δεν συνάγεται κατ' ανάγκη από τη θέση ότι η διαφορά μεταξύ εικονιστικών και συμβολικών αναπαραστάσεων έγκειται στον τρόπο που το αναπαραστασιακό περιεχόμενο αντιστοιχίζεται στην αναπαραστασιακή μορφή. Αυτό το βήμα καθίσταται αυτονόητο μόνο αν δεχθεί κανείς ότι τόσο οι εικόνες όσο και τα σύμβολα αναπαριστούν το *ίδιο είδος* περιεχομένου.

Η συγκεκριμένη παραδοχή δεν υποστηρίζεται ρητά από τους εμπλεκόμενους στη διαμάχη για τα νοητικά είδωλα, αλλά ούτε και απορρίπτεται. Υπάρχει μάλλον ως λανθάνουσα παραδοχή που μόνο έμμεσα μπορεί να συναχθεί. Για παράδειγμα, τόσο ο Kosslyn όσο και ο Pylyshyn, δηλαδή οι δύο κύριοι αντίπαλοι στη διαμάχη, υποστηρίζουν ότι δεν διαφωνούν ως προς το ότι στο φαινομενολογικό επίπεδο εμφανίζονται

είδωλα στη φαντασία. Με άλλα λόγια, συμφωνούν ως προς το είδος του αναπαραστασιακού περιεχομένου που δίδεται στη φαντασία. Όσες δε φορές φτάνουν να περιγράψουν το συγκεκριμένο αναπαραστασιακό περιεχόμενο μιλούν με όρους αντικειμένων και ιδιοτήτων. Για παράδειγμα, ο Kosslyn αναφερόμενος στις εικονιστικές αναπαραστάσεις δηλώνει ότι "ομοιάζουν" προς το αναπαραριστώμενο αντικείμενο ή αντικείμενα" (Kosslyn 1990, σελ. 86 -η έμφαση προστέθηκε). Αντίστοιχες αναφορές υπάρχουν και για τον τρόπο που εικονιστικές αναπαραστάσεις αναπαριστούν ιδιότητες²⁵¹. Όμως τα αντικείμενα και οι ιδιότητες είναι τα συστατικά του προτασιακού περιεχομένου, δηλαδή του περιεχομένου που κατεξοχήν εκφράζεται με συμβολικές αναπαραστάσεις.

Αν δεχθούμε αυτή την ανάλυση και τη θέση ότι το περιεχόμενο που τόσο οι εικόνες όσο και τα σύμβολα αναπαριστούν είναι προτασιακό, τότε ο ισχυρισμός του Pylyshyn και των συμμάχων του ότι το περιεχόμενο των ειδώλων της φαντασίας μπορεί να αναπαρασταθεί με συμβολικές αναπαραστάσεις καθίσταται τετριμμένα αληθές. Αν η διαμάχη για τα νοητικά είδωλα αφορά απλώς το ποιος είναι ο καταλληλότερος φορμαλισμός για την αναπαράσταση του προτασιακού περιεχομένου, τότε το βάρος της απόδειξης πραγματικά πέφτει στους υποστηρικτές των εικονιστικών αναπαραστάσεων. Για παράδειγμα, με τις συμβολικές αναπαραστάσεις μπορούμε, μεταξύ άλλων, να αναφερθούμε σε ειδή αντικειμένων (π.χ. το είδος τρίγωνο), σε αφηρημένα αντικείμενα (π.χ. αριθμός), σε ανύπαρκτα αντικείμενα, σε παρελθούσες ή επικείμενες καταστάσεις πραγμάτων, σε λογικές ιδιότητες (π.χ. γενίκευση), σε λογικές έννοιες (π.χ. σύζευξη, διάζευξη, άρνηση) ή, τέλος, σε μεμονωμένες ιδιότητες αντικειμένων. Όλα αυτά αποτελούν είδη προτασιακού περιεχομένου ή συστατικών αυτού. Κανένα, όμως, από τα συγκεκριμένα είδη δεν μπορεί να εκφραστεί αποκλειστικά με εικονιστικές αναπαραστάσεις. Ας πάρουμε ενδεικτικά το παράδειγμα της αναπαράστασης μεμονωμένων ιδιοτήτων ενός αντικειμένου. Έστω ότι θέλουμε να αναπαραστήσουμε με εικονιστική αναπαράσταση το προτασιακό περιεχόμενο που αναπαριστά η γλωσσική πρόταση "Ο Πέτρος είναι χονδρός". Εκ πρώτης όψεως, φαίνεται να μην υπάρχει κανένα πρόβλημα. Θα μπορούσε κανείς να πει ότι αρκεί να φωτογραφίσουμε τον Πέτρο ή να τον απεικονίσουμε σε έναν πίνακα. Όμως αν δεχθούμε την άποψη ότι μια εικόνα του Πέτρου αναπαριστά την κατάσταση πραγμάτων ότι ο Πέτρος είναι χονδρός, τότε με ποιον τρόπο μπορούμε να αναπαραστήσουμε εικονιστικά την κατάσταση πραγμάτων ότι ο Πέτρος είναι ψηλός; Πάλι με μια εικόνα του Πέτρου; Σε αυτή την περίπτωση όμως

²⁵¹ Για παράδειγμα, στην ίδια σελίδα με αυτήν της προηγούμενης παραπομπής ο Kosslyn αναφέρεται στον τρόπο που μια εικόνα αναπαριστά τη (σχεσιακή) ιδιότητα "πάνω".

δεν είναι δυνατόν να αναπαραστήσουμε μια μεμονωμένη ιδιότητα του Πέτρου με τη βοήθεια εικονιστικών αναπαραστάσεων²⁵².

Βλέπουμε, λοιπόν, ότι πολλά και βασικά είδη προτασιακού περιεχομένου δεν μπορούν να εκφραστούν με εικονιστικές αναπαραστάσεις. Το μόνο είδος προτασιακού περιεχομένου που φαίνεται να είναι άμεσα αναπαραστάσιμο με εικονιστικές αναπαραστάσεις είναι καταστάσεις πραγμάτων που αφορούν μεμονωμένα αντικείμενα και το σύνολο των οπτικά διαπιστώσιμων ιδιοτήτων τους. Ένα τέτοιο περιεχόμενο, βέβαια, μπορεί να εκφραστεί με μια λίστα γλωσσικών προτάσεων. Σε αυτή την περίπτωση, λοιπόν, αν δεχθούμε την κοινότυπη αντίληψη περί εικόνων και συμβόλων, θα έχουμε δύο *ισοδύναμες* από άποψη περιεχομένου αναπαραστάσεις που παρόλα αυτά ανήκουν σε διαφορετικό αναπαραστασιακό είδος.

Από τα παραπάνω συνάγεται ότι αν κανείς δεχθεί την παραδοχή ότι οι εικονιστικές αναπαραστάσεις αποτελούν απλώς ένα διαφορετικό τρόπο αντιστοίχισης του προτασιακού περιεχομένου στην αναπαραστασιακή μορφή, τότε οι εικονιστικές αναπαραστάσεις σίγουρα υπολείπονται των συμβολικών αναπαραστάσεων. Σ' αυτή την περίπτωση, το δίκιο θα ήταν μάλλον προς τη πλευρά των υποστηρικτών των νοητικών συμβόλων.

Το δεύτερο συμπέρασμα που εξάγεται, αν κανείς δεχθεί την προαναφερθείσα παραδοχή, είναι ότι οι εικονιστικές αναπαραστάσεις δεν βοηθούν στην κατανόηση του τρόπου που η πρακτική γνώση δεξιοτήτων μπορεί ενδεχομένως να αναπαρασταθεί. Κατά συνέπεια, η κριτική που ασκήσαμε στην ΑΥΘΝ στο τέταρτο κεφάλαιο ισχύει εξίσου αν, στη θέση των νοητικών συμβόλων, υιοθετήσουμε τα νοητικά είδωλα: η πρακτική γνώση δεξιοτήτων δεν μπορεί να συλληφθεί ούτε με νοητικά είδωλα.

Θεωρούμε ότι το βασικό ερώτημα που πρέπει να απαντηθεί απ' όσους θέλουν να υποστηρίξουν την ύπαρξη εικονιστικού είδους νοητικών αναπαραστάσεων είναι τι διαφοροποιεί μια εικονιστική αναπαράσταση από άλλα είδη αναπαραστάσεων; Αυτό αποτελεί ένα ανοικτό φιλοσοφικό ερώτημα και υπάρχουν αρκετές διαφορετικές απαντήσεις²⁵³. Αν ακολουθήσουμε την ανάλυση που κάναμε στο τρίτο κεφάλαιο, η απάντηση στο συγκεκριμένο ερώτημα θα πρέπει να αναζητηθεί στη σχέση του χρήστη με την αναπαράσταση. Προς αυτή την κατεύθυνση κινείται, ως ένα βαθμό, η θεωρία σημειογραφίας του N. Goodman (1968) η οποία είναι, ίσως, η πιο επεξεργασμένη από τις υπάρχουσες θεωρίες. Παρουσιάζουμε ενδεικτικά αυτή τη θεωρία στο πρώτο παράρτημα της διατριβής.

Αν κανείς ακολουθήσει αυτή τη γραμμή σκέψης η διαφορά των εικονιστικών αναπαραστάσεων από τις συμβολικές ανάγεται πρωταρχικά

²⁵² Βλ. Fodor 1981b.

²⁵³ Βλ., ενδεικτικά, Gombrich 1960, Goodman 1969, Wollheim 1980, Peacocke 1986, Schier 1986, Hopkins 1995 και Hyman 1997.

στη διαφορά των αντίστοιχων αναπαραστασιακών περιεχομένων²⁵⁴. Εκ πρώτης όψεως, αυτή η πρόταση φαίνεται πιο εποικοδομητική από την πρόταση που εντοπίζει τη διαφορά στις αντίστοιχες αναπαραστασιακές μορφές γιατί επιτρέπει να αναδειχθούν είδη περιεχομένου που συλλαμβάνονται από τις εικονιστικές αλλά όχι από τις συμβολικές αναπαραστάσεις. Ωστόσο, δεν θα διερευνήσουμε περαιτέρω το ερώτημα για το τι διαφοροποιεί μια εικονιστική αναπαράσταση από άλλα είδη αναπαραστάσεων γιατί, όπως είναι προφανές, αυτό το χρέος βαρύνει τους υποστηρικτές της αναπαραστασιακής θεωρίας του νου.

5.4 Επιλογος

Η διαμάχη για τα νοητικά είδωλα είναι ακόμα ενεργή στη γνωσιακή επιστήμη όπως, μεταξύ άλλων, αποδεικνύει η δημοσίευση μέσα στο 2001 ενός σημαντικού άρθρου από τον Pylyshyn στο *Behavioral and Brain Sciences*. Ωστόσο, θεωρούμε ότι η κοινή παραδοχή τόσο των υποστηρικτών όσο και των επικριτών των νοητικών ειδώλων ότι η διαμάχη αφορά τη μορφή των νοητικών αναπαραστάσεων, έχει θέσει την όλη συζήτηση σε λάθος βάση. Κατ' αυτόν τον τρόπο, αποσιωπούνται οι διαφορές που παρουσιάζει το αντιληπτικό περιεχόμενο –με την γενικότερη έννοια του όρου– από το περιεχόμενο των πεποηθήσεων και άλλων προτασιακών στάσεων που δεν συνοδεύονται από εποπτείες.

Στη φιλοσοφία του νου διεξάγεται τα τελευταία χρόνια ένας έντονος διάλογος για τα είδη του νοητικού περιεχομένου. Τα δύο αντίπαλα στρατόπεδα διαφωνούν ως προς το εάν, εκτός από το εννοιολογικό περιεχόμενο, υπάρχει και κάποιο άλλο είδος γνωσιακού περιεχομένου το οποίο είναι μη εννοιολογικό²⁵⁵. Στο τέταρτο κεφάλαιο αναφερθήκαμε στην πρακτική γνώση δεξιοτήτων η οποία, υπό μια έννοια, αποτελεί ένα είδος μη εννοιολογικού περιεχομένου. Θεωρούμε ότι η εξέταση της διαμάχης για το μη εννοιολογικό περιεχόμενο μπορεί να δώσει μια νέα ώθηση στη διαμάχη για τα νοητικά είδωλα. Ωστόσο, στο πλαίσιο αυτής της διατριβής δεν θα ασχοληθούμε περαιτέρω με τη διαμάχη για το μη εννοιολογικό περιεχόμενο καθώς η συγκεκριμένη συζήτηση στην αρχική της μορφή δεν συναρτάται με τη διαμάχη για την ύπαρξη διαφορετικών ειδών νοητικών αναπαραστάσεων. Τόσο ο Evans όσο και ο McDowell δεν συναρτούν το εννοιολογικό ή το μη εννοιολογικό περιεχόμενο με κάποιο είδος υποκείμενης νοητικής αναπαράστασης. Αυτό είναι ένα πρόσθετο βήμα που οφείλουν να κάνουν οι υποστηρικτές της αναπαραστασιακής θεωρίας του

²⁵⁴ Η συγκεκριμένη άποψη υποστηρίζεται από τον Haugeland (1991). Επίσης, παρεμφερής είναι η θέση του Peacocke (1986) που κάνει λόγο για "αναλογικό περιεχόμενο" προκειμένου να συλλάβει το μη προτασιακό αλλά αναπαραστάσιμο περιεχόμενο των εικόνων.

²⁵⁵ Βλ., ενδεικτικά, Evans 1982, Cussins 1990, Peacocke 1992 και McDowell 1994.

νου αν θέλουν να αντλήσουν επιχειρήματα για το ένα ή το άλλο είδος νοητικής αναπαράστασης.

Τέλος, θα υπενθυμίσουμε ότι για λόγους απλοποίησης του ζητήματος στο παρόν κεφάλαιο δεν εξετάσαμε καθόλου το ερώτημα "δυνάμει τίνος πράγματος αποκτά μια εσωτερική κατάσταση του εγκεφάλου αναπαραστασιακό περιεχόμενο;". Μάλιστα, προϋποθέσαμε προσωρινά ότι σ' αυτό το ερώτημα μπορεί να δοθεί μια κανονποιητική απάντηση και διερευνήσαμε από 'κει και πέρα τι προβλήματα προκύπτουν. Στο έβδομο κεφάλαιο θα επανέλθουμε σε αυτό το ερώτημα και θα δείξουμε ότι δεν μπορεί να απαντηθεί κανονποιητικά.

ΚΕΦΑΛΑΙΟ ΕΚΤΟ

Κατανομημένες Αναπαραστάσεις

6.1 Εισαγωγή

Η ΑΥΘΝ διατυπώνει το μεταφυσικό πλαίσιο μέσα στο οποίο αναπτύσσονται οι διάφορες επιμέρους θεωρίες της κλασικής γνωσιακής επιστήμης. Όπως είδαμε, συγκροτητικό στοιχείο αυτού του πλαισίου είναι η υιοθέτηση συμβολικών αναπαραστάσεων και τυπικών κανόνων. Έτσι, κάθε απόπειρα διαμόρφωσης μιας θεωρίας του νου που βασίζεται σε διαφορετικού τύπου αναπαραστάσεις έρχεται σε αντίθεση με την κλασική γνωσιακή επιστήμη και διαμορφώνει, δυνάμει, ένα εναλλακτικό πλαίσιο για τον τρόπο που πρέπει να ασκείται αυτή η επιστήμη. Ωστόσο, οι νέες προτάσεις είναι εξαιρετικά περιορισμένες. Μέχρι πριν λίγα χρόνια υπήρχε μία μόνο τέτοια εναλλακτική θεωρία που, όπως είδαμε στο προηγούμενο κεφάλαιο, βασιζόταν στην έννοια των εικονιστικών αναπαραστάσεων. Δεδομένου δε, ότι οι συμβολικές και οι εικονιστικές αναπαραστάσεις φαίνονταν να εξαντλούν το φάσμα των δυνατών τύπων αναπαράστασης²⁵⁶, θεωρήθηκε ότι δεν υπάρχει δυνατότητα ανάπτυξης κάποιου άλλου εναλλακτικού πλαισίου προς αυτή την κατεύθυνση.

Τα τελευταία χρόνια μια εντελώς νέα μορφή αναπαράστασης έχει προταθεί στο πλαίσιο της γνωσιακής επιστήμης. Πρόκειται για τη λεγόμενη "κατανεμημένη αναπαράσταση" (distributed representation). Η έννοια της κατανεμημένης αναπαράστασης έγινε ιδιαίτερα γνωστή το 1986

²⁵⁶ Ορισμένοι ερευνητές αναφέρονται και σε ένα τρίτο είδος αναπαράστασης τα λεγόμενα νοητικά μοντέλα (mental models). Κυριότερος υποστηρικτής των νοητικών μοντέλων είναι ο Johnson-Laird 1983 (βλ., επίσης, Vosniadou & Brewer 1992). Ωστόσο, τα νοητικά μοντέλα, κατά τη γνώμη μας, δεν αποτελούν αυτόνομο είδος αναπαραστάσεων, αλλά ένα υβρίδιο γλωσσικών και εικονιστικών αναπαραστάσεων. Για παράδειγμα, ο Johnson-Laird (1989, σελ. 488) απαντώντας στο ερώτημα "τι ακριβώς είναι ένα νοητικό μοντέλο;" επισημαίνει: "Εάν ο ερωτών απαιτεί έναν πρακτικό ορισμό, τότε ένα νοητικό μοντέλο μπορεί να οριστεί ως μια αναπαράσταση ενός σώματος γνώσης -είτε μακρόχρονης είτε βραχύχρονης -που πληροί τις ακόλουθες συνθήκες:

1. Η δομή της αντιστοιχεί στη δομή της κατάστασης που αναπαριστά
2. Μπορεί να συνίσταται μόνο από στοιχεία που αντιστοιχούν σε αντιληπτές οντότητες, στην οποία περίπτωση μπορεί να υλοποιείται ως ένα νοητικό είδωλο, αντιληπτό ή φανταστικό. Εναλλακτικά, μπορεί να περιέχει στοιχεία που αντιστοιχούν σε αφηρημένες έννοιες· η σημασία τους εξαρτάται ουσιαστικά από τις διαδικασίες χειρισμού των μοντέλων.
3. Αντίθετα με άλλες προτεινόμενες μορφές αναπαράστασης, δεν περιλαμβάνει μεταβλητές. . . Στη θέση της μεταβλητής ... ένα μοντέλο χρησιμοποιεί έκτυπα που αναπαριστούν ένα σύνολο ατόμων"

Αν δεχθούμε ότι τα νοητικά μοντέλα είναι ένα υβρίδιο εικονιστικών και συμβολικών αναπαραστάσεων, τότε η κριτική που έχουμε ασκήσει στην υιοθέτηση συμβολικών και εικονιστικών αναπαραστάσεων για την εξήγηση των νοητικών φαινομένων κληρονομείται και στην περίπτωση των νοητικών μοντέλων.

στο ιστορικό βιβλίο των Rumelhart και McClelland *Parallel Distributed Processing: Explorations in the Microstructures of Cognition* με το οποίο η έρευνα των νευρωνικών δικτύων επανήλθε στο προσκήνιο. Ωστόσο, η συγκεκριμένη έννοια υπήρχε στον ορίζοντα ήδη αρκετά νωρίτερα χωρίς όμως να τυγχάνει ιδιαίτερης προσοχής.

Στο παρόν κεφάλαιο, όπως και στο προηγούμενο, δεν θα ασχοληθούμε με το ερώτημα "δυνάμει τίνος πράγματος αποκτά μια εσωτερική κατάσταση του εγκεφάλου αναπαραστασιακό περιεχόμενο;". Θα θεωρήσουμε προσωρινά ως δεδομένο ότι οι κατανεμημένες αναπαραστάσεις έχουν όντως περιεχόμενο και θα διερευνήσουμε σε τι μπορεί να διαφοροποιείται το συγκεκριμένο είδος αναπαραστάσεων από τα υπόλοιπα και κυρίως τις συμβολικές αναπαραστάσεις. Το γενικότερο ερώτημα για το πώς και εάν μπορεί μια εσωτερική κατάσταση του εγκεφάλου να αποκτήσει αναπαραστασιακό περιεχόμενο θα εξεταστεί ανεξάρτητα στο έβδομο κεφάλαιο.

6.2 Γιατί τα νευρωνικά δίκτυα αφορούν τη γνωσιακή επιστήμη;

"Κατανεμημένες αναπαραστάσεις" ονομάζονται οι αναπαραστάσεις που διαθέτουν ορισμένοι τύποι των λεγόμενων τεχνητών νευρωνικών δικτύων (artificial neural networks). Προκειμένου, λοιπόν, να εξετάσουμε το τι θεωρείται κατανεμημένη αναπαράσταση θα πρέπει πρώτα να περιγράψουμε σε γενικές γραμμές πώς λειτουργεί ένα τεχνητό νευρωνικό δίκτυο. Στη βιβλιογραφία η προσέγγιση που βασίζεται στα τεχνητά νευρωνικά δίκτυα (στο εξής, χάριν συντομίας, νευρωνικά δίκτυα) αναφέρεται επίσης με τα ονόματα "παράλληλη κατανεμημένη επεξεργασία" (parallel distributed processing) και "συνδεδετισμός" (connectionism)²⁵⁷.

Τα νευρωνικά δίκτυα είναι ουσιαστικά συσχετιστές σχηματισμών (pattern associators). Συσχετίζουν εισόδους με εξόδους, για παράδειγμα γραπτό λόγο με προφορικό λόγο ή ρήματα στον ενεστώτα με αντίστοιχα ρήματα στον αόριστο. Το ενδιαφέρον έγκειται ωστόσο στον τρόπο που γίνεται αυτή η συσχέτιση. Σχηματικά μιλώντας, το μόνο που κάνει αυτή η μηχανή είναι να συσχετίζει κάποιες ακολουθίες αριθμών (τις εισόδους) με κάποιες άλλες ακολουθίες αριθμών (τις εξόδους), "λύνοντας" συστήματα εξισώσεων ύστερα από μια μακρά περίοδο εκμάθησης. Δεν υπάρχει

²⁵⁷ Παρότι οι όροι "συνδεδετισμός" και "παράλληλα κατανεμημένη επεξεργασία" έχουν πλέον καταστεί συνώνυμοι, αρχικά είχαν διαφορετικές σημασίες. Τα συνδεδετιστικά μοντέλα σχετίζονταν με την έρευνα του Ballard στο πανεπιστήμιο του Rochester, ενώ τα μοντέλα παράλληλα κατανεμημένης επεξεργασίας σχετίζονταν με την αντίστοιχη ερευνητική ομάδα του πανεπιστημίου του San Diego (βλ. Smolensky 1991, σελ. 225, σημ. 5). Βλ., επίσης, την άποψη του Sejnowski πάνω σε αυτό το θέμα σε μια συνέντευξή του που δημοσιεύεται στο Baumgartner & Payr, 1995, σελ. 224.

κάποιο επίπεδο περιγραφής της μηχανής στο οποίο η εσωτερική της λειτουργία να μπορεί να ερμηνευτεί με βάση αναπαραστάσεις ιδιοτήτων που αφορούν τους σχηματισμούς εισόδου. Επίσης, δεν υπάρχει ένα ανεξάρτητο επίπεδο προγράμματος ούτε κεντρική μονάδα επεξεργασίας. Για να γίνουν όλα αυτά περισσότερο κατανοητά θα πρέπει να αναλύσουμε λεπτομερικά τη λειτουργία των νευρωνικών δικτύων. Προηγουμένως, όμως, είναι σκόπιμο να εξετάσουμε τους λόγους για τους οποίους η μελέτη των νευρωνικών δικτύων θεωρείται ότι συμβάλλει στην έρευνα των γνωσιακών φαινομένων.

Μέχρις αυτό το σημείο έχουμε δει ότι οι ψηφιακοί υπολογιστές αποτέλεσαν το μόνο και αποκλειστικό τεχνολογικό μοντέλο στο οποίο βασίστηκε η κλασική γνωσιακή επιστήμη²⁵⁸. Μέσα σε ένα τέτοιο κλίμα, ο υπερτονισμός της αναλογίας νου - ψηφιακού υπολογιστή ήταν σχεδόν αναπόφευκτος. Ωστόσο, οι ψηφιακοί υπολογιστές διαθέτουν πολλές "γνωσιακές" ικανότητες που δεν παρατηρούνται στους ανθρώπους. Ενδεικτικά επισημαίνουμε τις ακόλουθες:

- A. Οι ψηφιακοί υπολογιστές εκτελούν μια ακολουθία αριθμητικών πράξεων σε πολύ μικρό χρόνο και χωρίς λάθη. Αντίθετα οι άνθρωποι είναι εξαιρετικά αργοί στους αριθμητικούς υπολογισμούς και όχι πάντα αλάθητοι.
- B. Οι ψηφιακοί υπολογιστές είναι ταχείς και αλάθητοι κατά την εκτέλεση σύνθετων παραγωγικών συλλογισμών ενώ, αντίθετα, η ανθρώπινη απόδοση μειώνεται δραματικά όσο αυξάνει ο αριθμός των προκειμένων.
- Γ. Οι ψηφιακοί υπολογιστές μπορούν να αποθηκεύσουν πάρα πολύ γρήγορα ένα τεράστιο αριθμό πληροφοριών υπό μορφή προτασιακής γνώσης ενώ, αντίθετα, στους ανθρώπους η απόκτηση προτασιακής γνώσης είναι μια επίπονη και χρονοβόρα διαδικασία.

²⁵⁸ Λέγοντας ψηφιακοί υπολογιστές ουσιαστικά αναφερόμαστε στους υπολογιστές με αρχιτεκτονική von Neumann (βλ. κεφάλαιο 1, ενότητα 1.3.7). Τα βασικά χαρακτηριστικά αυτών των υπολογιστών, όπως τα επισημαίνει ο Copeland (1993, σελ. 193), είναι:

1. Οι μηχανές von Neumann είναι σειριακοί υπολογιστές: εκτελούν τις εντολές του προγράμματος διαδοχικά
2. Κάθε έκφραση (ακολουθία συμβόλων) αποθηκεύεται σε συγκεκριμένη θέση της μνήμης
3. Η πρόσβαση σε μια αποθηκευμένη έκφραση γίνεται πάντοτε μέσω της αριθμητικής διεύθυνσης που έχει η θέση της έκφρασης
4. Υπάρχει μια μόνο μονάδα ελέγχου, η κεντρική μονάδα επεξεργασίας (cpu)

Αυτό που χαρακτηρίζει περισσότερο τους ψηφιακούς υπολογιστές ως μοντέλο του νου είναι ο χειρισμός (μετασχηματισμός, αποθήκευση και ανάκτηση) δομημένων συμβολικών εκφράσεων.

Όμως και οι άνθρωποι παρουσιάζουν αρκετές γνωσιακές ικανότητες που δεν απαντώνται στους ψηφιακούς υπολογιστές:

Α. Εκτελούν πολύ γρήγορα αντιληπτικές λειτουργίες. Για παράδειγμα, μπορούν να κάνουν άμεση κατηγοριοποίηση, δηλαδή, να αναγνωρίσουν άμεσα σε ποιον τύπο αντικειμένου ανήκει κάποιο επιμέρους αντικείμενο. Αντίθετα, για τους ψηφιακούς υπολογιστές μια τέτοια λειτουργία είναι χρονοβόρα και αμφιβόλου αποτελεσματικότητας.

Β. Οι άνθρωποι είναι ιδιαίτερα καινοί στο να εξάγουν συμπεράσματα στηριζόμενοι σε ασαφείς και ελλειπείς πληροφορίες. Συγκεκριμένα, σε αντίθεση με τους ψηφιακούς υπολογιστές, είναι πολύ καλοί στον επαγωγικό, τον απαγωγικό (abductive) και τον αναλογικό συλλογισμό. Θα πρέπει να επισημάνουμε ότι η διαφορά με τους ψηφιακούς υπολογιστές δεν έγκειται μόνο στην ταχύτητα εκτέλεσης των συλλογισμών αυτού του είδους, αλλά και στην ποιότητα των συμπερασμάτων.

Γ. Οι άνθρωποι λαμβάνουν υπόψη τους το πλαίσιο μέσα στο οποίο καλούνται να πράξουν και για αυτό τον λόγο προσαρμόζονται πολύ εύκολα και γρήγορα στις μεταβαλλόμενες συνθήκες του περιβάλλοντος. Αντίθετα, τα ρομπότ που βασίζονται σε ψηφιακούς υπολογιστές συνυπολογίζουν το πλαίσιο μέσα στο οποίο κινούνται μόνο στο βαθμό που αυτό έχει εκ των προτέρων αναπαρασταθεί εντός τους (βλ. κεφάλαιο 4, ενότητα 4.2.2.3.2). Έτσι, κάθε φορά που έρχονται αντιμέτωπα με ένα μη προβλεπόμενο πλαίσιο η απόδοσή τους φθίνει δραματικά²⁵⁹

Δ. Οι ασθενείς με επιλεκτικές βλάβες στον εγκέφαλο δεν χάνουν πάντα ολοκληρωτικά ορισμένες γνωσιακές τους ικανότητες: πολύ συχνά απλώς παρουσιάζουν μειωμένη απόδοση σε αυτές τις ικανότητες. Αντίθετα στους ψηφιακούς υπολογιστές οποιαδήποτε μικρή βλάβη έχει ως συνέπεια τον τερματισμό της λειτουργίας τους.

Οι παραπάνω εμπειρικές παρατηρήσεις δεν είναι φυσικά χωρίς αντίλογο. Οι υποστηρικτές της θέσης ότι ο νους είναι ψηφιακός υπολογιστής έχουν κάτι να πουν για κάθε μία από τις παραπάνω επισημάνσεις. Μια βασική τους γραμμική άμυνα είναι ότι ο νους είναι εξαιρετικά πολύπλοκος, τόσο πολύπλοκος που με την σημερινή τεχνολογία κανένας παρόμοιος υπολογιστής δεν είναι δυνατόν να κατασκευασθεί. Πιο συγκεκριμένα, συσχετίζουν την πολυπλοκότητα με την ποσότητα των επεξεργαζόμενων πληροφοριών και εν τέλει, με την υπολογιστική ισχύ²⁶⁰. Έτσι θεωρούν ότι οι ανθρώπινες γνωσιακές

²⁵⁹ Πρόκειται για αυτό που στα αγγλικά αποκαλείται brittleness, ήτοι ευθραυστότητα.

²⁶⁰ Είναι χαρακτηριστικό ότι πολλοί ερευνητές, κυρίως από τον χώρο της Τεχνητής Νοημοσύνης, βλέπουν ακόμα και τα νευρωνικά δίκτυα αποκλειστικά μέσα από αυτό το πρίσμα, θεωρούν δηλαδή ότι η μόνη διαφορά αυτής της νέας

ιδιότητες που δεν έχει μέχρι σήμερα καταστεί δυνατόν να προσομοιωθούν, δεν μπορεί παρά να οφείλονται στην ύπαρξη πρόσθετης προτασιακής γνώσης, είτε υπό τη μορφή δεδομένων, είτε υπό τη μορφή κανόνων. Ωστόσο, όπως είδαμε στο τέταρτο κεφάλαιο, αυτό το επιχείρημα παρουσιάζει ανυπέβλητα προβλήματα. Να προσέσουμε εδώ μόνο ότι η αύξηση της προτασιακής γνώσης τις περισσότερες φορές οδηγεί σε εκθετικού τύπου πολυπλοκότητα οπότε σε αυτές τις περιπτώσεις καμία αύξηση της υπολογιστικής ισχύος δεν θα ήταν επαρκής²⁶¹.

Πολλές από τις παραπάνω γνωσιακές ικανότητες και αδυναμίες που παρουσιάζουν οι άνθρωποι σε σχέση με τους ψηφιακούς υπολογιστές, απαντώνται και στα νευρωνικά δίκτυα. Χαρακτηριστικές ικανότητες των νευρωνικών δικτύων είναι η αναγνώριση οικογενειακών ομοιοτήτων, η αυτόματη κατηγοριοποίηση, η μάθηση μέσω παραδειγμάτων, η βαθμιαία μείωση της απόδοσης (*graceful degradation*)²⁶², η ύπαρξη μνήμης που κατευθύνεται από το περιεχόμενο (*content addressable memory*)²⁶³, η ανθεκτικότητα στον θόρυβο κα. Από την άλλη μεριά, η απόδοση των νευρωνικών δικτύων δεν είναι καλή σε γνωσιακές λειτουργίες όπως ο παραγωγικός συλλογισμός, που προσομοιώθηκαν με επιτυχία από την

αρχιτεκτονικής από την κλασική είναι ότι μας προσφέρει τη δυνατότητα πολύ μεγαλύτερου όγκου υπολογισμών. Ωστόσο, όπως θα δούμε, αυτή η προσέγγιση επισκιάζει το γεγονός ότι τα νευρωνικά δίκτυα έχουν τη δυνατότητα να κάνουν πολύ διαφορετικού τύπου επεξεργασία από την υπολογιστική επεξεργασία συμβολικών αναπαραστάσεων.

²⁶¹ Βλ. Glymour 1998.

²⁶² Για παράδειγμα, αν καταστραφεί ένα τμήμα του δικτύου δεν διαταράσσεται πλήρως η λειτουργία του. Η μείωση της απόδοσης του δικτύου είναι ανάλογη του μεγέθους του κατεστραμμένου τμήματος.

²⁶³ Στους κλασικούς υπολογιστές η ανάκτηση δεδομένων από τη μνήμη βασίζεται στη γνώση της διεύθυνσης στην οποία είναι αποθηκευμένα τα δεδομένα. Κάθε συγκεκριμένη θέση μνήμης σε έναν υπολογιστή έχει μια μοναδική διεύθυνση. Η διεύθυνση αποθήκευσης κάποιας αναπαράστασης δεν σχετίζεται με το περιεχόμενο της αναπαράστασης. Συνεπώς, δεν υπάρχει τρόπος άμεσου εντοπισμού της διεύθυνσης με βάση τη γνώση του περιεχομένου της αναπαράστασης. Για το λόγο αυτό, αν δεν γνωρίζουμε τη διεύθυνση μιας αναπαράστασης, απαιτείται μαζική έρευνα ολόκληρης της μνήμης μέχρι αυτή να εντοπιστεί. Τα νευρωνικά δίκτυα, αντίθετα, έχουν μνήμη κατευθυνόμενη από το περιεχόμενο. Αυτό συμβαίνει γιατί όλες οι πληροφορίες αποθηκεύονται στον ίδιο χώρο, οπότε ακόμα και ένας ελλειπτικός σχηματισμός ενεργοποίησης της εισόδου μπορεί να "εγκαλέσει" αυτόματα τον πιο αρμόζοντα προς αυτόν σχηματισμό εξόδου (βλ. Boden 1988, σελ. 251 και Copeland 1993, σελ. 188-189). Πάντως, θα πρέπει να επισημάνουμε προκαταρκτικά το εξής: είναι, ίσως, καταχρηστικό το να γίνεται λόγος για αναπαραστασιακό περιεχόμενο, στην καλύτερη περίπτωση πρόκειται για τον κατοπιρισμό του περιεχομένου στην αναπαραστασιακή μορφή. Το σημείο αυτό θα γίνει περισσότερο κατανοητό στο έβδομο κεφάλαιο και στον Επίλογο.

κλασική ΤΝ, αλλά όπως επισημάναμε, οι περισσότερες από αυτές δεν χαρακτηρίζουν ιδιαίτερα ούτε και τους ανθρώπους.

Επίσης τα νευρωνικά δίκτυα, σε αντίθεση με τους ψηφιακούς υπολογιστές, παρουσιάζουν αρκετές ομοιότητες με τη δομή και λειτουργία του εγκεφάλου. Δεν είναι τυχαίο ότι οι πρώτοι ερευνητές²⁶⁴ που κατασκεύασαν νευρωνικά δίκτυα εμπνεύστηκαν την όλη ιδέα από τον τρόπο που λειτουργούν και διασυνδέονται οι νευρώνες του εγκεφάλου. Φυσικά, η αναλογία μεταξύ τεχνητών νευρωνικών δικτύων και εγκεφάλου ισχύει μέχρι ένα επίπεδο ανάλυσης. Κατ' αρχάς στον εγκέφαλο υπάρχουν πολλοί διαφορετικοί τύποι νευρώνων (αστεροειδείς, πυραμιδικοί κ.α.) κάθε ένας από τους οποίους επιτελεί διαφορετικές λειτουργίες. Επίσης οι νευρώνες του εγκεφάλου είναι είτε διεγερτικοί (excitatory) είτε ανασταλτικοί (inhibitory), ενώ οι "νευρώνες" των τεχνητών νευρωνικών δικτύων έχουν ταυτόχρονα και τις δύο δυνατότητες²⁶⁵. Τέλος, οι περισσότερες μέθοδοι εκπαίδευσης των τεχνητών νευρωνικών δικτύων δεν φαίνεται να είναι ρεαλιστικές από βιολογική άποψη²⁶⁶. Ωστόσο παρά τις διαφορές φυσικών και τεχνητών νευρωνικών δικτύων, οι ομοιότητές τους εξακολουθούν να είναι αρκετά σημαντικές, ώστε να δικαιολογείται ως ένα βαθμό η θέση ότι ο συνδετισμός είναι περισσότερο ευλογοφανής από βιολογική άποψη απ' ό,τι η κλασική αρχιτεκτονική. Τι είναι όμως και πώς λειτουργεί ένα τεχνητό νευρωνικό δίκτυο;

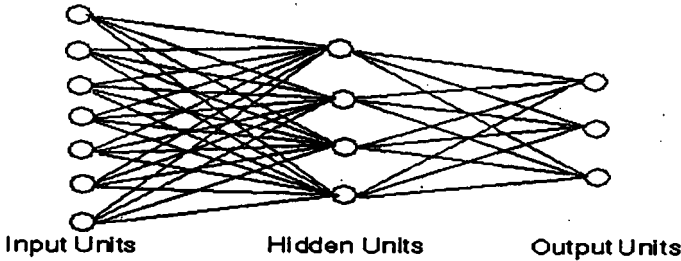
6.3 Τι είναι ένα νευρωνικό δίκτυο;

Τα νευρωνικά δίκτυα αποτελούνται από ένα μεγάλο αριθμό απλών επεξεργαστών που είναι συνδεδεμένοι μεταξύ τους. Οι επεξεργαστές καλούνται "κόμβοι" του δικτύου. Κάθε κόμβος υπολογίζει κάποια συνάρτηση που ονομάζεται συνάρτηση ενεργοποίησης (activation function). Στα απλούστερα νευρωνικά δίκτυα οι κόμβοι είναι διατεταγμένοι σε διαφορετικά επίπεδα (βλέπε σχήμα 1).

²⁶⁴ Βλ. McCulloch & Pitts, 1943.

²⁶⁵ Αυτή η διαφορά μπορεί να εξαλειφθεί αν υποθέσει κανείς ότι οι τεχνητοί νευρώνες αποτελούν προσομοιώσεις ολόκληρων ομάδων φυσικών νευρώνων.

²⁶⁶ Επίσης, ενώ οι άνθρωποι έχουν διαρκώς την δυνατότητα να αποκτούν νέα γνώση, η περίοδος εκπαίδευσης των τεχνητών νευρωνικών δικτύων είναι αυστηρά περιορισμένη. Μετά το πέρας αυτής καμιά νέα γνώση δεν μπορεί να "προστεθεί" στο δίκτυο χωρίς να απολεσθεί η προηγούμενη γνώση του δικτύου. Πρόκειται για το φαινόμενο της καταστροφικής παρεμβολής (catastrophic interference) (French 1999). Ο μόνος τρόπος για να διατηρήσει ένα νευρωνικό δίκτυο την παλιά γνώση κατά την εκπαίδευση για την προσθήκη νέας γνώσης είναι να εκπαιδευτεί εκ νέου και για την παλιά γνώση.

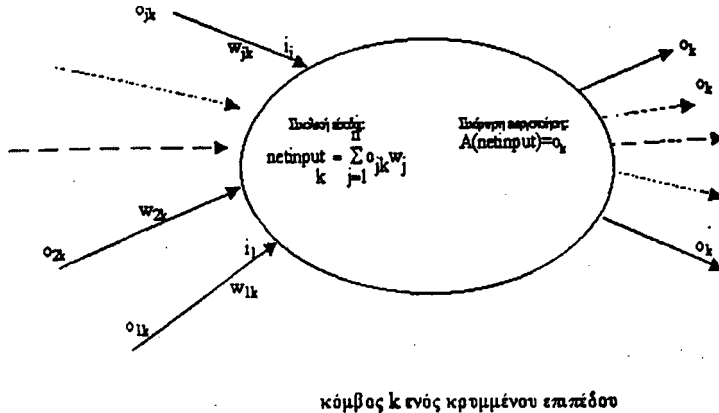


Σχήμα 1

Τα επίπεδα διακρίνονται σε τρεις κατηγορίες: το επίπεδο εισόδου, τα κρυμμένα επίπεδα²⁶⁷ και το επίπεδο εξόδου. Οι επεξεργαστές που ανήκουν στο ίδιο επίπεδο συνήθως δεν είναι συνδεδεμένοι μεταξύ τους. Κάθε ένας από τους επεξεργαστές ενός κρυμμένου επιπέδου είναι συνδεδεμένος με ορισμένους (ή όλους) τους επεξεργαστές του προηγούμενου και του επόμενου επιπέδου. Κάθε σύνδεση μεταξύ δύο επεξεργαστών δύο διαδοχικών επιπέδων μεταφέρει μια αριθμητική τιμή από τον έναν επεξεργαστή στον άλλον. Ακριβέστερα, η έξοδος του ενός επεξεργαστή πολλαπλασιάζεται με έναν αριθμό που καλείται "βάρος" και ο οποίος χαρακτηρίζει την συγκεκριμένη σύνδεση. Το αποτέλεσμα του γινομένου της εξόδου επί το βάρος της σύνδεσης αποτελεί μία από τις εισόδους των επεξεργαστών του επόμενου επιπέδου. Έτσι, σε κάθε επεξεργαστή των κρυμμένων επιπέδων ή του επιπέδου εξόδου φτάνουν οι έξοδοι πολλών (ή όλων) των επεξεργαστών του αμέσως προηγούμενου επιπέδου αφού πρώτα πολλαπλασιαστούν επί το βάρος της αντίστοιχης σύνδεσης. Το άθροισμα αυτών των γινομένων αποτελεί την είσοδο του κόμβου στον οποίο φτάνουν οι συγκεκριμένες έξοδοι.

Ας δούμε όμως όλα αυτά λίγο πιο συστηματικά εισάγοντας μαθηματικούς συμβολισμούς

²⁶⁷ Τα πρώτα νευρωνικά δίκτυα που δημιουργήθηκαν κατά τις δεκαετίες του '40 και του '50 δεν διέθεταν κρυμμένα επίπεδα. Το συγκεκριμένο γεγονός αποδείχθηκε μοιραίο για το μέλλον αυτών των πρώτων μοντέλων. Η κριτική των Minsky και Papert (1969) οδήγησε ουσιαστικά στην εγκατάλειψη αυτής της τεχνολογίας μέχρι τις αρχές της δεκαετίας του '80.



Σχήμα 2

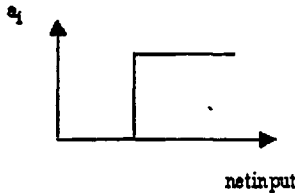
Έστω ότι οι έξοδοι των j κόμβων του επιπέδου εισόδου που φτάνουν στο k -στο κόμβο του αμέσως επόμενου επιπέδου είναι οι o_{1k}, o_{2k}, \dots και o_{jk} (βλ. σχήμα 2). Αν $w_{1k}, w_{2k}, \dots, w_{jk}$ είναι αντίστοιχα τα βάρη των j συνδέσεων καθενός από τους κόμβους του επιπέδου εισόδου με τον k -στο κόμβο του αμέσως επόμενου επιπέδου, τότε η είσοδος σε αυτόν τον κόμβο από κάθε μία από αυτές τις συνδέσεις είναι αντίστοιχα $o_{1k}w_{1k}, o_{2k}w_{2k}, \dots, o_{jk}w_{jk}$. Η συνολική είσοδος στον k -στο κόμβο του δεύτερου επιπέδου ($netinput_k$) είναι ίση με το άθροισμα όλων αυτών των επιμέρους εισόδων (βλ. σχήμα 2). Η συνολική είσοδος τροφοδοτείται στην συνάρτηση ενεργοποίησης. Η έξοδος της συνάρτησης ενεργοποίησης αποτελεί την έξοδο που ο συγκεκριμένος κόμβος στέλνει προς τους κόμβους του επόμενου επιπέδου. Την έξοδο αυτή θα την ονομάζουμε και βαθμό ενεργοποίησης του κόμβου.

6.4 Είδη νευρωνικών δικτύων

Τα τελευταία είκοσι χρόνια έχει αναπτυχθεί μια πληθώρα διαφορετικών ειδών νευρωνικών δικτύων. Δύο είναι κυρίως οι τρόποι με τους οποίους ταξινομούνται τα νευρωνικά δίκτυα: είτε με βάση τη συνάρτηση ενεργοποίησης των κόμβων είτε με βάση το είδος της σύνδεσης μεταξύ των κόμβων.

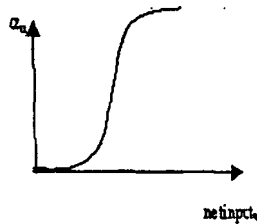
Η συνάρτηση ενεργοποίησης μπορεί καταρχάς να είναι διακριτή ή συνεχής. Στην περίπτωση που είναι διακριτή επιστρέφει, συνήθως, τις τιμές 0 και 1. Αυτό επιτυγχάνεται με κάποια συνάρτηση βήματος η οποία για κάθε είσοδο που ξεπερνά μια συγκεκριμένη τιμή (τιμή κατωφλίου) δίνει ως έξοδο την τιμή 1 και για κάθε άλλη περίπτωση δίνει την τιμή 0. Σε πιο σύνθετες περιπτώσεις, η είσοδος του κόμβου εισάγεται σε μια συνάρτηση πιθανότητας και ανάλογα με το αποτέλεσμα προκύπτει η έξοδος. Γενικά, μια διακριτή συνάρτηση ενεργοποίησης έχει το

πλεονέκτημα ότι ελαχιστοποιεί την επίδραση μικρών εισόδων που συνήθως είναι αποτέλεσμα θορύβου του συστήματος.



Αν τώρα η συνάρτηση ενεργοποίησης είναι συνεχής η έξοδος του εκάστοτε κόμβου μπορεί να λάβει οποιαδήποτε πραγματική τιμή από κάποιο υποσύνολο των πραγματικών αριθμών. Ειδικότερα, η συνάρτηση ενεργοποίησης σε αυτή την περίπτωση είναι συνήθως είτε γραμμική (π.χ. $a_i = c \cdot \text{netinput}_i$) όπου η έξοδος είναι ευθέως ανάλογη της εισόδου, είτε σιγμοειδής π.χ.

$$a_u = \frac{1}{1 + e^{-\frac{\text{netinput}_i - \theta_i}{T}}}$$



Η σιγμοειδής συνάρτηση ενεργοποίησης συνδυάζει τα πλεονεκτήματα της συνάρτησης βήματος και της γραμμικής συνάρτησης: για μέτριες τιμές εισόδου η έξοδος είναι σχεδόν ανάλογη προς την είσοδο, ενώ για πολύ χαμηλές και πολύ υψηλές τιμές εισόδου η έξοδος τείνει αντίστοιχα στο 0 και το 1.

Ένας άλλος - ίσως σημαντικότερος - τρόπος κατηγοριοποίησης των νευρωνικών δικτύων είναι με βάση τον τύπο σύνδεσης των κόμβων. Σύμφωνα με αυτό το κριτήριο, τα νευρωνικά δίκτυα χωρίζονται σε δύο μεγάλες κατηγορίες: τα δίκτυα πρόσθιας τροφοδότησης (feedforward networks) και τα αναδρομικά δίκτυα (recurrent networks). Τα πρώτα είναι τα δίκτυα στα οποία η κατεύθυνση της πληροφορίας μεταξύ των κόμβων είναι μονόδρομη: η ενεργοποίηση μεταδίδεται από τον έναν κόμβο στον άλλο χωρίς αυτή να περνά πάνω από μία φορά από κάποιο κόμβο. Έτσι, στα δίκτυα πρόσθιας τροφοδότησης μπορούμε να διακρίνουμε διαφορετικά επίπεδα κόμβων που ενεργοποιούνται διαδοχικά, ξεκινώντας από το επίπεδο εισόδου, περνώντας στα κρυμμένα επίπεδα και καταλήγοντας στο επίπεδο εξόδου. Αντίθετα, στα αναδρομικά δίκτυα δεν μπορεί να γίνει με σαφήνεια αυτή η διάκριση γιατί η σύνδεση μεταξύ των κόμβων δεν είναι μονόδρομη, δηλαδή η μετάδοση της ενεργοποίησης δεν πηγαινει μόνο προς τη μία κατεύθυνση. Αποτέλεσμα αυτού είναι τα αναδρομικά δίκτυα να εμφανίζουν έναν ή περισσότερους κύκλους και να μην μπορούν έτσι να διακριθούν σε ξεχωριστά επίπεδα διότι σε κάθε

κύκλο επεξεργασίας οι κόμβοι ενεργοποιούνται περισσότερες της μιας φορές.²⁶⁸ Χαρακτηριστικά παραδείγματα αναδρομικών νευρωνικών δικτύων αποτελούν το δίκτυο Hopfield²⁶⁹ (με γραμμική συνάρτηση ενεργοποίησης) και η μηχανή Boltzmann²⁷⁰ (με στοχαστική συνάρτηση ενεργοποίησης). Δεν θα επεκταθούμε στην ανάλυση των αναδρομικών νευρωνικών δικτύων γιατί αφενός η μαθηματική τους περιγραφή είναι αρκετά σύνθετη και αφετέρου δεν συμβάλλουν περισσότερο από τα δίκτυα πρόσθιας τροφοδότησης στην κατανόηση της έννοιας της κατανεμημένης αναπαράστασης που αποτελεί και το βασικό στόχο αυτού του κεφαλαίου. Έτσι λοιπόν, στο υπόλοιπο του κεφαλαίου κάθε αναφορά στα νευρωνικά δίκτυα θα αφορά τα δίκτυα πρόσθιας τροφοδότησης.

6.5 Πώς μαθαίνουν τα νευρωνικά δίκτυα;

Μια από τις σημαντικότερες ιδιότητες των νευρωνικών δικτύων είναι ότι έχουν τη δυνατότητα να εκπαιδεύονται. Όταν ένα νευρωνικό δίκτυο μαθαίνει να εκτελεί ικανοποιητικά ένα έργο τότε ουσιαστικά αλλάζει η απόδοσή του, δηλαδή για τις εκάστοτε εισόδους παράγει διαφορετικές εξόδους απ' ό,τι παρήγαγε πριν εκπαιδευτεί. Αυτή η αλλαγή στην συμπεριφορά του νευρωνικού δικτύου οφείλεται στην αλλαγή της εσωτερικής του δομής. Τρία είναι τα στοιχεία που χαρακτηρίζουν τη δομή ενός νευρωνικού δικτύου: ο αριθμός των κόμβων, η συνάρτηση ενεργοποίησης των κόμβων και τα βάρη των συνδέσεων μεταξύ των κόμβων. Κατά συνέπεια, οποιαδήποτε μεταβολή της δομής του νευρωνικού δικτύου αφορά κάποιο από αυτά τα στοιχεία. Ωστόσο, ο αριθμός των κόμβων είναι συνήθως σταθερός²⁷¹ και το ίδιο συμβαίνει επίσης για τη συνάρτηση ενεργοποίησης των κόμβων²⁷². Έτσι, οποιαδήποτε μεταβολή κατά την εκπαίδευση ενός νευρωνικού δικτύου σχετίζεται με την μεταβολή του βάρους των συνδέσεων.

Τα τελευταία κυρίως χρόνια έχουν αναπτυχθεί πολλές διαφορετικές τεχνικές εκπαίδευσης των νευρωνικών δικτύων. Γενικά, μπορούμε να τις διακρίνουμε σε δύο κατηγορίες: την επιβλεπόμενη μάθηση (supervised learning) και την μη επιβλεπόμενη μάθηση (unsupervised learning). Στην πρώτη κατηγορία εντάσσονται τεχνικές στις

²⁶⁸ Στο πλαίσιο ενός κύκλου επεξεργασίας η κατάσταση ενεργοποίησης του δικτύου κατά την χρονική στιγμή t εξαρτάται από την κατάσταση ενεργοποίησης του την χρονική στιγμή $t-1$. Για τον λόγο αυτό τα αναδρομικά νευρωνικά δίκτυα προσφέρονται άμεσα για να μελετηθούν ως δυναμικά συστήματα (βλ. Παράρτημα 2).

²⁶⁹ Βλ. Hopfield 1982.

²⁷⁰ Βλ. Hinton & Sejnowski 1986.

²⁷¹ Εκτός αν πρόκειται για τα λεγόμενα δυναμικά δίκτυα, τα οποία έχουν τη δυνατότητα κατά την περίοδο εκπαίδευσης να προσθέτουν νέους κόμβους.

²⁷² Εκτός από ορισμένες περιπτώσεις στις οποίες υπάρχει η δυνατότητα μεταβολής κάποιου συντελεστή της εξίσωσης ενεργοποίησης.

οποίες καθορίζεται ρητά εκ των προτέρων η επιθυμητή έξοδος για την εκάστοτε εισοδο²⁷³. Αντίθετα, στις τεχνικές μη επιβλεπόμενης μάθησης (πχ. ανταγωνιστική μάθηση (competitive learning)) το νευρωνικό δίκτυο μεταβάλλει τα βάρη των συνδέσεων χωρίς να γνωρίζει εκ των προτέρων τις επιθυμητές εξόδους, απλά "αναπαριστά τους εκάστοτε σχηματισμούς εισόδου κατά ένα τρόπο που αντανακλά τη στατιστική δομή ολόκληρου του συνόλου των σχηματισμών εισόδου" (Dayan 1999).

Δεδομένου ότι στο υπόλοιπο του κεφαλαίου θα ασχοληθούμε αποκλειστικά με δίκτυα πρόσθιας τροφοδότησης είναι σκόπιμο να περιγράψουμε κάπως πιο αναλυτικά την κυριότερη μέθοδο εκπαίδευσης αυτών των δικτύων. Η μέθοδος ανήκει στην κατηγορία της επιβλεπόμενης μάθησης και ονομάζεται μέθοδος της αντίστροφης διάδοσης του σφάλματος (back-propagation of error). Εφαρμόζεται στα δίκτυα πρόσθιας τροφοδότησης με κρυμμένα επίπεδα. Η ανακάλυψη αυτής της μεθόδου αποτέλεσε την κινητήρια δύναμη για την ανάπτυξη και μελέτη δικτύων με κρυμμένα επίπεδα.

Η μέθοδος της αντίστροφης διάδοσης του σφάλματος σε γενικές γραμμές λειτουργεί με τον ακόλουθο τρόπο: έστω ένα δίκτυο πρόσθιας τροφοδότησης τριών επιπέδων με τυχαία βάρη συνδέσεων (αρνητικά και θετικά). Αρχικά το δίκτυο τροφοδοτείται με έναν σχηματισμό εισόδου (ουσιαστικά ένα διάνυσμα αριθμών πλήθους n , όπου n ο αριθμός των κόμβων του επιπέδου εισόδου) και υπολογίζει τον αντίστοιχο σχηματισμό εξόδου. Κατόπιν, ο σχηματισμός εξόδου συγκρίνεται με το επιθυμητό διάνυσμα εξόδου και η διαφορά που προκύπτει υψώνεται στο τετράγωνο. Ο μέσος όρος όλων των αριθμών που αποτελούν αυτό το διάνυσμα ονομάζεται μέσο τετραγωνισμένο σφάλμα (mean squared error). Ακριβώς αυτός ο αριθμός αποτελεί το μέτρο της εκπαίδευσης του δικτύου. Στόχος, λοιπόν, είναι η ελαχιστοποίηση του μέσου τετραγωνισμένου σφάλματος.

Η διαδικασία που ακολουθείται στη συνέχεια μπορεί να περιγραφεί κάπως απλουστευτικά ως εξής: κρατάμε όλα τα βάρη εκτός από ένα σταθερά και εξετάζουμε αν μια πολύ μικρή μεταβολή του συγκεκριμένου βάρους (η οποία προκύπτει από έναν αλγόριθμο που χρησιμοποιεί τη συνάρτηση του σφάλματος) μειώνει το μέσο τετραγωνισμένο σφάλμα. Αν όντως μειώνεται το σφάλμα τότε διατηρούμε τη νέα τιμή του βάρους, αλλιώς αφήνουμε την αρχική τιμή. Η διαδικασία αυτή επαναλαμβάνεται διαδοχικά για όλα τα υπόλοιπα βάρη και έτσι καταλήγουμε σε ένα νέο δίκτυο με διαφορετικά βάρη συνδέσεων το οποίο έχει μικρότερο μέσο τετραγωνισμένο σφάλμα.

²⁷³ Επίσης, στην ίδια κατηγορία εντάσσεται και η λεγόμενη μάθηση μέσω ενίσχυσης (reinforcement learning) κατά την οποία δηλώνεται στο δίκτυο το εάν η εκάστοτε έξοδος είναι κοντά στην επιθυμητή έξοδο χωρίς όμως να καθορίζεται η ακριβής τιμή της τελευταίας.

Η όλη διαδικασία επαναλαμβάνεται για όλα τα υπόλοιπα ζεύγη εισόδου-εξόδου που περιλαμβάνονται στο σύνολο εκμάθησης του δικτύου. Αν το μέσο τετραγωνισμένο σφάλμα δεν έχει φτάσει σε μια αποδεκτή τιμή επαναλαμβάνεται ολόκληρος ο κύκλος εκμάθησης τόσες φορές όσες είναι αρκετές για να επιτευχθεί ο στόχος.

Η μέθοδος της αντίστροφης διάδοσης του σφάλματος είναι πλήρως αυτοματοποιημένη, δηλαδή μπορεί να εκτελεστεί από έναν υπολογιστή που έχει τροφοδοτηθεί με τα ζεύγη εισόδου-εξόδου στα οποία θα εκπαιδευτεί το δίκτυο.

6.6 *Μελέτη ενός νευρωνικού δικτύου αναγνώρισης προσώπων*

Προκειμένου να αποκτήσουμε μια πιο συγκεκριμένη εικόνα του πώς λειτουργούν και τι επιτυγχάνουν τα δίκτυα πρόσθιας τροφοδότησης είναι σκόπιμο να εξετάσουμε αναλυτικά ένα συγκεκριμένο δίκτυο. Ένα πολύ καλό παράδειγμα τέτοιου δικτύου αποτελεί το δίκτυο αναγνώρισης προσώπων που ανέπτυξε ο Garrison Cottrell (1991). Το εν λόγω δίκτυο προσομοιώνει με αρκετή επιτυχία μια ικανότητα που είναι ιδιαίτερα ανεπτυγμένη στους ανθρώπους. Το πλεονέκτημα του δικτύου που ανέπτυξε ο Cottrell είναι ότι αποτελεί ένα σχετικά απλό και εποπτικό παράδειγμα. Δεν είναι τυχαίο εξάλλου ότι το συγκεκριμένο δίκτυο υιοθετείται ως κύριο παράδειγμα από τον Paul Churchland (1995) για να εξηγήσει τις έννοιες της διανυσματικής κωδικοποίησης και της διανυσματικής επεξεργασίας, έννοιες που, σύμφωνα με αυτόν, συνδέονται άμεσα με την έννοια της κατανεμημένης αναπαράστασης.

6.6.1 *Ανάλυση του προβλήματος*

Αν σκεφτούμε ότι τα οπτικά μας όργανα είναι ευαίσθητα μόνο σε διαφορές χρώματος και φωτεινότητας γίνεται φανερό ότι η ικανότητα αναγνώρισης προσώπων είναι μια πολύ σύνθετη γνωσιακή ικανότητα. Δεν πρόκειται για την αντίληψη κάποιου απλού ερεθίσματος που μπορεί να επιτελεστεί στο επίπεδο των οπτικών αισθητηρίων οργάνων. Όπως επισημαίνει και ο Churchland (1995, σελ. 35): "... κανένα κύτταρο του αμφιβληστροειδούς δεν ανταποκρίνεται σε κάποια από τις ποικίλες όψεις ενός προσώπου που είναι σημαντικές για την αναγνώριση προσώπων". Οι πληροφορίες που είναι σημαντικές για την αναγνώριση προσώπων δεν μπορεί παρά να βρίσκονται στον συνολικό σχηματισμό ενεργοποίησης των κυττάρων του αμφιβληστροειδούς. Το ερώτημα, λοιπόν, είναι με ποια διαδικασία εξάγονται και ποιος είναι αυτές οι πληροφορίες. Η ερευνητική υπόθεση που διατυπώνει ο Churchland είναι ότι η αναγνώριση προσώπων επιτελείται μέσω διαδοχικών μετατροπών του εκάστοτε σχηματισμού ενεργοποίησης των κυττάρων του αμφιβληστροειδούς από διαδοχικούς πληθυσμούς νευρώνων του εγκεφάλου. Μέσα από αυτή τη διαδικασία φτάνουμε σε έναν πληθυσμό νευρώνων στον χρονικό λοβό που είναι υπεύθυνοι αποκλειστικά για την αναγνώριση προσώπων: "Οι πολλές συναπτικές συνδέσεις μεταξύ των κυττάρων του αμφιβληστροειδούς και των "προσωπικών κυττάρων" στο χρονικό λοβό, φιλτράρουν και

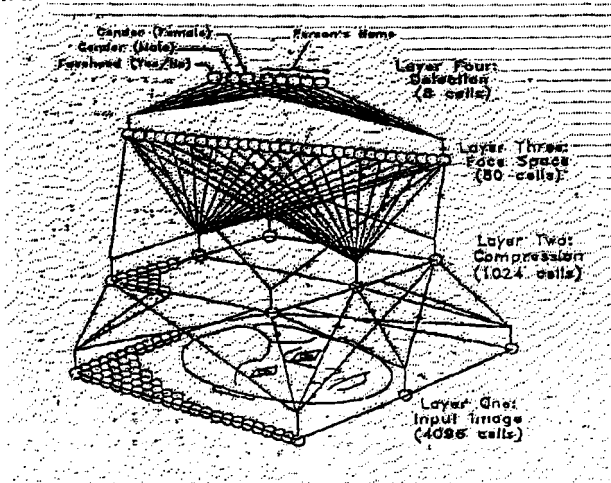
μετασχηματίζουν την εισερχόμενη πληροφορία κατά τέτοιο τρόπο που τα "προσωπικά" κύτταρα ανταποκρίνονται μόνο στις σημαντικές διαστάσεις της δομής των προσώπων που άρρητα κωδικοποιούνται στο συνολικό σχηματισμό ενεργοποίησης του αμφιβληστροειδούς" (Churchland 1995, σελ. 36)

Το δίκτυο του Cottrell παρέχει ένα πρώτο μοντέλο για το πώς μπορεί να επιτελείται μια τέτοια διαδικασία.

6.6.2 Περιγραφή του δικτύου

Το δίκτυο εικονίζεται στο σχήμα 3. Αποτελείται από τέσσερα επίπεδα. Το επίπεδο εισόδου προσομοιώνει τον αμφιβληστροειδή και είναι ουσιαστικά ένα πλέγμα από 4096 "κύτταρα" - κόμβους. Κάθε κόμβος μπορεί να λάβει 256 διαφορετικούς βαθμούς ενεργοποίησης που αντιστοιχούν ουσιαστικά σε βαθμούς φωτεινότητας. Έτσι το επίπεδο εισόδου τροφοδοτείται με φωτογραφίες που τεμαχίζονται σε 4096 εικονοστοιχεία (pixels). Στο δεύτερο επίπεδο γίνεται μια συμπίεση των πληροφοριών σε 1024 κύτταρα. Το τρίτο επίπεδο είναι αυτό όπου κατεξοχήν γίνεται η αναγνώριση προσώπων και αποτελείται από 80 κόμβους. Τέλος, το επίπεδο εξόδου είναι αυτό με το οποίο επικοινωνούμε με το δίκτυο. Αποτελείται από οκτώ κόμβους. Ο πρώτος από τα αριστερά απαντά στο ερώτημα ο εάν πρόκειται για φωτογραφία προσώπου ή όχι. Ο δεύτερος και ο τρίτος για το αν πρόκειται για φωτογραφία άνδρα ή γυναίκας αντίστοιχα. Και, τέλος, οι υπόλοιποι πέντε κόμβοι εξάγουν κάποιον αριθμητικό συνδυασμό που αντιστοιχεί στο όνομα του προσώπου της εισαγόμενης φωτογραφίας.

Σχήμα 3



6.6.3 Εκπαίδευση του δικτύου

Το δίκτυο έχει εκπαιδευτεί με τη μέθοδο της αντιστροφής τροφοδότησης του σφάλματος σε ένα σύνολο 77 φωτογραφιών. Από αυτές 64 ήταν φωτογραφίες προσώπων και οι υπόλοιπες 13 δεν περιείχαν πρόσωπα. Οι φωτογραφίες προσώπων ανήκαν σε 11 διαφορετικά πρόσωπα.

6.6.4 Απόδοση του δικτύου

Μετά τη λήξη της εκπαίδευσης το δίκτυο παρουσίαζε 100% επιτυχία στην αναγνώριση των φωτογραφιών στις οποίες εκπαιδεύτηκε. Επίσης, παρουσίαζε 98% επιτυχία στην αναγνώριση νέων φωτογραφιών των έντεκα προσώπων.

Προκειμένου να ελεγχθεί η ικανότητα του δικτύου να γενικεύει τροφοδοτήθηκε με φωτογραφίες νέων προσώπων και σκηνών. Παρατηρήθηκε ότι είχε 100% επιτυχία στην αναγνώριση του εάν πρόκειται για φωτογραφία προσώπου ή όχι και 81% επιτυχία στην αναγνώριση του φύλου.

Μια τρίτου τύπου εξέταση αφορούσε την ικανότητα του δικτύου να αναγνωρίζει οικείες φωτογραφίες προσώπων από τις οποίες έλειπε ένα σημαντικό τμήμα τους. Για το σκοπό αυτό τροφοδοτήθηκε με φωτογραφίες στις οποίες καλύπτονταν μια οριζόντια ζώνη που ισοδυναμούσε με το 1/5 της επιφάνειάς τους. Αυτό που παρατηρήθηκε ήταν ότι η απόδοση του δικτύου παρέμενε σχεδόν ανεπηρέαστη εκτός από τις περιπτώσεις στις οποίες είχε καλυφθεί το τμήμα των φωτογραφιών που απεικόνιζε το μέτωπο. Για αυτές τις περιπτώσεις η απόδοση του δικτύου έπεσε στο 71%, ένα αποτέλεσμα εντελώς αναπάντεχο που δείχνει ίσως ότι η διαμόρφωση των μαλλιών πάνω στο μέτωπο παίζει ένα σημαντικό ρόλο στον τρόπο που το δίκτυο κατορθώνει να ταξινομήσει τα πρόσωπα.

Πώς μπορεί κανείς να εξηγήσει τον τρόπο με τον οποίο το δίκτυο επιτυγχάνει όλες αυτές τις αξιοσημείωτες επιδόσεις; Η θέση του Churchland είναι ότι η ουσία του ζητήματος έγκειται στον τρόπο που το δίκτυο κωδικοποιεί στο εσωτερικό του τα χαρακτηριστικά των προσώπων. Ειδικότερα, η αναγνώριση επιτυγχάνεται στο τρίτο επίπεδο του δικτύου.

6.6.5 Πώς κωδικοποιούνται οι πληροφορίες στο εσωτερικό του δικτύου;

Ένας τρόπος να διερευνήσουμε τι γίνεται στο τρίτο επίπεδο του δικτύου είναι να εξετάσουμε τι πληροφορίες κωδικοποιεί κάθε ένας από τους κόμβους ξεχωριστά. Η βασική ιδέα εδώ είναι ότι κάθε κόμβος ενδεχομένως να λειτουργεί ως ανιχνευτής συγκεκριμένων χαρακτηριστικών. Πώς, όμως, μπορεί να ελεγχθεί αυτή η υπόθεση; Η λύση που προτείνει ο Cottrell είναι σχετικά απλή: για τη μέγιστη ενεργοποίηση

ενός κόμβου του κρυμμένου επιπέδου μπορούμε να υπολογίσουμε το προτεινόμενο ερέθισμά του, αφού γνωρίζουμε τα βάρη όλων των συνδέσεων αυτού με τους κόμβους εισόδου²⁷⁴. Το αποτέλεσμα είναι, εκ πρώτης όψεως, μη αναμενόμενο. Κάθε κόμβος του κρυμμένου επιπέδου δεν ανιχνεύει κάποιο εντοπισμένο χαρακτηριστικό των προσώπων, αλλά αναπαριστά μια ολόκληρη "προσωποειδή" μορφή. Αυτή η μορφή παρότι μοιάζει με πρόσωπο δεν αντιστοιχεί σε κάποιο συγκεκριμένο και, όπως λέει ο Churchland (1995, σελ. 47), "ενσωματώνει μια ποικιλία σαφώς ολιστικών χαρακτηριστικών ή διαστάσεων των προσώπων, διαστάσεων για τις οποίες η καθημερινή γλώσσα δεν διαθέτει επαρκές λεξιλόγιο."²⁷⁵ Η Metcalfe, συνεργάτιδα του Cottrell, ονόμασε το προτεινόμενο ερέθισμα κάθε κόμβου του κρυμμένου επιπέδου "όλον" (holon) για να εκφράσει τον ολιστικό του χαρακτήρα²⁷⁶.

Ωστόσο, αν σκεφτούμε πώς λειτουργεί ένα νευρωνικό δίκτυο, ο ολιστικός χαρακτήρας των προτεινόμενων ερεθισμάτων κάθε κόμβου του κρυμμένου επιπέδου φαίνεται να εξηγείται άμεσα. Σε κάθε τέτοιο κόμβο φτάνουν πληροφορίες από όλους τους κόμβους του επιπέδου εισόδου, συνεπώς από όλα τα εικονοστοιχεία της εκάστοτε φωτογραφίας. Είναι εντελώς αναμενόμενο, λοιπόν, ότι κάθε κόμβος του κρυμμένου επιπέδου θα διαθέτει ορισμένες πληροφορίες από όλη την έκταση της φωτογραφίας. Αντίστροφα, οι πληροφορίες από κάθε εικονοστοιχείο μιας φωτογραφίας πηγαιίνουν σε όλους τους κόμβους του κρυμμένου επιπέδου. Ήδη, με αυτές

²⁷⁴ Αυτό το σημείο είναι, εκ πρώτης όψεως, προβληματικό. Έστω ότι στον κόμβο k_i η έξοδος της συνάρτησης ενεργοποίησης $f(x_i)$ είναι y_i . Με βάση αυτή την έξοδο, μπορούμε να υπολογίσουμε την είσοδο x_i . Το πρόβλημα ξεκινά από 'δω και πέρα. Ξέρουμε ότι το x_i είναι ένας γραμμικός συνδυασμός των βαρών που καταλήγουν στον κόμβο k_i , δηλαδή ότι ισχύει $x_i = \sum_j w_{ij}$. Στόχος μας είναι να βρούμε τα a_j δηλαδή τις ενεργοποιήσεις των κόμβων εισόδου. Είναι προφανές ότι η παραπάνω εξίσωση έχει άπειρες λύσεις. Το ζήτημα αυτό σίγουρα απαιτεί περαιτέρω διερεύνηση η οποία, ωστόσο, υπερβαίνει το πλαίσιο του κεφαλαίου.

²⁷⁵ Ένα ερώτημα δημιουργείται από την επισήμανση του Churchland (1995, σελ. 47) ότι "... διαφορετικές φωτογραφίες του ίδιου προσώπου παρουσιάζόμενες στο επίπεδο εισόδου θα παράγουν ουσιαστικά το ίδιο διάνυσμα ενεργοποίησης κατά μήκος των κυττάρων του δεύτερου επιπέδου, επιτρέποντας έτσι στα κύτταρα εξόδου στο τρίτο επίπεδο να αναγνωρίσουν ορθά το συγκεκριμένο άτομο". Αν σε κάθε κόμβο του κρυμμένου επιπέδου αντιστοιχεί ένα μοναδικό προτεινόμενο ερέθισμα τότε πώς γίνεται να αποκτά μέγιστη ενεργοποίηση με διαφορετικές φωτογραφίες του ίδιου προσώπου; Αυτό θα ισχυρε μόνο αν το προτεινόμενο ερέθισμα υπήρχε σε κάθε μία από τις διαφορετικές φωτογραφίες - πράγμα που εκ πρώτης όψεως δεν φαίνεται να ισχύει. Ένα ανάλογο πρόβλημα δημιουργείται όταν συγκρίνουμε τα προτεινόμενα ερεθίσματα κόμβων διαφορετικών δικτύων που έχουν εκπαιδευτεί με το ίδιο σύνολο φωτογραφιών αλλά είχαν διαφορετικά αρχικά βάρη και τα οποία καταλήγουν να έχουν την ίδια σχεδόν απόδοση.

²⁷⁶ Αναφέρεται στο Churchland 1995, σελ. 48.

τις παρατηρήσεις, έχουμε αρχίσει να θίγουμε το ερώτημα για τη φύση των κατανεμημένων αναπαραστάσεων.

6.6.5.1 Η σημασιολογία του χώρου φάσεων

Ο Churchland (1995, σελ. 48), παρότι διαπιστώνει την ιδιάζουσα αυτή σχέση της μορφής ενεργοποίησης των κόμβων του κρυμμένου επιπέδου με ό,τι υποτίθεται ότι αναπαριστούν, δεν δίνει ιδιαίτερη βαρύτητα, αλλά ανατρέχει σε αυτό που ονομάζει "σημασιολογία του χώρου των φάσεων" (state space semantics). Για την ακρίβεια, ο Paul και η Patricia Churchland θεωρούν κάθε αναφορά σε κατανεμημένες αναπαραστάσεις ως ισοδύναμη με αυτό που οι ίδιοι αποκαλούν διανυσματική κωδικοποίηση. Χαρακτηριστικό είναι το ακόλουθο απόσπασμα:

"Την τελευταία δεκαετία, η έρευνα γύρω από το ερώτημα πώς οι νευρώνες αναπαριστούν έχει στραφεί στην ιδέα της *διανυσματικής κωδικοποίησης*. Γνωστή επίσης ως "κατανεμημένη αναπαράσταση", "αναπαράσταση στο χώρο των φάσεων" και "πολυδιαστατική αναπαράσταση", η έννοια της διανυσματικής κωδικοποίησης έχει καταστεί τόσο ο δούρειος ίππος για τη κριτική της τοπικής κωδικοποίησης, όσο και ένα νέο πλαίσιο για να ερμηνευτούν εντός του τα νευροβιολογικά δεδομένα" (Churchland, & Sejnowski 1992, σελ. 163)²⁷⁷

Η διανυσματική κωδικοποίηση μπορεί να παρασταθεί γεωμετρικά στον χώρο των φάσεων, δηλαδή σε ένα n -διάστατο χώρο, όπου n ο αριθμός των κόμβων του κρυμμένου επιπέδου. Συγκεκριμένα, κάθε διάσταση αυτού του χώρου αναπαριστά τον βαθμό ενεργοποίησης a ενός κόμβου του κρυμμένου επιπέδου. Αν, για παράδειγμα, υπάρχουν πέντε διαφορετικές τιμές ενεργοποίησης που μπορεί να λάβει κάποιος κόμβος, τότε η διάσταση με την οποία αναπαρίσταται αυτός ο κόμβος έχει αντίστοιχα πέντε τιμές. Έτσι, κάθε μορφή ενεργοποίησης των κόμβων του κρυμμένου επιπέδου προσδιορίζει ένα μοναδικό σημείο στο n -διάστατο χώρο φάσεων που έχει ως συντεταγμένες τούς εκάστοτε βαθμούς ενεργοποίησης όλων των κόμβων (a_1, a_2, \dots, a_n). Ο Churchland θεωρεί ότι οι κόμβοι του κρυμμένου επιπέδου ενός νευρωνικού δικτύου αναπαριστούν διάφορες καταστάσεις τοποθετώντας κάθε μια από αυτές σε κάποιο σημείο ενός n -διάστατου χώρου. Το ίδιο, σύμφωνα πάντα με τον Churchland, κάνει και ο εγκέφαλος:

²⁷⁷ Βλ. επίσης Churchland, Paul 1998, σελ. 6: "ένας συγκεκριμένος *σχηματισμός* ενεργοποίησης = ένα συγκεκριμένο *διάνυσμα* ενεργοποίησης = ένα συγκεκριμένο *σημείο* στο χώρο ενεργοποιήσεων. Όλοι αυτοί είναι διαφορετικοί τρόποι για να μιλήσει κανείς για το ίδιο πράγμα: τη μονάδα αναπαράστασης, ως φυσικά και όχι ως σημασιολογικά περιγραφόμενη"

"Η βασική ιδέα ... είναι ότι ο εγκέφαλος αναπαριστά διάφορες όψεις της πραγματικότητας μέσω μιας θέσης σε έναν κατάλληλο χώρο φάσεων" (Churchland, Paul 1989a, σελ. 78-79)

Επίσης, οι νοητικές διαδικασίες δεν θεωρούνται τίποτε άλλο παρά μετασχηματισμοί των συντεταγμένων των αναπαραστάσεων:

"...ο εγκέφαλος εκτελεί υπολογισμούς με τέτοιες αναπαραστάσεις μέσω γενικών μετασχηματισμών συντεταγμένων από τον ένα χώρο φάσεων στον άλλον" (Churchland, Paul 1989a, σελ. 78-79)

Είναι φανερό ότι ο τρόπος που ο Churchland κατανοεί τις νοητικές διαδικασίες εξαρτάται ουσιαδώς από τον τρόπο που κατανοεί τις αναπαραστάσεις στα νευρωνικά δίκτυα. Έτσι, οποιαδήποτε κριτική της θεώρησης του Churchland περί αναπαραστάσεων αφορά και την αντίστοιχη θεώρηση περί νοητικών διαδικασιών. Γι' αυτό θα αρκестούμε στην ανάδειξη των προβλημάτων που παρουσιάζουν οι υποτιθέμενες αναπαραστάσεις του χώρου των φάσεων.

Καταρχάς, οι αναπαραστάσεις στο χώρο των φάσεων, προθεωρητικά μιλώντας, δεν έχουν τίποτε το κατανοημένο: κάθε περιεχόμενο αντιστοιχείται σε ένα και μοναδικό σημείο του χώρου και αντίστροφα κάθε σημείο του χώρου αναπαριστά ένα και μοναδικό περιεχόμενο. Οτιδήποτε, λοιπόν, και αν είναι οι αναπαραστάσεις στον χώρο των φάσεων το σίγουρο είναι ότι δεν πρόκειται για κατανοημένες αναπαραστάσεις.

Όμως υπάρχει ένα ακόμα σημαντικότερο πρόβλημα: τα διάφορα σημεία του χώρου των φάσεων δεν αποτελούν τον τρόπο που το νευρωνικό δίκτυο αναπαριστά κάτι, αλλά τον τρόπο που εμείς αναπαριστούμε τις διάφορες καταστάσεις του νευρωνικού δικτύου. Μάλιστα, δεν πρόκειται ούτε καν για κάποιον ιδιαίτερα πρωτότυπο τρόπο απεικόνισης: κάθε δυναμικό σύστημα μπορεί να περιγραφεί με την βοήθεια ενός χώρου φάσεων. Αποτελεί κοινό τόπο στις φυσικές επιστήμες ότι προκειμένος να μελετηθεί ένα σύνθετο σύστημα απαιτείται ο εντοπισμός ενός συνόλου μεταβλητών με βάση τις οποίες σχηματίζεται μια εξίσωση που περιγράφει την συμπεριφορά του συστήματος και η οποία με τη σειρά της μπορεί να απεικονισθεί σε ένα n -διάστατο χώρο, όπου n ο αριθμός των διακριτών μεταβλητών. Για παράδειγμα, μπορούμε να περιγράψουμε την συμπεριφορά ενός αερίου με ένα τρισδιάστατο διάγραμμα πίεσης, όγκου και θερμοκρασίας. Κάθε σημείο αυτού του χώρου μπορεί να αναπαριστά καταστάσεις του αερίου. Και οποιαδήποτε μεταβολή της κατάστασης του αερίου αναπαρίσταται με την μετόπιση από ένα σημείο του τρισδιάστατου χώρου σε κάποιο άλλο. Ωστόσο, κανείς δεν θα ισχυριζόταν ότι το ίδιο το αέριο προκειμένου να φτάσει από τη μια κατάσταση στην άλλη χρησιμοποιεί αναπαραστάσεις! Ό,τι ισχύει για τα αέρια ισχύει κατ' αντιστοιχία και για τα νευρωνικά δίκτυα όσον αφορά τον χώρο των φάσεων.

6.7 Τι είναι κατανεμημένη αναπαράσταση;

Ας επιχειρήσουμε, λοιπόν, εκ νέου να εντοπίσουμε σε τι συνίστανται οι κατανεμημένες αναπαραστάσεις. Πώς, με άλλα λόγια, αναπαριστάται η γνώση στα νευρωνικά δίκτυα μέσω κατανεμημένων αναπαραστάσεων; Το ζήτημα αυτό αποτελεί σήμερα την καρδιά της έρευνας στα νευρωνικά δίκτυα. Χαρακτηριστική είναι η απάντηση του Sejnowski, ενός πολύ σημαντικού ερευνητή στην περιοχή των νευρωνικών δικτύων, στο ερώτημα "πώς θα έμοιαζε μια αναπαράσταση κατανεμημένη σε χιλιάδες τμήματα;":

"Αυτό αποτελεί σήμερα το επίκεντρο της έρευνας στα μοντέλα παράλληλα κατανεμημένης επεξεργασίας και τα συνδεδεστικά μοντέλα. Μας προκύπτουν πολλές αναπάντεχες απαντήσεις που ποτέ δεν θα μπορούσαμε να μαντέψουμε αν βασιζόμασταν στο μοντέλο της μνήμης που βρίσκει κανείς σε ένα ψηφιακό υπολογιστή, δηλαδή, μια πολύ συγκεκριμένη και ακριβής θέση για κάθε γεγονός" (Baumgartner & Payr, 1995, σελ. 222)

6.7.1 Πέντε προσεγγίσεις των κατανεμημένων αναπαραστάσεων

Ένα σημαντικό πρόβλημα που διαπιστώνει κανείς άμεσα αν εξετάσει τη βιβλιογραφία για τα νευρωνικά δίκτυα είναι ότι οι ερευνητές ενώ χρησιμοποιούν κατά κόρον τον όρο "κατανεμημένη αναπαράσταση", ελάχιστοι από αυτούς κάνουν τον κόπο να προσδιορίσουν τι σημαίνει. Αλλά ακόμα και όσοι ανέλαβαν αυτό το έργο κατέληξαν σε προτάσεις οι οποίες είναι είτε αρκετά ασαφείς, ώστε να μην επδέχονται μία μοναδική ερμηνεία, είτε ανεπαρκείς. Έτσι, το ερώτημα για το τι συνιστά κατανεμημένη αναπαράσταση παραμένει μέχρι σήμερα εν πολλοίς ανοικτό.

Ο van Gelder, ο σημαντικότερος ίσως μελετητής των κατανεμημένων αναπαραστάσεων, στο αντίστοιχο λήμμα στην εγκυκλοπαίδεια των γνωστικών επιστημών του MIT²⁷⁸ διακρίνει πέντε διαφορετικές προσεγγίσεις όσον αφορά το τι είναι κατανεμημένη αναπαράσταση. Ας δούμε αναλυτικότερα κάθε μία από αυτές.

1. **Βασική έννοια κατανομής:** η αναπαραστασιακή μορφή εκτείνεται σε πολλές υπολογιστικές μονάδες και κάθε υπολογιστική μονάδα εμπλέκεται στην αναπαράσταση πολλών οντοτήτων.

Ένας αφετηριακός ορισμός της έννοιας της κατανεμημένης αναπαράστασης βρίσκεται στο κεφάλαιο με τίτλο "Distributed Representations" της θεωρούμενης ως βιβλίου του συνδεδεσμού *PDP: Explorations in the Microstructure of Cognition* (Rumelhart, McClelland

²⁷⁸ Wilson & Keil 1999.

and the PDP Research Group 1986): "Κάθε οντότητα αναπαρίσταται μέσω μιας μορφής ενεργοποίησης που εκτείνεται (distributed over) σε πολλές υπολογιστικές μονάδες, και κάθε υπολογιστική μονάδα εμπλέκεται (is involved) στην αναπαράσταση πολλών διαφορετικών οντοτήτων" (Hinton, McClelland & Rumelhart 1986, σελ. 77). Όπως εύστοχα επισημαίνει ο van Gelder (1991, σελ. 36) προκειμένου ο συγκεκριμένος ορισμός να αποφύγει την κυκλικότητα πρέπει η λέξη "distributed over" να κατανοηθεί ως "εκτείνεται" και όχι ως "κατανέμεται".

2. **Υπέρθεση (superposition):** διαφορετικά περιεχόμενα αναπαρίστανται με μία και την αυτή αναπαραστασιακή μορφή. Μία άλλη ιδέα που προτείνεται από πολλούς ερευνητές²⁷⁹ είναι ότι το χαρακτηριστικό γνώρισμα που διαφοροποιεί τις κατανεμημένες αναπαραστάσεις από όλα τα υπόλοιπα είδη αναπαραστάσεων είναι η υπέρθεση, το γεγονός δηλαδή ότι πολλά διαφορετικά περιεχόμενα μπορούν να αναπαρασταθούν με μία και την αυτή αναπαραστασιακή μορφή. Για παράδειγμα, ένα νευρωνικό δίκτυο, όταν έχει πλέον περάσει την περίοδο εκπαίδευσης, αναπαριστά όλα τα διαφορετικά περιεχόμενα με ένα μοναδικό και αμετάβλητο σύνολο βαρών.

3. **Ισοδυναμικότητα (equipotentiality)**²⁸⁰: κάθε τμήμα της αναπαραστασιακής μορφής έχει το ίδιο περιεχόμενο με εκείνο που έχει η αναπαραστασιακή μορφή στο σύνολό της

Ορισμένοι μεμονωμένοι ερευνητές²⁸¹ ανιχνεύουν την ιδιαιτερότητα των κατανεμημένων αναπαραστάσεων στην ισοδυναμικότητά τους. Η ιδέα της ισοδυναμικότητας προέρχεται από τα ολογράφηματά. Ένα ολογράφημα μιας σκηνης είναι μια αποτύπωση της σκηνης σε μια φωτογραφική πλάκα κατά τέτοιο τρόπο ώστε αν πέσει πάνω στην πλάκα μια ακτίνα laser (μονοχρωματικό φως με υψηλό βαθμό συμφωνίας) παρουσιάζεται στον χώρο μπροστά από τη πλάκα τρισδιάστατη η σκηνη²⁸². Ωστόσο, αν κοιτάξουμε την φωτογραφική πλάκα με κοινό φως δεν θα δούμε την εικόνα της σκηνης αλλά μια θολούρα. Μια χαρακτηριστική ιδιότητα των ολογραφημάτων είναι ότι

²⁷⁹ Βλ. ενδεικτικά Murdock 1979 και Kohonen 1984.

²⁸⁰ Η έννοια αυτή εισήχθη από τον Lashley (1963/1929). Συγκεκριμένα, ο Lashley διατύπωσε της αρχή της νευρωνικής ισοδυναμικότητας η οποία είναι: "η εμφανιζόμενη ικανότητα κάθε άθικτου τμήματος μιας λειτουργικής περιοχής να εκτελεί, με ή χωρίς μείωση της απόδοσης, τις λειτουργίες που χάνονται από την καταστροφή του όλου" (Lashley 1963/1929, σελ. 25)

²⁸¹ Plate 1993.

²⁸² Το ολογράφημα μιας σκηνης αποτελεί ουσιαστικά την φωτογράφιση του στάσιμου κύματος που δημιουργείται από τη συμβολή μιας ακτίνας laser που αντανακλάται από τη σκηνη και μιας άλλης ακτίνας του ίδιου μήκους κύματος που πέφτει κατευθείαν στη φωτογραφική πλάκα.

κάθε²⁸³ τμήμα τους περιέχει πληροφορίες για το σύνολό της σκηνής που αναπαρίσταται στο ολογράφημα²⁸⁴. Ακριβώς αυτή η ιδιότητα ονομάζεται ισοδυναμικότητα. Για παράδειγμα, αν από ένα ολογράφημα που αποτυπώνει ένα σπίτι κόψουμε το πάνω μισό τμήμα του και το φωτίσουμε με μια ακτίνα laser, τότε δεν θα παρουσιαστεί στο χώρο το κάτω μισό του σπιτιού αλλά ολόκληρο το σπίτι, με τη μόνη διαφορά ότι θα είναι λιγότερο ευδιάκριτο. Σύμφωνα με τη συγκεκριμένη προσέγγιση λοιπόν, τα ολογραφήματα αποτελούν ένα παράδειγμα κατανεμημένων αναπαραστάσεων. Μια αναπαράσταση είναι κατανεμημένη αν κάθε τμήμα της αναπαριστά ό,τι και η αναπαράσταση στο σύνολό της. Η ύπαρξη αυτής της ιδιότητας φαίνεται να εξηγεί το γιατί τα νευρωνικά δίκτυα εξακολουθούν να λειτουργούν ικανοποιητικά ακόμα και όταν καταστραφεί ένα τμήμα τους.

4. Σημασιολογικά μικροχαρακτηριστικά (semantic microfeatures): κάθε κόμβος αναπαριστά ένα διαφορετικό χαρακτηριστικό της προς αναπαράσταση οντότητας

Σύμφωνα με αυτή την προσέγγιση, οι προς αναπαράσταση οντότητες αναλύονται σε ένα μεγάλο αριθμό χαρακτηριστικών. Μια τόσο λεπτομερής ανάλυση δεν απαντά συνήθως στον τρόπο που κατανοούνται οι οντότητες στην καθημερινή ζωή, γι' αυτό και τα συγκεκριμένα χαρακτηριστικά συνηθίζεται να ονομάζονται μικροχαρακτηριστικά. Ένα στοιχειώδες παράδειγμα²⁸⁵ τέτοιου τύπου ανάλυσης αναφέρει ο Smolensky (1988b, σελ. 332) προκειμένου να συγκεκριμενοποιήσει την έννοια της κατανεμημένης αναπαράστασης. Το παράδειγμα αφορά το πώς αναπαρίσταται ένα φλιτζάνι που περιέχει καφέ με τη βοήθεια εννέα κόμβων. Κάθε ένας από τους κόμβους αναπαριστά την ύπαρξη ή μη ενός μικροχαρακτηριστικού ανάλογα με το αν είναι ενεργοποιημένος ή όχι. Στο

²⁸³ Για την ακρίβεια κάθε τμήμα τους που είναι μεγαλύτερο από το μήκος κύματος των συμβαλλόντων ακτινοβολιών.

²⁸⁴ Ακριβέστερα, κάθε τμήμα ενός ολογραφήματος αποθηκεύει πληροφορίες για ολόκληρη τη σκηνή αλλά από τη δικιά του οπτική γωνία. Το σημαντικό, βέβαια, είναι ότι δύο γειτονικές οπτικές γωνίες παρέχουν πάνω-κάτω το ίδιο περιεχόμενο. Δεν χρειάζεται να ανατρέξουμε στα ολογραφήματα για να παρατηρήσουμε αυτό το φαινόμενο: όταν κοιτάμε ένα τοπίο τότε σε μια πολύ περιορισμένη επιφάνεια που είναι τα μάτια μας φτάνουν πληροφορίες από μια τεράστια έκταση. Αν μετακινήσουμε λίγο, τότε στα μάτια μας φτάνουν πάλι οι ίδιες σχεδόν πληροφορίες. Με άλλα λόγια, σε κάθε περιοχή της τάξεως της έκτασης των ματιών μας που βρίσκεται εντός ενός σχετικά μεγάλου χώρου φτάνουν οπτικές πληροφορίες από κάθε σημείο του παρατηρούμενου τοπίου.

²⁸⁵ Επίσης, για κάπως πιο ρεαλιστικά παραδείγματα, βλέπε την αναπαράσταση τύπων δωματίων από τους Rumelhart, Smolensky, McClelland & Hinton (1986) και την αναπαράσταση ρηματικών και ονοματικών εκφράσεων από τους McClelland & Kawamoto (1986).

παρακάτω σχήμα παρουσιάζονται συγκεντρωτικά οι κόμβοι και τα μικροχαρακτηριστικά που αυτοί αναπαριστούν (ενεργοποιημένοι θεωρούνται οι κόμβοι που εμφανίζονται με μαύρες κουκίδες).

Πίνακας 1: Κατανεμημένη αναπαράσταση φλιτζανιού με καφέ

Κόμβοι	Μικροχαρακτηριστικά
•	Όρθιο δοχείο
•	Ζεστό υγρό
ο	Γυαλί σε επαφή με ξύλο
•	Κυρτή επιφάνεια από πορσελάνη
•	Καυτερή μυρωδιά
•	Καφέ υγρό σε επαφή με πορσελάνη
ο	Στενόμακρο ασημένιο αντικείμενο
•	Χερούλι σε μέγεθος δακτύλου
•	Καφέ υγρό με κυρτές πλευρές και πάτο

Ο Smolensky αναπτύσσει το συγκεκριμένο παράδειγμα για να δείξει ότι οι κατανεμημένες αναπαραστάσεις διαθέτουν συστατική δομή η οποία όμως διαφέρει ριζικά από την συντακτική δομή των γλωσσικών αναπαραστάσεων. Ένα φλιτζάνι με καφέ δεν αναπαρίσταται με δύο ξεχωριστά σύμβολα ένα για το φλιτζάνι και ένα για τον καφέ που τοποθετούνται το ένα δίπλα στο άλλο. Αυτό δεν σημαίνει ωστόσο ότι η κατανεμημένη αναπαράσταση που φαίνεται στον πίνακα 1 δεν περιλαμβάνει με κάποιο άλλο τρόπο την αναπαράσταση του καφέ. Ο Smolensky προκειμένου να προσδιορίσει την εν λόγω αναπαράσταση δέχεται την ακόλουθη πρόταση του Pylyshyn: "Θα πάρουμε μια κατανεμημένη αναπαράσταση ενός φλιτζανιού με καφέ και θα αφαιρέσουμε από αυτή μια κατανεμημένη αναπαράσταση ενός φλιτζανιού χωρίς καφέ και θα ονομάσουμε ό,τι απομένει 'συνδεδεστική αναπαράσταση του καφέ'." (Smolensky 1988b, σελ. 332). Στους πίνακες 2 και 3 βλέπουμε αντίστοιχα τις κατανεμημένες αναπαραστάσεις ενός φλιτζανιού χωρίς καφέ και του καφέ.

Πίνακας 2: Κατανεμημένη αναπαράσταση φλιτζανιού χωρίς καφέ

Κόμβοι	Μικροχαρακτηριστικά
•	Όρθιο δοχείο
ο	Ζεστό υγρό
ο	Γυαλί σε επαφή με ξύλο
•	Κυρτή επιφάνεια από πορσελάνη
ο	Καυτερή μυρωδιά
ο	Καφέ υγρό σε επαφή με πορσελάνη
ο	Στενόμακρο ασημένιο αντικείμενο
•	Χερούλι σε μέγεθος δακτύλου
ο	Καφέ υγρό με κυρτές πλευρές και πάτο

Πίνακας 3: Κατανεμημένη αναπαράσταση του καφέ

Κόμβοι	Μικροχαρακτηριστικά
ο	Όρθιο δοχείο
•	Ζεστό υγρό
ο	Γυαλί σε επαφή με ξύλο
ο	Κυρτή επιφάνεια από πορσελάνη
•	Καυτερή μυρωδιά
•	Καφέ υγρό σε επαφή με πορσελάνη
ο	Στενόμακρο ασημένιο αντικείμενο
ο	Χερούλι σε μέγεθος δακτύλου
•	Καφέ υγρό με κυρτές πλευρές και πάτο

Αυτό που πρέπει να κρατήσουμε από το παραπάνω παράδειγμα είναι ότι η συγκεκριμένη κατανεμημένη αναπαράσταση του καφέ εξαρτάται ουσιωδώς από το πλαίσιο μέσα στο οποίο χρησιμοποιείται. Η συγκεκριμένη αναπαράσταση του καφέ περιέχει στοιχεία από την αλληλεπίδραση του καφέ με το φλιτζάνι μέσα στο οποίο βρίσκεται. Ο καφές δεν είναι πάντα ένα καφέ υγρό σε επαφή με πορσελάνη ούτε ένα καφέ υγρό με κυρτές πλευρές και πάτο, αλλά ούτε και ένα ζεστό υγρό. Όπως επισημαίνει ο Smolensky (1988b, σελ. 334): "... η αναπαράσταση του καφέ που εξάγουμε από την κατασκευή που βασίζεται στο φλιτζάνι με καφέ οδηγεί σε μια διαφορετική αναπαράσταση του καφέ από εκείνη που εξάγουμε από άλλες κατασκευές οι οποίες έχουν a priori ισόδυναμο καθεστώς. Αυτό σημαίνει ότι αν θέλεις να μιλήσεις για συνδυαστική αναπαράσταση του καφέ σε αυτό το κατανεμημένο σχήμα, πρέπει να μιλήσεις για *οικογένεια μορφών κατανεμημένης ενεργοποίησης*. Αυτό που συγκρατεί όλες τις επιμέρους αναπαραστάσεις του καφέ μαζί δεν είναι τίποτα άλλο από μια *οικογενειακή ομοιότητα*".

Φυσικά, θα πρέπει να παρατηρήσουμε ότι η θέση του Smolensky ότι οι έννοιες που αναπαρίστανται εξαρτώνται ουσιωδώς από το πλαίσιο μέσα στο οποίο χρησιμοποιούνται, προκύπτει αποκλειστικά από το είδος του παραδείγματος που επιλέχτηκε. Μπορούμε κάλλιστα να φανταστούμε άλλα παραδείγματα όπου οι έννοιες δεν εξαρτώνται από το πλαίσιο χρήσης τους. Έτσι, αν το προς κωδικοποίηση περιεχόμενο είναι τα κεφαλαία γράμματα του ελληνικού αλφαβήτου, τότε μπορούμε να θεωρήσουμε ως μικροχαρακτηριστικά ευθύγραμμα και κυκλικά τμήματα συγκεκριμένων προσανατολισμών με βάση τα οποία μπορούν να ανακατασκευαστούν όλα τα γράμματα.

5. "Αδρομερής" κωδικοποίηση (coarse coding): κάθε κόμβος αναπαριστά περισσότερα του ενός χαρακτηριστικά και κάθε χαρακτηριστικό αναπαρίσταται από περισσότερους του ενός κόμβους

Μια άλλη απόπειρα εξήγησης των κατανεμημένων αναπαραστάσεων γίνεται μέσω της έννοιας της "αδρομερούς" κωδικοποίησης. Σύμφωνα με αυτή την μέθοδο κωδικοποίησης, κάθε κόμβος του επιπέδου εισόδου αναπαριστά περισσότερα του ενός χαρακτηριστικά (εξ' ου και ο όρος "αδρομερής κωδικοποίηση"). Επίσης,

αντίστροφα, κάθε χαρακτηριστικό αναπαρίσταται από περισσότερους του ενός κόμβους. Η ιδιαιτερότητα της μεθόδου έγκειται στο ότι η τομή των υποσυνόλων των χαρακτηριστικών που αναπαριστούν γειτονικοί κόμβοι δεν είναι το κενό σύνολο. Χαρακτηριστικό είναι το παρακάτω χωρίο από το κλασικό άρθρο των Hinton, McClelland & Rumelhart (1986) "Distributed Representations": "Μια κατανεμημένη αναπαράσταση χρησιμοποιεί μια μονάδα για ένα σύνολο πραγμάτων (items) και κωδικοποιεί έμμεσα (implicitly) ένα συγκεκριμένο πράγμα ως την τομή των συνόλων που αντιστοιχούν στις ενεργοποιημένες μονάδες." (σελ. 264)

Ας δούμε ένα απλοϊκό παράδειγμα²⁸⁶. Έστω ότι η προς αναπαράσταση οντότητα έχει πέντε χαρακτηριστικά (α, β, γ, δ και ε) και ότι οι διαθέσιμοι κόμβοι είναι πέντε (κ₁, κ₂, κ₃, κ₄ και κ₅). Αν εφαρμόσουμε τη "μέθοδο της αδρομερούς" κωδικοποίησης μπορούμε να αναθέσουμε σε κάθε κόμβο την αναπαράσταση τριών χαρακτηριστικών:

στον κόμβο κ₁ των χαρακτηριστικών α, β, γ

στον κόμβο κ₂ των χαρακτηριστικών β, γ, δ

στον κόμβο κ₃ των χαρακτηριστικών γ, δ, ε

στον κόμβο κ₄ των χαρακτηριστικών δ, ε, α

στον κόμβο κ₅ των χαρακτηριστικών ε, α, β

Έτσι, λοιπόν, βλέπουμε ότι ενώ η ενεργοποίηση ενός μεμονωμένου κόμβου δεν καθορίζει με βεβαιότητα ποια χαρακτηριστικά αναπαρίστανται, η ταυτόχρονη ενεργοποίηση περισσότερων κόμβων δίνει, αντίθετα, την δυνατότητα να προσδιορισθεί πλήρως ποια είναι τα αναπαριστώμενα χαρακτηριστικά. Για παράδειγμα, η ενεργοποίηση των κ₁, κ₂ και κ₃ προσδιορίζει μετά βεβαιότητας ότι αναπαρίσταται το χαρακτηριστικό γ, δηλαδή η τομή των συνόλων {α, β, γ}, {β, γ, δ} και {γ, δ, ε}²⁸⁷.

Η μέθοδος της αδρομερούς κωδικοποίησης έλκει την καταγωγή της από τον τρόπο που το οπτικό σύστημα κωδικοποιεί τον εγγύς ερεθισμό σε ένα πρώιμο στάδιο. Συγκεκριμένα, κάθε υποδοχέας του ματιού είναι ευαίσθητος σε περισσότερα του ενός είδη ερεθισμάτων, π.χ. σε ευθείες συγκεκριμένους προσανατολισμούς. Το σύνολο των ερεθισμάτων στα οποία είναι ευαίσθητος ένας υποδοχέας ονομάζεται πεδίο υποδοχής (receptive

²⁸⁶ Για ένα περισσότερο χρήσιμο παράδειγμα βλ. Touretzky & Hinton 1988.

²⁸⁷ Ο Smolensky (1991, σελ. 213) θεωρεί ότι η αδρομερής κωδικοποίηση ή, με τη διαία του ορολογία, η συζευκτική (conjunctive) κωδικοποίηση, αποτελεί μια επιμέρους περίπτωση της γενικότερης τεχνικής κωδικοποίησης με βάση το τανυστικό γινόμενο (tensor product). Η τεχνική αυτή έχει ως στόχο να επιλύσει το πρόβλημα του πώς μια σύνθετη δομή μπορεί να αντιστοιχηθεί σε ένα διάγραμμα που παριστά βαθμούς ενεργοποίησης κόμβων κατά τέτοιο τρόπο ώστε η συστατική δομή να είναι διαθέσιμη προς επεξεργασία από το νευρωνικό δίκτυο. Μέσα σε αυτό το πλαίσιο, μια κατανεμημένη αναπαράσταση κατανοείται ως μια αντιστοίχιση ενός συνόλου διακριτών δομών σε έναν διανυσματικό χώρο (Smolensky 1991, σελ. 212).

field). Η ενεργοποίηση ενός μεμονωμένου υποδοχέα δεν παρέχει ακριβείς πληροφορίες για τον προσανατολισμό της ανιχνευόμενης ευθείας. Αν όμως ενεργοποιηθούν όλοι οι υποδοχείς που έχουν δυνατότητα να ανιχνεύουν, μεταξύ άλλων, ευθείες του συγκεκριμένου προσανατολισμού τότε καθορίζεται με βεβαιότητα ο προσανατολισμός της ευθείας.

6.7.2 Κριτική αποτίμηση των πέντε προσεγγίσεων της έννοιας της κατανομής

Αναπτύξαμε λεπτομερειακά τις πέντε προσεγγίσεις που αναφέρονται στην έγκυρη και σύγχρονη εγκυκλοπαίδεια της γνωσιακής επιστήμης του MIT όσον αφορά το τι είναι κατανεμημένη αναπαράσταση. Ωστόσο, στο συγκεκριμένο λήμμα δεν επιχειρείται κάποια σύγκριση των προτεινόμενων προσεγγίσεων, ούτε, πολύ περισσότερο, προκρίνεται κάποια από αυτές²⁸⁸. Αυτό φυσικά δεν είναι καθόλου τυχαίο μια και το τι είναι κατανεμημένη αναπαράσταση παραμένει ένα ανοικτό ερευνητικό ζήτημα. Οι πέντε προσεγγίσεις, εκ πρώτης όψης, μοιάζουν τόσο διαφορετικές μεταξύ τους, ώστε δεν φαίνεται να υπάρχει κάποιο κοινό μέτρο σύγκρισης. Στο υπόλοιπο του κεφαλαίου θα επιχειρήσουμε να συστηματοποιήσουμε όσο το δυνατόν τη συζήτηση για το τι είναι κατανεμημένη αναπαράσταση εντοπίζοντας κάποιες διακρίσεις οι οποίες, όπως θα δείξουμε, είναι ιδιαίτερα διαφωτιστικές για την σύγκριση και αξιολόγηση των παραπάνω προσεγγίσεων.

Πριν προχωρήσουμε, ωστόσο, μπορούμε άμεσα να απλοποιήσουμε το πρόβλημα απορρίπτοντας την πρώτη από τις πέντε προσεγγίσεις ως ανεπαρκή. Η προσέγγιση αυτή στηρίζεται σε δύο συνθήκες. Η πρώτη συνθήκη είναι ότι η μορφή των κατανεμημένων αναπαράστάσεων εκτείνεται σε πολλές αναπαραστασιακές μονάδες²⁸⁹ και

²⁸⁸ Ωστόσο, σε ένα προγενέστερο κείμενο του ο van Gelder (1991) προκρίνει την έννοια της υπέρθεσης ως την ουσία της κατανομής. Το γεγονός ότι στο κείμενο που γράφει για την εγκυκλοπαίδεια του MIT δεν παίρνει θέση δείχνει ότι το όλο ζήτημα κάθε άλλο παρά έχει διευθετηθεί.

²⁸⁹ Είναι "ανησυχητικά" μεγάλος ο αριθμός των ερευνητών που εξηγούν την έννοια της κατανεμημένης αναπαράστασης με βάση αυτή τη συνθήκη. Βλ. ενδεικτικά McLeod, Plunkett & Rolls (1998, σελ. 62-63): "Κατανεμημένη αναπαράσταση είναι εκείνη στην οποία όλα τα στοιχεία στον σχηματισμό (pattern) χρησιμοποιούνται για την κωδικοποίηση ενός συγκεκριμένου ερεθίσματος ... κανένα στοιχείο δεν μπορεί να χρησιμοποιηθεί [μεμονωμένα] για τον καθορισμό του ερεθίσματος. Εφόσον η πληροφορία για το ποιο ερέθισμα είναι παρόν κατανέμεται στον πληθυσμό των στοιχείων, αποκαλείται κατανεμημένη αναπαράσταση", Kosslyn & Hatfield (1984, σελ. 1030): "Στον εγκέφαλο, η καλύτερη τρέχουσα εικασία είναι ότι η πληροφορία δεν αποθηκεύεται σε μια δεδομένη θέση. Μάλλον, η πληροφορία εμφανίζεται να κατανέμεται σε πολυάριθμες θέσεις", Rolls & Treves (1998, σελ. 31): "Μια κατανεμημένη αναπαράσταση είναι εκείνη η αναπαράσταση στην οποία η πυροδότηση ή

η δεύτερη συνθήκη ότι κάθε υπολογιστική μονάδα εμπλέκεται στην αναπαράσταση πολλών οντοτήτων.

Η πρώτη συνθήκη δεν συλλαμβάνει κάποια ειδοποιό διαφορά των κατανεμημένων αναπαραστάσεων από τις υπόλοιπες. Για παράδειγμα, είναι συχνό φαινόμενο μια συμβολική αναπαράσταση να αποθηκεύεται σε έναν ψηφιακό υπολογιστή κομματιαστά σε διαφορετικά σημεία της μνήμης. Το γεγονός αυτό, ωστόσο, δεν επηρεάζει την φύση της αναπαράστασης η οποία παραμένει συμβολική²⁹⁰. Μια προφανής διαφορά με τις κατανεμημένες αναπαραστάσεις είναι ότι αν χαθεί ένα μικρό τμήμα της διασκορπισμένης συμβολικής αναπαράστασης τότε χάνεται, στην καλύτερη περίπτωση, ένα αντίστοιχο τμήμα του περιεχομένου της. Αντίθετα, στις κατανεμημένες αναπαραστάσεις δεν χάνεται κάποιο τμήμα του περιεχομένου απλώς μειώνεται η ευκρίνεια του περιεχομένου συνολικά. Η πρώτη συνθήκη λοιπόν σίγουρα δεν είναι επαρκής για τον ορισμό των κατανεμημένων αναπαραστάσεων²⁹¹, ενδεχομένως όμως να αποτελεί αναγκαία συνθήκη - θέση που θα διερευνηθεί παρακάτω.

Επίσης, όσον αφορά τη δεύτερη συνθήκη που λέει ότι κάθε υπολογιστική μονάδα εμπλέκεται στην αναπαράσταση πολλών οντοτήτων, αυτή προσδιορίζει ουσιαστικά μια μορφή υπέρθεσης οπότε υπάγεται στην δεύτερη προσέγγιση.

Μια κατανεμημένη αναπαράσταση, εφόσον θεωρείται αναπαράσταση, οφείλει να διαθέτει τα κύρια χαρακτηριστικά μιας αναπαράστασης εν γένει: αναπαραστασιακή μορφή και αναπαραστασιακό περιεχόμενο. Από μια πρώτη εξέταση των τελευταίων τεσσάρων προσεγγίσεων που προτείνονται στο λήμμα της εγκυκλοπαίδειας του MIT, διαπιστώνουμε ότι τόσο η μορφή όσο και το περιεχόμενο των αναπαραστάσεων ποικίλουν.

Καταρχάς, παρατηρούμε ότι ως αναπαραστασιακή μορφή άλλοτε θεωρείται ο βαθμός ενεργοποίησης των κόμβων²⁹² και άλλοτε τα βάρη των συνδέσεων²⁹³. Μάλιστα οι Churchland και Grush για να μην αποκλείσουν την μία μορφή εις βάρος της άλλης στο λήμμα που γράφουν στην εγκυκλοπαίδεια του MIT, φτάνουν στην ακόλουθη συμβιβαστική πρόταση:

ενεργοποίηση όλων των στοιχείων στο διάγραμμα χρησιμοποιείται για την κωδικοποίηση ενός συγκεκριμένου ερεθίσματος".

²⁹⁰ Βλ. Rich & Knight 1991, σελ. 520.

²⁹¹ Βλ. και van Gelder (1991), σελ. 38.

²⁹² Π.χ. βλ. Churchland, Paul (1998), σελ. 6: "Στα μοντέλα γνώσης που βασίζονται στα νευρωνικά δίκτυα, η πρωταρχική μονάδα αναπαράστασης και ο πρωταρχικός φορέας του σημασιολογικού περιεχομένου είναι ο συνολικός σχηματισμός των ταυτόχρονων επιπέδων ενεργοποίησης συγκεντρωμένων νευρώνων ενός δεδομένου πληθυσμού".

²⁹³ Π.χ. βλ. Smolensky 1988a, σελ. 780: "Η γνώση σε ένα συνδυαστικό σύστημα έγκειται στα βάρη των συνδέσεών του".

"Τι είναι, με νευρωνικούς όρους, οι αναπαραστάσεις; Ενώ η προσέγγιση που βασίζεται στην τεχνητή νοημοσύνη εξισώνει τις αναπαραστάσεις με τα σύμβολα, έναν όρο καλώς ορισμένο στο πλαίσιο των συμβατικών υπολογιστών, οι συνδετιστές συνειδητοποίησαν ότι το "σύμβολο" ουσιαστικά δεν ορίζεται σε νευροβιολογικό πλαίσιο. Κατά συνέπεια, επεδίωξαν να αναπτύξουν μια νέα θεωρία αναπαράστασης κατάλληλη για την νευροβιολογία. Έτσι, υπέθεσαν ότι τρέχουσες αναπαραστάσεις (αυτές που συμβαίνουν τώρα) είναι σχηματισμοί ενεργοποιήσεων των μονάδων ενός νευρωνικού δικτύου που προσδιορίζονται ως ένα διάνυσμα $\langle x, y, z, \dots \rangle$, όπου κάθε στοιχείο του διανύσματος καθορίζει το επίπεδο ενεργοποίησης ενός κόμβου. Αντίθετα, οι αποθηκευμένες αναπαραστάσεις πιστεύεται ότι εξαρτώνται από τον σχηματισμό των βαρών μεταξύ των μονάδων. Με νευρωνικούς όρους, αυτά τα βάρη είναι η ισχύς των συναπτικών συνδέσεων μεταξύ των νευρώνων" (Churchland, P. S. & Grush 1999)

Από την άλλη μεριά, όσον αφορά το περιεχόμενο των κατανεμημένων αναπαραστάσεων, βλέπουμε ότι σε ορισμένες περιπτώσεις αυτό που αναπαρίσταται θεωρείται ότι είναι οντότητες (ή, γενικότερα, καταστάσεις πραγμάτων) στο σύνολό τους, ενώ σε άλλες ότι είναι χαρακτηριστικά ή μικροχαρακτηριστικά των οντοτήτων (ή καταστάσεων πραγμάτων).

Τέλος, το όλο πρόβλημα επιτείνεται από τις διαφορετικές απόψεις που κυκλοφορούν όσον αφορά τη σχέση της μορφής των κατανεμημένων αναπαραστάσεων με το περιεχόμενό τους. Οι προσεγγίσεις που παρουσιάσαμε παραπάνω αποτελούν, κάθε μια τους, έναν διαφορετικό τρόπο συσχέτισης μορφής και περιεχομένου. Είναι πολύ σημαντικό να επισημάνουμε σε αυτό το σημείο ότι οι περισσότεροι ερευνητές θεωρούν ότι η ειδοποιός διαφορά των κατανεμημένων αναπαραστάσεων από τα υπόλοιπα είδη αναπαράστασης θα πρέπει να αναζητηθεί ακριβώς σε αυτή την ειδική σχέση μορφής - περιεχομένου²⁹⁴. Όπως και στην περίπτωση των νοητικών συμβόλων και των νοητικών ειδώλων, οι κατανεμημένες αναπαραστάσεις εκλαμβάνονται ως ένας διαφορετικός τύπος κατοπτρισμού (mirroring) του αναπαραστασιακού περιεχομένου στην αναπαραστασιακή μορφή. Μια τέτοια παραδοχή, ωστόσο, προϋποθέτει ότι υπάρχει κάποια συναίνεση ως προς το ποια είναι η μορφή και ποιο το περιεχόμενο των κατανεμημένων αναπαραστάσεων - κάτι που όπως είδαμε δεν ισχύει. Εκ πρώτης όψεως, υπάρχουν τέσσερες διαφορετικές περιπτώσεις που πρέπει να διακριθούν:

Πίνακας 4

²⁹⁴ Βλ. van Gelder 1999.

	Περίπτωση 1 ^α	Περίπτωση 2 ^α	Περίπτωση 3 ^α	Περίπτωση 4 ^α
Μορφή	Βαθμός ενεργοποίησης των κόμβων	Βαθμός ενεργοποίησης των κόμβων	Βάρη των συνδέσεων	Βάρη των συνδέσεων
Περιεχόμενο	Χαρακτηριστικά της οντότητας (ή κατάστασης πραγμάτων)	Οντότητα (ή κατάσταση πραγμάτων) στο σύνολό της	Χαρακτηριστικά της οντότητας (ή κατάστασης πραγμάτων)	Οντότητα (ή κατάσταση πραγμάτων) στο σύνολό της

Κάθε μία από τις παραπάνω περιπτώσεις επιδέχεται ένα διαφορετικό σύνολο τρόπων συνδυασμού μορφής και περιεχομένου. Ωστόσο, υπάρχει και μια άλλη διάκριση η οποία, όπως θα δείξουμε, είναι πολύ σημαντική για την διευκρίνιση της έννοιας της κατανεμημένης αναπαράστασης. Πρόκειται για τη διάκριση μεταξύ των αναπαραστάσεων του επιπέδου εισόδου / εξόδου και των αναπαραστάσεων των κρυμμένων επιπέδων. Ένα μεγάλο μέρος της σύγχυσης γύρω από την έννοια της κατανεμημένης αναπαράστασης οφείλεται στο γεγονός ότι οι ερευνητές αναφέρονται, χωρίς να κάνουν διάκριση, άλλοτε στις αναπαραστάσεις των επιπέδων εισόδου / εξόδου και άλλοτε σε εκείνες των κρυμμένων επιπέδων.

Γιατί όμως υποστηρίζουμε ότι οι αναπαραστάσεις του επιπέδου εισόδου / εξόδου πρέπει να διακριθούν από τις αναπαραστάσεις των κρυμμένων επιπέδων; Ένας πρώτος λόγος αφορά τη μορφή των αναπαραστάσεων. Συγκεκριμένα, όταν πρόκειται για τις αναπαραστάσεις του επιπέδου εισόδου / εξόδου, ως μορφή των αναπαραστάσεων δεν μπορεί να θεωρηθεί παρά μόνο ο βαθμός ενεργοποίησης των κόμβων, καθότι οι συνδέσεις ξεκινούν μετά από την είσοδο και πριν από την έξοδο. Αντίθετα στην περίπτωση των αναπαραστάσεων των κρυμμένων επιπέδων ως μορφή μπορεί να θεωρηθεί είτε ο βαθμός ενεργοποίησης των κόμβων είτε τα βάρη των συνδέσεων. Έτσι, λοιπόν, σε ό,τι αφορά τις κατανεμημένες αναπαραστάσεις του επιπέδου εισόδου / εξόδου, οι προς εξέταση περιπτώσεις από τον Πίνακα 4 είναι μόνο οι δύο πρώτες.

Ένας δεύτερος λόγος, πολύ σημαντικότερος κατά τη γνώμη μας, για τον οποίο θα πρέπει να διακριθούν οι αναπαραστάσεις που χρησιμοποιούνται στο επίπεδο εισόδου / εξόδου από εκείνες των κρυμμένων επιπέδων, είναι ότι ενώ οι πρώτες καθορίζονται ρητά από τον προγραμματιστή, οι δεύτερες προκύπτουν από την εσωτερική δυναμική του ίδιου του νευρωνικού δικτύου. Προφανώς, ο λόγος που οι αναπαραστάσεις των επιπέδων εισόδου / εξόδου πρέπει να ορισθούν ρητά από τον προγραμματιστή είναι ότι μέσω αυτών των επιπέδων "επικοινωνεί" το νευρωνικό δίκτυο με τον χρήστη του. Πρέπει κατά συνέπεια να έχει ορισθεί με σαφήνεια ο κώδικας επικοινωνίας. Αυτό που επισημαίνουμε, με άλλα λόγια, είναι ότι στην περίπτωση των αναπαραστάσεων του επιπέδου εισόδου / εξόδου είναι σαφές τόσο το περιεχόμενο και η μορφή τους, όσο και η μεταξύ τους σχέση, επειδή ακριβώς οι συγκεκριμένες αναπαραστάσεις πρέπει να ορισθούν ρητά από τον προγραμματιστή. Έτσι, λοιπόν, το πρόβλημα του τι και πώς

αναπαρίσταται στα επίπεδα εισόδου / εξόδου είναι, στην εκάστοτε περίπτωση, ουσιαστικά ένα λυμένο πρόβλημα. Αντίθετα, στην περίπτωση των αναπαραστάσεων των κρυμμένων επιπέδων δεν είναι ο προγραμματιστής που καθορίζει το τι και πώς αναπαρίσταται, αλλά ούτε και χρειάζεται να γίνει κάτι τέτοιο προκειμένου να μπορεί κανείς να χρησιμοποιήσει ένα νευρωνικό δίκτυο. Οι ερευνητές χρησιμοποιούν κατά κόρον και με αποτελεσματικότητα τα τεχνητά νευρωνικά δίκτυα χωρίς να χρειάζεται να γνωρίζουν το τι και πώς αναπαρίσταται στα κρυμμένα επίπεδα. Για το λόγο αυτό το πραγματικό πρόβλημα με τις κατανεμημένες αναπαραστάσεις δεν αφορά τις αναπαραστάσεις των επιπέδων εισόδου και εξόδου αλλά εκείνες των κρυμμένων επιπέδων.

Με τις παραπάνω δύο επισημάνσεις πιστεύουμε ότι καθίσταται σαφές γιατί δεν θα πρέπει να συγχέουμε τις αναπαραστάσεις που χρησιμοποιούνται στα επίπεδα εισόδου και εξόδου με τις κατανεμημένες αναπαραστάσεις των κρυμμένων επιπέδων.

Υπό το πρίσμα αυτής της διάκρισης, αν στρέψουμε ξανά την προσοχή μας στις τέσσερες υπό εξέταση προσεγγίσεις του τι είναι κατανεμημένη αναπαράσταση, διαπιστώνουμε ότι οι δύο τελευταίες (σημασιολογικά μικροχαρακτηριστικά και αδρομερής κωδικοποίηση) αποτελούν πρωταρχικά τρόπους κωδικοποίησης του περιεχομένου στα επίπεδα εισόδου και εξόδου. Αυτός είναι και ο λόγος που οι εν λόγω προσεγγίσεις είναι τόσο συγκεκριμένες και σαφείς. Όπως ήδη επισημάναμε, πρόκειται, ουσιαστικά, για δύο λυμένα προβλήματα. Μάλιστα, αν εξετάσουμε λίγο καλύτερα τη μέθοδο της αδρομερούς κωδικοποίησης διαπιστώνουμε ότι αυτή, στη βάση της, στηρίζεται σε μια έννοια τοπικής υπέρθεσης. Ανατρέχοντας στο σχετικό παράδειγμα, οι πέντε αναπαραστάσεις των α, β, γ, δ και ε μερικώς επικαλύπτονται: κάθε κόμβος συμμετέχει στην αναπαράσταση περισσότερων του ενός χαρακτηριστικών. Όσον αφορά δε την αναπαράσταση οντοτήτων μέσω σημασιολογικών μικροχαρακτηριστικών, είναι λάθος να θεωρείται καν ως παράδειγμα κατανεμημένης αναπαράστασης: εφόσον κάθε μικροχαρακτηριστικό κωδικοποιείται από έναν μεμονωμένο κόμβο -την ελάχιστη δηλαδή μονάδα αναπαράστασης που μπορεί να χρησιμοποιηθεί σε ένα νευρωνικό δίκτυο-, η αναπαράσταση του μικροχαρακτηριστικού είναι εντελώς εντοπισμένη. Με άλλα λόγια, κάθε περιεχόμενο (εν προκειμένω, κάθε μικροχαρακτηριστικό) απονέμεται σε μια και μοναδική μορφή (εν προκειμένω, τον βαθμό ενεργοποίησης ενός κόμβου) και αντίστροφα, ο βαθμός ενεργοποίησης κάθε κόμβου αναπαριστά ένα και μοναδικό μικροχαρακτηριστικό. Μια αναπαράσταση οντοτήτων μέσω σημασιολογικών μικροχαρακτηριστικών είναι υπ' αυτή την έννοια τόσο κατανεμημένη όσο και η συμβολική περιγραφή "άρρεν, ενήλικος, ανύπανδρος" που χρησιμοποιείται για να ορίσει την έννοια εργένης²⁹⁵.

²⁹⁵ Βλ. Fodor & Pylyshyn 1988.

Κατά συνέπεια, η αναπαράσταση οντοτήτων στο επίπεδο εισόδου με τη βοήθεια σημασιολογικών μικροχαρακτηριστικών θα πρέπει να θεωρηθεί ως μια ειδική περίπτωση εντοπισμένης αναπαράστασης.

Απομένουν λοιπόν προς εξέταση οι έννοιες της υπέρθεσης και της ισοδυναμικότητας οι οποίες φαίνεται να χαρακτηρίζουν γενικότερα τις κατανεμημένες αναπαράστασεις. Στο βαθμό μάλιστα που στα κρυμμένα επίπεδα και, γενικότερα, στο εσωτερικό του δικτύου υποτίθεται ότι έχουμε γνήσια κατανεμημένες αναπαράστασεις, οι συγκεκριμένες έννοιες θα πρέπει να αποτελέσουν τον βασικό μας οδηγό.

Τόσο η υπέρθεση όσο και η ισοδυναμικότητα προσδιορίζουν τρόπους αντιστοίχισης μορφής και περιεχομένου. Σχέση υπέρθεσης υφίσταται όταν διαφορετικά περιεχόμενα αντιστοιχίζονται στην ίδια μορφή, ενώ σχέση ισοδυναμικότητας όταν διαφορετικά τμήματα της μορφής αναπαριστούν ολόκληρο το περιεχόμενο, δηλαδή ό,τι αναπαριστά η μορφή στο σύνολό της.

Και οι δύο έννοιες επιδέχονται διαβάθμιση. Αν ονομάσουμε M_i και Π_i το τυχόν τμήμα της μορφής M και του περιεχομένου Π της κατανεμημένης αναπαράστασης αντίστοιχα, τότε υπάρχουν οι ακόλουθες κύριες δυνατότητες:

1. Ορισμένα M_i αναπαριστούν ορισμένα από τα Π_i ενός περιεχομένου (τοπική υπέρθεση)
2. Κάθε M_i αναπαριστά ορισμένα από τα Π_i ενός περιεχομένου
3. Κάθε M_i αναπαριστά όλα τα Π_i ενός περιεχομένου
4. Κάθε M_i αναπαριστά όλο το Π (ισοδυναμικότητα)
5. Κάθε M_i αναπαριστά μέρος από τα Π_i όλων των αναπαριστώμενων περιεχομένων
6. Κάθε M_i αναπαριστά όλα τα Π_i όλων των αναπαριστώμενων περιεχομένων
7. Κάθε M_i αναπαριστά όλα τα αναπαριστώμενα περιεχόμενα
8. Η μορφή M στο σύνολό της αναπαριστά όλα τα αναπαριστώμενα περιεχόμενα (υπέρθεση)

Από τις οκτώ αυτές περιπτώσεις φαίνεται να υπάρχει κάποιου είδους σχέση μεταξύ ισοδυναμικότητας και υπέρθεσης. Ωστόσο, αν ξεκινήσουμε από τις πρωτοτυπικές μορφές ισοδυναμικότητας και υπέρθεσης (βλ. 4 και 8 αντίστοιχα) δεν διαπιστώνεται άμεσα αυτή η συγγένεια γιατί η μεν ισοδυναμικότητα έχει να κάνει με την αναπαράσταση ενός περιεχομένου, η δε υπέρθεση με την αναπαράσταση πολλών. Αν όμως εξετάσουμε τις υπόλοιπες περιπτώσεις εκεί δεν είναι πλέον σαφές το εάν πρόκειται για περιπτώσεις ισοδυναμικότητας ή υπέρθεσης. Ποια είναι λοιπόν η σχέση ισοδυναμικότητας και υπέρθεσης; Θα δείξουμε ότι η ισοδυναμικότητα μπορεί να θεωρηθεί ως μια ειδική περίπτωση υπέρθεσης.

Ένας τρόπος να δειχθεί αυτή η θέση είναι να ξεκινήσουμε από την κλασική περίπτωση ισοδυναμικότητας που είναι η τέταρτη από τις παραπάνω περιπτώσεις: κάθε M_i αναπαριστά όλο το Π , ό,τι δηλαδή αναπαριστά το M στο σύνολό του. Αν τώρα το Π θεωρηθεί ως σύνολο όλων των Π_i που το αποτελούν, τότε σε κάθε M_i θα μπορούσε να θεωρηθεί ότι γίνεται υπέρθεση όλων των Π_i , μια και κάθε M_i έχει πληροφορίες για όλο το Π άρα και για κάθε Π_i . Ένας άλλος τρόπος να δειχθεί η ίδια θέση είναι αν θεωρήσουμε ως M_i το M . Μια τέτοια υπόθεση δικαιολογείται διότι σύμφωνα με τον ορισμό της κλασσικής περίπτωσης ισοδυναμικότητας κάθε M_i αναπαριστά όλο το Π , άρα και όταν $M_i = M$. Σε αυτή την περίπτωση και πάλι κάθε Π_i υπερτίθεται στο M . Άρα λοιπόν, η ισοδυναμικότητα αποτελεί μια ειδική περίπτωση υπέρθεσης²⁹⁶.

Επίσης δεν ισχύει το αντίστροφο, δηλαδή δεν αποτελεί κάθε περίπτωση υπέρθεσης περίπτωση ισοδυναμικότητας. Για παράδειγμα, αν πάρουμε την κλασσική περίπτωση υπέρθεσης σύμφωνα με την οποία η M αναπαριστά όλα τα Π , βλέπουμε ότι δεν συνεπάγεται από αυτήν ότι κάθε M_i αναπαριστά όλα τα Π . Με άλλα λόγια, στον ορισμό της υπέρθεσης προϋποτίθεται ότι απαιτούνται όλα τα τμήματα της μορφής M προκειμένου να αναπαρασταθούν όλα τα περιεχόμενα.

Συνεπώς η έννοια της υπέρθεσης φαίνεται να είναι πρωταρχικότερη από την έννοια της ισοδυναμικότητας και να μην την προϋποθέτει. Θα πρέπει να επισημάνουμε, ωστόσο, ότι το συμπέρασμά μας αυτό βασίζεται στην παραδοχή ότι κάθε Π μπορεί να αναλυθεί σε ένα σύνολο Π_i . Σε περιπτώσεις όμως που δεν ισχύει η συγκεκριμένη παραδοχή δεν ισχύει και η θέση ότι η έννοια της ισοδυναμικότητας μπορεί να εξηγηθεί μέσω της έννοιας της υπέρθεσης. Βλέπουμε, με άλλα λόγια, ότι το είδος του περιεχομένου που θέλουμε να αναπαραστήσουμε επηρεάζει τους δυνατούς τρόπους αντιστοίχισής του στην αναπαραστασιακή μορφή. Θα πρέπει λοιπόν να εξετάσουμε πιο λεπτομερειακά τα είδη περιεχομένου που μπορούν να αναπαρασταθούν από τις κατανεμημένες αναπαραστάσεις. Προτού όμως προχωρήσουμε προς αυτή την κατεύθυνση είναι σκόπιμο να εξετάσουμε τις έννοιες της ισοδυναμικότητας και της υπέρθεσης από πρακτική σκοπιά, δηλαδή να διερευνήσουμε αν ικανοποιούν τους πρακτικούς περιορισμούς που θέτει ο τρόπος λειτουργίας των νευρωνικών δικτύων.

Ας ξεκινήσουμε με την έννοια της υπέρθεσης: η μορφή M στο σύνολό της αναπαριστά όλα τα αναπαριστώμενα περιεχόμενα. Ένα πρόβλημα που δημιουργείται με αυτή τη διατύπωση είναι ότι δεν μπορεί να εξηγήσει το πώς κατορθώνουν να λειτουργούν νευρωνικά δίκτυα στα οποία έχει καταστραφεί κάποιο τμήμα τους. Για να έχουμε υπέρθεση, απαιτείται το σύνολο της αναπαραστασιακής μορφής. Το συγκεκριμένο

²⁹⁶ Βλ. van Gelder (1991, σελ. 51-52) για μια παρεμφερή απόδειξη της θέσης ότι η ισοδυναμικότητα συνεπάγεται την υπέρθεση.

πρόβλημα, εκ πρώτης όψεως, φαίνεται να λύνεται αν ισχύει η ιδιότητα της ισοδυναμικότητας: κάθε M_i αναπαριστά όλο το Π , ό,τι δηλαδή αναπαριστά το M στο σύνολό του. Ωστόσο και αυτή η ιδιότητα δεν μπορεί να θεωρηθεί ότι αρμόζει για τον χαρακτηρισμό των τεχνητών νευρωνικών δικτύων γιατί από μια άποψη είναι υπερβολικά ισχυρή και από μία άλλη υπερβολικά ασθενής:

α. Είναι υπερβολικά ισχυρή διότι πρακτικώς είναι αδύνατο να θεωρήσουμε ότι κάθε M_i αναπαριστά όλο το Π . Κανένα δίκτυο δεν λειτουργεί αν καταστραφούν, για παράδειγμα, όλοι εκτός από έναν οι κόμβοι του κρυμμένου επιπέδου. Το πόσοι κόμβοι χρειάζονται για να διατηρήσει μια ικανοποιητική λειτουργία είναι ένα εμπειρικό ζήτημα. Πάντως αυτό που πρέπει να κρατήσουμε όσον αφορά τον χαρακτηρισμό της ιδιότητας που οφείλουν να πληρούν οι αναπαραστάσεις στα τεχνητά νευρωνικά δίκτυα είναι ότι αντί να μιλάμε για "κάθε M_i " θα πρέπει να μιλάμε "για κάθε επαρκώς μεγάλο M_i "

β. Είναι υπερβολικά ασθενής διότι αφορά μόνο ένα περιεχόμενο, ενώ τα τεχνητά νευρωνικά δίκτυα, όταν καταστραφεί κάποιο τμήμα τους, συνεχίζουν να λειτουργούν ικανοποιητικά για όλα τα περιεχόμενα στα οποία έχουν εκπαιδευτεί.

Άρα, λοιπόν, τελικά η ιδιότητα που χαρακτηρίζει τις αναπαραστάσεις στα τεχνητά νευρωνικά δίκτυα δεν είναι ούτε η υπέρθεση ούτε η ισοδυναμικότητα, αλλά μια ιδιότητα που μπορεί να διατυπωθεί ως εξής:

Κάθε επαρκώς μεγάλο M_i αναπαριστά όλα τα Π .

Θα ονομάσουμε αυτή την ιδιότητα "ισχυρή υπέρθεση".

Ένας επιπλέον εμπειρικός περιορισμός που τίθεται από τη λειτουργία των δικτύων πρόσθιας τροφοδότησης και ο οποίος πρέπει να ληφθεί υπόψη στον προσδιορισμό της ισχυρής υπέρθεσης είναι ότι όταν καταστραφεί ένα τμήμα των δικτύων, παρότι αυτά συνεχίζουν να λειτουργούν, η απόδοσή τους είναι χειρότερη από εκείνη που είχαν πριν την καταστροφή. Αν θέλαμε να διατυπώσουμε αυτόν τον περιορισμό με όρους αναπαράστασης μπορούμε σε πρώτη φάση να σκεφτούμε τα ολογραφήματα: όπως είδαμε η αποκοπή μέρους της φωτογραφικής πλάκας οδηγεί στην απώλεια της ευκρίνειας ολόκληρου του περιεχομένου. Αυτό που χάνεται δηλαδή από την καταστροφή μέρους της φωτογραφικής πλάκας δεν είναι μέρος του περιεχομένου αλλά μέρος της ευκρίνειας του περιεχομένου συνολικά. Ο van Gelder (1991), προκειμένου να εκφράσει αυτή τη διάκριση με γενικότερους όρους ούτως ώστε να αφορά και την περίπτωση των νευρωνικών δικτύων, εισήγαγε τους όρους "σημασιολογικός χαρακτήρας" και "σημασιολογική ποιότητα". Ο σημασιολογικός χαρακτήρας του περιεχομένου έχει να κάνει με αυτό που διατηρείται κατά την καταστροφή μέρους του νευρωνικού δικτύου, ενώ η σημασιολογική ποιότητα του περιεχομένου έχει να κάνει με αυτό που αντιστοίχα χάνεται. Η μείωση της σημασιολογικής ποιότητας του περιεχομένου εξαρτάται άμεσα από την ποσότητα του δικτύου που έχει καταστραφεί. Ουσιαστικά η ιδέα αυτή αποτελεί μια επαναδιατύπωση, στο

πλαίσιο των νευρωνικών δικτύων, του νόμου της μαζικής δράσης (law of mass action) του Lashley (1963/1929) σύμφωνα με τον οποίο η αποτελεσματικότητα της μάθησης είναι συνάρτηση της συνολικής μάζας του εγκεφαλικού ιστού. Η ισχύς αυτού του νόμου εξασφαλίζει την βαθμιαία μείωση της απόδοσης που παρατηρείται στα νευρωνικά δίκτυα. Ενσωματώνοντας τη διάκριση μεταξύ σημασιολογικού χαρακτήρα και σημασιολογικής ποιότητας στον προσδιορισμό της ισχυρής υπέρθεσης, δηλαδή της ιδιότητας που χαρακτηρίζει τις αναπαραστάσεις στα νευρωνικά δίκτυα, φτάνουμε στην ακόλουθη διατύπωση:

Κάθε επαρκώς μεγάλο M_i αναπαριστά τον σημασιολογικό χαρακτήρα όλων των Π και όσο μεγαλύτερο είναι το M_i τόσο καλύτερη είναι η σημασιολογική ποιότητα των Π που αναπαρίστανται.

Η συγκεκριμένη διατύπωση αποτελεί έναν γενικό και όσο το δυνατόν πιο αφηρημένο χαρακτηρισμό των κατανεμημένων αναπαραστάσεων²⁹⁷. Η έννοια της ισχυρής υπέρθεσης, έχοντας προέλθει από τον συνυπολογισμό των εμπειρικών περιορισμών που τίθενται από τη λειτουργία των τεχνητών νευρωνικών δικτύων, συλλαμβάνει καλύτερα το χαρακτηριστικό γνώρισμα των κατανεμημένων αναπαραστάσεων. Έτσι, αν κανείς παραμείνει στην λογική του van Gelder αλλά και των υπολοίπων ερευνητών στο χώρο του συνδετισμού που αναζητούν την "ουσία" των κατανεμημένων αναπαραστάσεων στην ειδικού τύπου σχέση μορφής - περιεχομένου, πιστεύουμε ότι μια προσέγγιση που θα βασιστεί στην έννοια της ισχυρής υπέρθεσης είναι σαφώς πιο βελτιωμένη από όσες στηρίζονται είτε στην έννοια της υπέρθεσης είτε στην έννοια της ισοδυναμικότητας.

Ωστόσο, η πρόταση ότι η "ουσία" των κατανεμημένων αναπαραστάσεων έγκειται στην ισχυρή υπέρθεση, προκειμένου να συγκεκριμενοποιηθεί, πρέπει να αναλυθεί περαιτέρω. Συγκεκριμένα θα πρέπει να διερευνηθεί τι είδους αναπαραστασιακή μορφή και τι είδους αναπαραστασιακό περιεχόμενο προσιδιάζουν σε αυτή. Η διατύπωση στην οποία καταλήξαμε κάνει απλώς λόγο για κάποια μορφή M και κάποιο περιεχόμενο Π εν γένει. Τα ερωτήματα που απομένει να απαντηθούν είναι τι θα μπορούσε να παίζει το ρόλο της μορφής στα νευρωνικά δίκτυα και τι είδους περιεχόμενο μπορεί να εκφραστεί από τις κατανεμημένες αναπαραστάσεις;

Σε πρώτη φάση θα εξετάσουμε τι είδους αναπαραστασιακή μορφή μπορεί να υποστηρίξει την εμφάνιση ισχυρής υπέρθεσης. Αφού διερευνήσουμε αυτό το ερώτημα, θα επιχειρήσουμε να προσεγγίσουμε το

²⁹⁷ Η αναφορά στην μείωση της σημασιολογικής ποιότητας του περιεχομένου συναρτήσει του τμήματος του δικτύου που έχει καταστραφεί είναι σημαντική γιατί, μεταξύ άλλων, απαγορεύει την εφαρμογή της ιδιότητας της ισχυρής υπέρθεσης στην τετριμμένη περίπτωση της πλεοναστικής αναπαράστασης, δηλαδή της αναπαράστασης ενός περιεχομένου η οποία αποτελείται από μικρότερες αναπαραστάσεις με το ίδιο περιεχόμενο (βλ. van Gelder 1991).

πολύ δυσκολότερο ερώτημα για το αν και κατά πόσο, μέσω της ισχυρής υπέρθεσης, αναπαρίσταται κάποιο συγκεκριμένο είδος περιεχομένου.

6.7.3 Τι είδους αναπαραστασιακή μορφή προϋποτίθεται στην έννοια της ισχυρής υπέρθεσης.

Όπως είδαμε υπάρχουν δύο υπονήφιοι για τον ρόλο της αναπαραστασιακής μορφής στα νευρωνικά δίκτυα: ο βαθμός ενεργοποίησης των κόμβων και τα βάρη των συνδέσεων μεταξύ των κόμβων.

6.7.3.1 Αναπαραστασιακή μορφή: ο βαθμός ενεργοποίησης των κόμβων

Ας ξεκινήσουμε με τον βαθμό ενεργοποίησης των κόμβων. Οι περισσότεροι ερευνητές θεωρούν ότι η αναπαραστασιακή μορφή των κατανεμημένων αναπαραστάσεων είναι ο βαθμός ενεργοποίησης των κόμβων. Υπάρχουν πολλοί επιφανειακοί λόγοι που καθιστούν την πεποίθηση αυτή τόσο διαδεδομένη. Καταρχάς, όπως είδαμε, για την κωδικοποίηση των σχηματισμών εισόδου και εξόδου σε ένα νευρωνικό δίκτυο χρησιμοποιούνται αποκλειστικά οι βαθμοί ενεργοποίησης των κόμβων εισόδου και εξόδου. Δεδομένου δε ότι οι αναπαραστάσεις εισόδου και εξόδου δεν διακρίνονται ως προς το είδος τους από τις αναπαραστάσεις των κρυμμένων επιπέδων, θεωρείται ως αυτονόητο ότι οι τελευταίες θα πρέπει και αυτές να έχουν του ίδιου τύπου αναπαραστασιακή μορφή. Ένας δεύτερος λόγος, ίσως λιγότερο σημαντικός, είναι ότι υπάρχουν νευρωνικά δίκτυα, τα λεγόμενα εντοπιστικά δίκτυα (localist networks), στα οποία οι ενεργοποιήσεις εσωτερικών κόμβων χρησιμοποιούνται ρητά ως αναπαραστάσεις. Στα συγκεκριμένα δίκτυα, ωστόσο, αντίθετα με ό,τι συμβαίνει με τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης, οι αναπαραστάσεις είναι τοπικές, δηλαδή ο βαθμός ενεργοποίησης κάθε κόμβου μεμονωμένα αποτελεί μια αυτόνομη αναπαράσταση²⁹⁸. Τέλος, ένας τρίτος λόγος που ο βαθμός ενεργοποίησης των κόμβων του κρυμμένου επιπέδου εκλαμβάνεται ως η αναπαραστασιακή μορφή των κατανεμημένων αναπαραστάσεων είναι ότι προσφέρει ένα σύνολο διακριτών καταστάσεων που μπορούν να περιγραφούν με τη βοήθεια διανυσμάτων και να παρουσιαστούν εποπτικά ως ένα n -διάστατο διάνυσμα.

Το ερώτημα που θα εξετάσουμε τώρα είναι αν ο βαθμός ενεργοποίησης των κόμβων του κρυμμένου επιπέδου μπορεί όντως να αποτελεί την αναπαραστασιακή μορφή των κατανεμημένων αναπαραστάσεων, δεδομένου ότι η ιδιότητα που χαρακτηρίζει τις τελευταίες είναι η ισχυρή υπέρθεση. Υπάρχουν δύο ενδεχόμενα:

²⁹⁸ Ένα κλασικό παράδειγμα εντοπιστικού δικτύου αναφέρεται στο McClelland, Rumelhart & Hinton (1986) και αφορά την κωδικοποίηση πληροφοριών για τα μέλη δύο συμμοριών.

- A. διαφορετικά περιεχόμενα αναπαρίστανται στο ίδιο σύνολο κόμβων, κάθε φορά με διαφορετικό συνδυασμό ενεργοποιήσεων των κόμβων.
 B. διαφορετικά περιεχόμενα αναπαρίστανται στο ίδιο σύνολο κόμβων με τον ίδιο συνδυασμό ενεργοποιήσεων των κόμβων.

Δεδομένου ότι η ύπαρξη ισχυρής υπέρθεσης συνεπάγεται την ύπαρξη απλής υπέρθεσης, θα δείξουμε ότι τόσο η περίπτωση A όσο και η περίπτωση B δεν είναι συμβατές με την ύπαρξη απλής υπέρθεσης, και άρα, κατά συνέπεια, ούτε με την ύπαρξη ισχυρής υπέρθεσης.

Θα ξεκινήσουμε από την περίπτωση A. Το γεγονός ότι χρησιμοποιείται ο ίδιος αναπαραστασιακός φορέας για όλα τα περιεχόμενα δεν θα πρέπει να μας παραπλανά, γιατί αυτό που εν προκειμένω έχει σημασία είναι το εάν χρησιμοποιείται η ίδια αναπαραστασιακή μορφή για την αναπαράσταση όλων των περιεχομένων, κάτι που εξ υποθέσεως δεν ισχύει. Ένας τρόπος να δούμε τη διαφορά είναι αν φανταστούμε μια ψηφιακή οθόνη που δείχνει την ώρα. Το γεγονός ότι όλες οι διαφορετικές αναπαραστάσεις που εμφανίζονται στην οθόνη χρησιμοποιούν τον ίδιο αναπαραστασιακό φορέα σίγουρα δεν τις καθιστά κατανεμημένες²⁹⁹. Αυτό είναι ένα συχνό λάθος που ακόμα και ο van Gelder δεν φαίνεται ορισμένες φορές να συνειδητοποιεί. Για παράδειγμα, στο κλασικό του άρθρο "What is the "D" in "PDP"? A survey of the concept of distribution", αναφερόμενος στην έννοια της υπέρθεσης, λέει χαρακτηριστικά:

"Διαισθητικά, τα αναπαριστώμενα δύο διακριτών πραγμάτων υπερτίθενται εάν είναι ισοεκτατά - εάν, με άλλα λόγια, καταλαμβάνουν το ίδιο τμήμα των πόρων που διατίθενται για την αναπαραστασιακή λειτουργία. Έτσι, στα συνδετιστικά δίκτυα μπορούμε να έχουμε διαφορετικά πράγματα που αποθηκεύονται ως μορφές ενεργοποίησης πάνω στο ίδιο σύνολο μονάδων" (van Gelder 1991, σελ 43)

Το γεγονός ότι χρησιμοποιείται το ίδιο σύνολο πόρων (π.χ. οι κόμβοι του κρυμμένου επιπέδου) σημαίνει απλά ότι χρησιμοποιείται ο ίδιος αναπαραστασιακός φορέας. Όμως αυτό που έχει σημαίνοντα ρόλο δεν είναι ο φορέας αλλά η αναπαραστασιακή μορφή, δηλαδή ο βαθμός ενεργοποίησης των κόμβων, ο οποίος είναι για κάθε περιεχόμενο διαφορετικός.

²⁹⁹ Ένα άλλο παράδειγμα είναι ο τρόπος με τον οποίον καταχωρούνται οι συμβολικές αναπαραστάσεις στον κεντρικό επεξεργαστή ενός ψηφιακού υπολογιστή: ένας καταχωρητής (register) μπορεί να λάβει ένα σύνολο διακριτών καταστάσεων (συνδυασμών 0 και 1) ανάλογα με το ποια κατάσταση αναπαριστά. Η αναπαραστασιακή μορφή εν προκειμένω, δηλαδή η σημαίνουσα μορφή, είναι η μορφή "ενεργοποίησης" του καταχωρητή και όχι ο καταχωρητής εν γένει.

Ας έρθουμε τώρα στην περίπτωση Β: διαφορετικά περιεχόμενα αναπαριστανται στο ίδιο σύνολο κόμβων με τον ίδιο συνδυασμό ενεργοποιήσεων των κόμβων. Αυτή, στο βαθμό που ισχύει, αποτελεί εξ ορισμού περίπτωση υπέρθεσης. Ωστόσο, η εν λόγω περίπτωση δεν είναι ρεαλιστική: κάθε νευρωνικό δίκτυο αποκτά διαφορετικό βαθμό ενεργοποίησης στους κρυμμένους κόμβους για κάθε διαφορετική εισόδο³⁰⁰.

Βλέπουμε λοιπόν ότι αν θεωρήσουμε ως αναπαραστασιακή μορφή των κατανεμημένων αναπαραστάσεων τον βαθμό ενεργοποίησης των κόμβων του κρυμμένου επιπέδου, τότε δεν υπάρχει δυνατότητα υπέρθεσης διαφορετικών περιεχομένων στην ίδια μορφή, πόσο μάλλον δυνατότητα ισχυρής υπέρθεσης. Άρα, λοιπόν, αν δεχθούμε ότι η ισχυρή υπέρθεση είναι η ιδιότητα που χαρακτηρίζει τις κατανεμημένες αναπαραστάσεις, τότε είναι αδύνατον η αναπαραστασιακή μορφή αυτών να είναι ο βαθμός ενεργοποίησης των κόμβων του κρυμμένου επιπέδου.

6.7.3.2 Αναπαραστασιακή μορφή: τα βάρη των συνδέσεων μεταξύ των κόμβων

Ας εξετάσουμε τώρα τον δεύτερο υποψήφιο για τον ρόλο της αναπαραστασιακής μορφής, δηλαδή τα βάρη των συνδέσεων μεταξύ των κόμβων. Ως γνωστόν, αυτά παραμένουν σταθερά μετά το πέρας της περιόδου εκπαίδευσης. Αν λοιπόν όντως τα βάρη των συνδέσεων αποτελούν την αναπαραστασιακή μορφή των κατανεμημένων

³⁰⁰ Σε αυτή την περίπτωση μπορεί να ενταχθεί και η προσπάθεια του Smolensky (1991) να ορίσει την υπέρθεση όπως ορίζεται στα μαθηματικά, δηλαδή ως πρόσθεση διανυσμάτων. Συγκεκριμένα, ο Smolensky θεώρησε ότι αν έχουμε δύο διανύσματα v_1 και v_2 των βαθμών ενεργοποίησης των κόμβων που αναπαριστούν αντίστοιχα τα περιεχόμενα Π_1 και Π_2 τότε το διάνυσμα $v_1 + v_2$ αναπαριστά τα περιεχόμενα Π_1 και Π_2 . Η πρόταση του Smolensky ωστόσο έχει πολλά προβλήματα. Καταρχάς, ισχύει εναντίον της του γενικό επιχείρημα που διατυπώσαμε στο κυρίως κείμενο: ο βαθμός ενεργοποίησης των κόμβων του κρυμμένου επιπέδου είναι διαφορετικός για κάθε διαφορετικό σχηματισμό εισόδου. Επίσης δεν είναι σαφές ότι η πρόταση του Smolensky ισχύει για τους κόμβους του κρυμμένου επιπέδου, μια και όλα τα παραδείγματα που δίνει αφορούν τους κόμβους εισόδου. Τέλος, υπάρχει πρόβλημα σε σχέση με τη δυνατότητα διάκρισης των περιεχομένων που έχουν κατ' αυτόν τον τρόπο υπερθεθεί, γιατί με δεδομένο το διάνυσμα $v_1 + v_2$, δεν υπάρχει τρόπος να καθορισθεί ότι αυτό έχει προέλθει από τα v_1 και v_2 συγκεκριμένα και κατά συνέπεια ότι αναπαριστά τα Π_1 και Π_2 . Για να δώσουμε ένα απλό παράδειγμα, αν μας δοθεί ο αριθμός 576 όπως αυτός έχει προέλθει από το άθροισμα δύο αριθμών, δεν υπάρχει τρόπος να προσδιορίσουμε ποιοι είναι οι συγκεκριμένοι αριθμοί. Με άλλα λόγια, η έννοια της διανυσματικής υπέρθεσης δεν μπορεί να χρησιμεύσει για τον ορισμό των κατανεμημένων αναπαραστάσεων γιατί δεν υπάρχει τρόπος ανάκτησης των διακριτών περιεχομένων που έχουν υπερθεθεί (βλ. και Fodor & McLaughlin 1991).

αναπαραστάσεων, τότε φαίνεται να αντιμετωπίζουμε μια πρωτοτυπική περίπτωση υπέρθεσης: όλα τα διαφορετικά περιεχόμενα που μπορεί να αναπαραστήσει το δίκτυο αναπαρίστανται με μια και μοναδική μορφή. Επίσης, ακόμα και αν ένα τμήμα αυτών των συνδέσεων καταστραφεί το δίκτυο εξακολουθεί να λειτουργεί ικανοποιητικά. Αυτό σημαίνει ότι αν δεχτούμε ότι η λειτουργία του νευρωνικού δικτύου βασίζεται στη χρήση αναπαραστάσεων, η καταστροφή μέρους των συνδέσεων δεν καταστρέφει πλήρως τις εν λόγω αναπαραστάσεις. Άρα πληρούται όχι μόνο η ιδιότητα της υπέρθεσης, αλλά και εκείνη της ισχυρής υπέρθεσης.

Μην έχοντας άλλο υποψήφιο και μέχρι αποδείξεως του εναντίου, πρέπει λοιπόν να δεχτούμε ότι η αναπαραστασιακή μορφή των κατανεμημένων αναπαραστάσεων είναι τα βάρη των συνδέσεων μεταξύ των κόμβων. Εξάλλου υπάρχει και ένας ανεξάρτητος λόγος που υποδεικνύει ότι τα βάρη των συνδέσεων σχετίζονται άμεσα με τη γνώση που "αποθηκεύει" το δίκτυο: Κατά την περίοδο εκπαίδευσης του νευρωνικού δικτύου μόνο τα βάρη των συνδέσεων μεταβάλλονται απ' όλη τη δομή του δικτύου³⁰¹. Κατά συνέπεια, μπορούμε βάσιμα να υποθέσουμε ότι αν το δίκτυο αποθηκεύει κάποιου είδους γνώση υπό τη μορφή αναπαραστάσεων, η αναπαράσταση αυτής της γνώσης εξαρτάται άμεσα από τα βάρη των συνδέσεων.

Τι είδους γνώση, όμως, αποθηκεύεται σε ένα νευρωνικό δίκτυο; Με άλλα λόγια, τι είδους περιεχόμενα αναπαριστούν οι κατανεμημένες αναπαραστάσεις;

6.7.4 Τι είδους περιεχόμενο έχουν οι κατανεμημένες αναπαραστάσεις;

Μέχρι τώρα έχουμε αναφερθεί με πολύ γενικούς όρους στο περιεχόμενο των κατανεμημένων αναπαραστάσεων. Η διάκριση μεταξύ οντοτήτων στο σύνολό τους και οντοτήτων ως συνόλου μικροχαρακτηριστικών, παρότι συλλαμβάνει τους δύο βασικούς τρόπους με τους οποίους οι ίδιοι οι συνδεδειστές κατανοούν το περιεχόμενο, είναι υποτυπώδης και, όπως θα δείξουμε, προβληματική. Θα πρέπει, λοιπόν, να θέσουμε εκ νέου το ερώτημα τι είδους περιεχόμενο έχουν οι κατανεμημένες αναπαραστάσεις που χρησιμοποιούν τα νευρωνικά δίκτυα;

Πολλοί ερευνητές φαίνεται να υποθέτουν ότι η απάντηση σε αυτό το ερώτημα εξαρτάται άμεσα από το είδος του περιεχομένου που αναπαρίστανται στην *είσοδο* του νευρωνικού δικτύου. Μάλιστα, ορισμένοι προχωρούν ένα βήμα παραπέρα και μιλούν για τις κατανεμημένες αναπαραστάσεις στο εσωτερικό του δικτύου σαν να επρόκειτο ουσιαστικά για αναπαραστάσεις των σχηματισμών εισόδου. Για παράδειγμα, σε ένα από τα καλύτερα και πιο σύγχρονα εγχειρίδια για τα νευρωνικά δίκτυα, οι

³⁰¹ Να υπενθυμίσουμε ότι εξετάζουμε δίκτυα πρόσθιας τροφοδότησης στα οποία δεν μεταβάλλεται η κλίση της συνάρτησης ενεργοποίησης.

συγγραφείς, αναφερόμενοι στον τρόπο που ένα δίκτυο πρόσθιας τροφοδότησης μπορεί να επιλύσει το πρόβλημα του λογικού συνδέσμου XOR³⁰², σημειώνουν χαρακτηριστικά: "το παράδειγμα με το XOR απέδειξε ότι ένα πολυεπίπεδο δίκτυο μπορεί να σχηματίσει νέες αναπαραστάσεις ενός συνόλου σχηματισμών εισόδου στις κρυμμένες μονάδες του" (McLeod, Plunkett & Rolls 1998, σελ. 108). Καταρχάς, να επισημάνουμε ότι οι συγγραφείς εν προκειμένω θεωρούν ως μορφή των κατανεμημένων αναπαραστάσεων τον βαθμό ενεργοποίησης των κόμβων. Αυτό, όπως είδαμε, δεν μπορεί να είναι ορθό, αν η ιδιότητα που χαρακτηρίζει τις κατανεμημένες αναπαραστάσεις είναι η ισχυρή υπέρθεση. Θα παραβλέψουμε, ωστόσο, προσωρινά αυτό το στοιχείο προκειμένου να εστιάσουμε την προσοχή μας αποκλειστικά στο είδος του περιεχομένου που υποτίθεται ότι αναπαρίσταται στο εσωτερικό του δικτύου. Οι αναπαραστάσεις στο εσωτερικό του δικτύου, σύμφωνα με την παραπάνω προσέγγιση, δεν θα ήταν τίποτα άλλο από αναπαραστάσεις των αναπαραστάσεων εισόδου. Με άλλα λόγια, θεωρείται ότι ο βαθμός ενεργοποίησης των κόμβων του κρυμμένου επιπέδου αναπαριστά το περιεχόμενο που αναπαρίσταται με τον βαθμό ενεργοποίησης των κόμβων του επιπέδου εισόδου. Όμως κάτι τέτοιο φαίνεται να είναι εντελώς περιττό. Ποιος ο λόγος να δημιουργηθεί μια αναπαράσταση της αναπαράστασης εισόδου³⁰³; Τα δίκτυα πρόσθιας τροφοδότησης προκειμένου να είναι χρήσιμα πρέπει να αναπαριστούν στο εσωτερικό τους κάτι διαφορετικό από την αναπαράσταση εισόδου. Για παράδειγμα, αν η δουλειά του δικτύου είναι να κάνει γενικεύσεις, τότε οι εσωτερικές αναπαραστάσεις του έχουν ως περιεχόμενο το είδος στο οποίο υπάγονται οι εκάστοτε αναπαραστάσεις εισόδου. Με αυτή τη λογική, αν θεωρήσουμε το δίκτυο αναγνώρισης προσώπων που παρουσιάσαμε στην αρχή αυτού του κεφαλαίου, θα ήταν λάθος να πούμε ότι οι αναπαραστάσεις στο εσωτερικό του δικτύου αναπαριστούν τις διάφορες φωτογραφίες που παρουσιάζονται στην είσοδό του. Αυτό που αναπαριστούν οι εσωτερικές αναπαραστάσεις του δικτύου είναι το εάν η εκάστοτε φωτογραφία που παρουσιάζεται στην είσοδο ανήκει σε πρόσωπο ή όχι καθώς και το φύλο και το όνομα του εμφανιζόμενου προσώπου. Άρα, λοιπόν, θα πρέπει να δεχτούμε ότι το περιεχόμενο των εσωτερικών αναπαραστάσεων είναι διαφορετικό από το περιεχόμενο των αναπαραστάσεων εισόδου. Η άποψη

³⁰² Πρόκειται για το αποκλειστικό ή (exclusive or): η έκφραση $P \cdot XOR \cdot Q$ είναι αληθής αν μόνο ένα από τα P και Q είναι αληθή, ενώ σε όλες τις άλλες περιπτώσεις είναι ψευδής. Το νευρωνικό δίκτυο που προσομοιώνει το λογικό σύνδεσμο XOR αποτελείται από δύο κόμβους εισόδου που αναπαριστούν αντίστοιχα τις τιμές αληθείας των P και Q , και έναν κόμβο εξόδου που αναπαριστά την αληθοτιμή του $P \cdot XOR \cdot Q$.

³⁰³ Φυσικά ένας λόγος θα μπορούσε να είναι η συμπίεση των πληροφοριών. Ωστόσο, το ενδιαφέρον και το ζητούμενο είναι τι γίνεται μετά το στάδιο της συμπίεσης.

ότι το περιεχόμενο που αναπαριστούν οι κατανεμημένες αναπαραστάσεις είναι είτε οντότητες στο σύνολό τους, είτε οντοτήτες ως σύνολο μικροχαρακτηριστικών προέρχεται από την εσφαλμένη προκατάληψη ότι στο εσωτερικό του δικτύου αναπαρίσταται το περιεχόμενο των εισόδου.

Τι θα μπορούσε όμως να είναι το περιεχόμενο των εσωτερικών αναπαραστάσεων; Η απάντηση έχει ήδη δοθεί σε ό,τι αναφέραμε παραπάνω για το δίκτυο αναγνώρισης προσώπων: Αυτό που αναπαριστούν οι εσωτερικές αναπαραστάσεις του δικτύου είναι το εάν η εκάστοτε φωτογραφία που παρουσιάζεται στην είσοδο ανήκει σε πρόσωπο ή όχι, καθώς και το φύλο και το όνομα του εμφανιζόμενου προσώπου. Το εάν η φωτογραφία είναι φωτογραφία προσώπου ή όχι καθώς και το φύλο και το όνομα του προσώπου αποτελούν ουσιαστικά την αναπαράσταση εξόδου του δικτύου. Κατά συνέπεια, θα μπορούσαμε να πούμε ότι αυτό που αναπαριστούν οι εσωτερικές αναπαραστάσεις ενός νευρωνικού δικτύου είναι σχέσεις εισόδων με εξόδους. Με άλλα λόγια, αντιστοιχίζουν εισόδους σε εξόδους.

Η συγκεκριμένη θέση βρίσκει πρόσθετη υποστήριξη στο γεγονός ότι η μορφή των κατανεμημένων αναπαραστάσεων, σύμφωνα με όσα υποστηρίξαμε στην προηγούμενη ενότητα, δεν είναι ο βαθμός ενεργοποίησης των κόμβων του κρυμμένου επιπέδου, αλλά τα βάρη των συνδέσεων. Διότι τα βάρη των συνδέσεων διαμορφώνονται από τα ζεύγη εισόδου-εξόδου στα οποία εκπαιδεύεται το δίκτυο. Για την ακρίβεια, διαμορφώνονται από την εκάστοτε είσοδο και τη διαφορά της αντίστοιχης εξόδου από την επιθυμητή έξοδο. Άρα λοιπόν, σε μια πρώτη προσέγγιση θα πρέπει να δεχτούμε ότι οποιοδήποτε κι αν είναι το περιεχόμενο που φέρουν τα βάρη των συνδέσεων, αυτό σίγουρα δεν αφορά μόνο τις εισόδους αλλά και τις εξόδους. Μάλιστα, δεν αφορά τις εισόδους και τις εξόδους γενικώς και αορίστως, αλλά ζεύγη εισόδων και εξόδων. Αυτό δεν σημαίνει, κατ' ανάγκη, ότι τα βάρη των συνδέσεων αναπαριστούν τις ίδιες τις εισόδους και εξόδους των εκάστοτε ζευγών. Τα βάρη των συνδέσεων μπορεί απλά να θεωρηθούν ότι αναπαριστούν τον κανόνα μετασχηματισμού συγκεκριμένων εισόδων σε συγκεκριμένες εξόδους, χωρίς ταυτόχρονα να απαιτείται η αναπαράσταση και των εκάστοτε εισόδων και εξόδων.

Ένας άλλος τρόπος να νοήσουμε αυτή τη σχέση είναι να θεωρήσουμε ότι ολόκληρο το νευρωνικό δίκτυο υλοποιεί μια πολύ σύνθετη συνάρτηση - μια συνάρτηση στην οποία αν δοθεί κάποια είσοδος παράγεται η κατάλληλη έξοδος. Σύμφωνα με αυτή την ιδέα το νευρωνικό δίκτυο κατά την περίοδο εκπαίδευσης τροποποιεί τα βάρη των συνδέσεων για να μεταβάλλει κατάλληλα την συνάρτηση που υπολογίζει, ούτως ώστε για την εκάστοτε είσοδο να εξαγάγει την αρμόζουσα έξοδο. Φτάνουμε έτσι σε μια ιδέα που προωθείται απ' όσους επιχειρούν να μελετήσουν τα γνωστικά φαινόμενα με τη βοήθεια της θεωρίας δυναμικών

συστημάτων³⁰⁴. Στο παρόν κεφάλαιο δεν θα χρησιμοποιήσουμε τη μαθηματική ορολογία που προσφέρει η θεωρία δυναμικών συστημάτων³⁰⁵, αλλά θα παραμεινουμε στην διατύπωση ότι αυτό που αναπαριστούν τα βάρη των συνδέσεων σε ένα νευρωνικό δίκτυο είναι ο κανόνας μετασχηματισμού εισόδων σε εξόδους.

Η συγκεκριμένη διατύπωση, ωστόσο, δεν είναι χωρίς προβλήματα. Κατ'αρχάς, θα μπορούσε κανείς να αντιτείνει ότι είναι τόσο γενική ώστε ισχύει όχι μόνο για τα νευρωνικά δίκτυα αλλά και για τους συνήθεις ψηφιακούς υπολογιστές. Γιατί, υπό μια έννοια, και οι ψηφιακοί υπολογιστές αντιστοιχίζουν εισόδους σε εξόδους. Ωστόσο, υπάρχουν τουλάχιστον τρεις σημαντικές διαφορές όσον αφορά τον τρόπο που τα νευρωνικά δίκτυα και οι ψηφιακοί υπολογιστές επιτελούν αυτή την αντιστοίχιση:

1. τα νευρωνικά δίκτυα, σε αντίθεση με τους ψηφιακούς υπολογιστές, συνεχίζουν να λειτουργούν ακόμα και αν έχει καταστραφεί μέρος των συνδέσεών τους
2. η λειτουργία των νευρωνικών δικτύων, σε αντίθεση με εκείνη των ψηφιακών υπολογιστών, δεν μπορεί να αναλυθεί σε σημασιολογικά ερμηνεύσιμες υπολειτουργίες. Ούτε, επίσης, η λειτουργία ενός δικτύου στο σύνολό του μπορεί να περιγραφεί υπό την μορφή ενός προτασιακού κανόνα που εκφράζει σχέσεις των ιδιοτήτων της εισόδου με ιδιότητες της εξόδου του δικτύου.
3. Ένας ψηφιακός υπολογιστής έχει τη δυνατότητα για συγκεκριμένες εισόδους να εξάγει συγκεκριμένες εξόδους. Δεδομένης μιας εισόδου, αν κάποια από τις εσωτερικές λειτουργίες του ψηφιακού υπολογιστή τροποποιηθεί απειροελάχιστα, τότε ο υπολογιστής, σε αντίθεση με τα νευρωνικά δίκτυα, δεν θα καταλήξει στην ίδια έξοδο - το πιθανότερο μάλιστα είναι ότι το πρόγραμμα δεν θα λειτουργήσει καθόλου.

Μια δεύτερη ένσταση στη θέση ότι τα βάρη των συνδέσεων σε ένα νευρωνικό δίκτυο αναπαριστούν τον κανόνα μετασχηματισμού συγκεκριμένων εισόδων σε συγκεκριμένες εξόδους είναι ότι, αν την δεχτούμε, ουσιαστικά ερμηνεύουμε τα νευρωνικά δίκτυα ως "συμπεριφοριστικές" μηχανές, δηλαδή μηχανές που το μόνο που κάνουν είναι να συσχετίζουν απλώς εισόδους με εξόδους. Ωστόσο, ένα μαύρο κουτί του οποίου η εσωτερική λειτουργία είναι εντελώς ακατανόητη δεν συμβάλλει καθόλου στην κατανόηση των γνωσιακών φαινομένων. Η ένσταση αυτή, εν πολλοίς, είναι βάσιμη. Όπως επισημάναμε και παραπάνω, στα νευρωνικά δίκτυα δεν είναι δυνατή η απασύνθεση της λειτουργίας τους σε αυτόνομα υποσυστήματα καθένα από τα οποία επεξεργάζεται την είσοδο ή σημασιολογικά ερμηνεύσιμα τμήματά της. Ελάχιστη αναπαραστασιακή μονάδα είναι το σύνολο των συνδέσεων του

³⁰⁴ Βλ. παράρτημα 2.

³⁰⁵ Βλ. van Gelder 1995.

δικτύου. Τμήματα των συνδέσεων του δικτύου δεν έχουν σημασιολογική ερμηνεία. Αντίστοιχα, η μονάδα περιεχομένου "κατοπτρίζεται" στο σύνολο των συνδέσεων του δικτύου.

Αν υπάρχουν κάποια χαρακτηριστικά των νευρωνικών δικτύων που, ενδεχομένως, τα διαφοροποιούν από απλές "συμπεριφοριστικές" μηχανές αυτά έχουν να κάνουν α) με τη δυνατότητα τους να λειτουργούν κανονοποιητικά με νέα ερεθίσματα στα οποία δεν έχουν εκπαιδευτεί και β) με το ότι η έξοδός τους δεν είναι υποχρεωτικά κάποιου είδους (εξωτερική) συμπεριφορά. Για παράδειγμα, ένα νευρωνικό δίκτυο μπορεί να χρησιμοποιηθεί για την προσομοίωση ενός γνωσιακού υποσυστήματος, όπως το αντιληπτικό σύστημα- που, υπό μια θεώρηση, δεν εμπλέκει το κινητικό σύστημα.

Η πρόταση στην οποία καταλήξαμε ότι τα βάρη των συνδέσεων σε ένα νευρωνικό δίκτυο αναπαριστούν τον κανόνα μετασχηματισμού εισόδων σε εξόδους υποδεικνύει ότι το περιεχόμενο που "κατοπτρίζεται" τα νευρωνικά δίκτυα είναι μη-προτασιακό και, ακριβέστερα, μια μορφή κhow-how (βλ. κεφάλαιο 4, ενότητα 4.3). Ο εν λόγω κανόνας δεν μπορεί να διατυπωθεί προτασιακά, αλλά με όρους δεξιότητων³⁰⁶. Πρόκειται για μία γνώση που δεν μπορεί να διακριθεί από τον χρήστη της και να περιγραφεί μέσω αναπαραστάσεων που θα ίσχυαν ανεξάρτητα από τον κάτοχο του περιεχομένου. Γι' αυτό και το κhow-how είναι τόσο δύσκολο να μεταφερθεί από τον έναν άνθρωπο στον άλλο. Σε αντίθεση με την προτασιακή γνώση που μεταφέρεται με τον γραπτό ή τον προφορικό λόγο, η μη-προτασιακή γνώση μπορεί να μεταφερθεί μόνο μέσα από μια επίπονη διαδικασία ανάπτυξης κάποιας σωματικής ή πνευματικής δεξιότητας. Το κhow-how είναι ένα ουσιαδώς ενσώματο περιεχόμενο. Επίσης, δεδομένου ότι το κhow-how είναι μια γνώση που έχει να κάνει με τις ικανότητες δράσης και αντίληψης του υποκειμένου, δεν αφορά κάποιες οντότητες του εξωτερικού κόσμου, αλλά το τρόπο που το υποκείμενο σχετίζεται πρωταρχικά με το σώμα του και μέσω αυτού με τον κόσμο. Γι' αυτό, όπως επισημάναμε στο τέταρτο κεφάλαιο (ενότητα 4.3.2), το συγκεκριμένο περιεχόμενο, τελικά, δεν αναπαριστά κάτι από τον εξωτερικό κόσμο και ούτε μπορεί να αναλυθεί με όρους αναφοράς. Εφόσον, όμως το περιεχόμενο που "κατοπτρίζεται" στα νευρωνικά δίκτυα δεν διαθέτει κάποια αναφορική διάσταση, μπορούμε να συμπεράνουμε, σύμφωνα με όσα υποστηρίξαμε στο τρίτο κεφάλαιο, ότι οι επονομαζόμενες "κατανεμημένες αναπαραστάσεις" δεν είναι στην πραγματικότητα αναπαραστάσεις. Τα νευρωνικά δίκτυα προσομοιώνουν³⁰⁷ κάποιο είδος

³⁰⁶ Πρβλ. Taylor 1985, υποσημ. 2, σελ. 146.

³⁰⁷ Είναι σκόπιμο να επαναλάβουμε σε αυτό το σημείο ότι στο παρόν κεφάλαιο δεν εξετάσαμε το γενικότερο ερώτημα "δυνάμει τίνος πράγματος μια εσωτερική κατάσταση του εγκεφάλου ή του οργανισμού αποκτά περιεχόμενο". Μάλιστα, από μεθοδολογική άποψη, θεωρήσαμε προσωρινά ότι το συγκεκριμένο ερώτημα

περιεχομένου που όμως δεν αφορά τον εξωτερικό κόσμο, αλλά τη σχέση της αντίληψης με την πράξη, δηλαδή την επιδέξια δραστηριότητα (skillful activity).

6.7.5 Τι είδους περιεχόμενα αναπαρίστανται στις εισόδους και τις εξόδους ενός νευρωνικού δικτύου;

Ωστόσο, στα νευρωνικά δίκτυα κατοπτρίζεται και κάποιο είδος περιεχομένου που μπορεί να αναλυθεί με όρους αναφοράς: πρόκειται για το περιεχόμενο που κατοπτρίζεται στην είσοδο και την έξοδο των δικτύων. Όπως επισημίσαμε σε προηγούμενη παράγραφο, οι εισόδοι και οι εξόδοι σε ένα νευρωνικό δίκτυο χρησιμοποιούνται για να επικοινωνεί ο χρήστης με αυτό. Κατά συνέπεια, οι αναπαραστάσεις που χρησιμοποιούνται στα επίπεδα εισόδου και εξόδου πρέπει να είναι άμεσα ερμηνεύσιμες. Αυτό σημαίνει ότι οι εν λόγω αναπαραστάσεις θα πρέπει να είναι απομονώσιμες από το νευρωνικό δίκτυο που τις χρησιμοποιεί³⁰⁸, ούτως ώστε να μπορούν να χρησιμοποιηθούν και από τους χρήστες που επικοινωνούν με το δίκτυο. Για παράδειγμα, στο δίκτυο αναγνώρισης προσώπων που εξετάσαμε, η είσοδος ήταν εικονιστική αναπαράσταση και η έξοδος γλωσσική αναπαράσταση.

Ας δούμε όμως λίγο πιο προσεκτικά τι συμβαίνει με τις αναπαραστάσεις εισόδου και εξόδου στα νευρωνικά δίκτυα. Καταρχάς, η αναπαραστασιακή μορφή αυτών των αναπαραστάσεων είναι ο βαθμός ενεργοποίησης των κόμβων εισόδου / εξόδου. Από μόνο του το διάγραμμα των βαθμών ενεργοποίησης των κόμβων εισόδου / εξόδου δεν έχει τίποτα το γλωσσικό ή εικονιστικό. Αυτό καθορίζεται μόνο αν καθορισθεί η σχέση του διανύσματος ενεργοποίησης με την εικονιστική ή γλωσσική αναπαράσταση την οποία κωδικοποιεί. Έτσι, για παράδειγμα, στο δίκτυο αναγνώρισης προσώπων του Cottrell, υπάρχει η σύμβαση ότι ο βαθμός ενεργοποίησης κάθε κόμβου του επιπέδου εισόδου αναπαριστά κάποιο συγκεκριμένο εννοιολογικό της φωτογραφίας. Αντίστοιχα, για την κωδικοποίηση μιας γλωσσικής αναπαράστασης ορίζεται κάποια σαφής σχέση αντιστοίχησης είτε των λέξεων, είτε των γραμμάτων, είτε

απαντώντα ικανοποιητικά. Με άλλα λόγια προϋποθέσαμε ότι τα νευρωνικά δίκτυα έχουν όντως κάποιο είδος περιεχομένου και εξετάσαμε από κει και πέρα σε τι διαφοροποιείται αυτό από το περιεχόμενο των συμβολικών αναπαραστάσεων. Στο επόμενο κεφάλαιο θα θέσουμε υπό αμφισβήτηση αυτή την παραδοχή και τελικά θα καταλήξουμε στον Επίλογο ότι όλα τα υποτιθέμενα είδη αναπαραστάσεων που υιοθετεί η γνωσιακή επιστήμη αποτελούν απλώς διαφορετικά είδη εγγραφών (βλ. κεφάλαιο 3), δηλαδή αποτελούν προσομοιώσεις διαφορετικών ειδών περιεχομένου. Προαναγγέλλοντας αυτήν την κριτική χρησιμοποιούμε εδώ την έκφραση "προσομοιώνουν κάποιο είδος περιεχομένου" αντί της έκφρασης "έχουν κάποιο είδος περιεχομένου".

³⁰⁸ Άρα λουτόν δεν πρόκειται για κάποια μορφή know-how, σύμφωνα με ό,τι αναφέραμε στην προηγούμενη ενότητα.

οποιοδήποτε άλλο χαρακτηριστικού γνωρίσματος της γλωσσικής αναπαράστασης με τους βαθμούς ενεργοποίησης των κόμβων. Μόνο αν γνωρίζουμε αυτή την αντιστοίχιση μπορούμε να πούμε ότι ο βαθμός ενεργοποίησης των κόμβων εισόδου / εξόδου αναπαριστά κάποιο εικονιστικό ή συμβολικό περιεχόμενο. Έτσι, ο βαθμός ενεργοποίησης των κόμβων εισόδου / εξόδου δεν αποτελεί κάποιο αυτόνομο είδος αναπαράστασης με την έννοια ότι μπορεί να κατανοηθεί ανεξάρτητα από κάποια εικονιστική ή γλωσσική αναπαράσταση. Ούτε κατά συνέπεια, μπορεί να εκφράσει κάποιου είδους περιεχόμενο που δεν είναι εικονιστικό ή συμβολικό. Γι' αυτό είναι ορθότερο να κατανοήσουμε τις αναπαραστάσεις των επιπέδων εισόδου και εξόδου ως μια μορφή κώδικα. Οι κωδικοποιημένες αναπαραστάσεις προκύπτουν πάντα μέσω κάποιου συστηματικού μετασχηματισμού των *μορφικών* χαρακτηριστικών γλωσσικών ή εικονιστικών αναπαραστάσεων³⁰⁹. Δεν πρόκειται για κάποια παρέμβαση που έχει ως στόχο την μεταβολή του περιεχομένου των αναπαραστάσεων, αλλά την απόκρυψή του. Για παράδειγμα, ένας απλός τρόπος κρυπτογράφησης γλωσσικών αναπαραστάσεων είναι η αντικατάσταση των γραμμάτων με άλλα γράμματα ή σύμβολα σύμφωνα με κάποιον κανόνα αντιστοίχισης³¹⁰. Γνωρίζοντας την μέθοδο κρυπτογράφησης μπορούμε να παρέμβουμε στον κώδικα και να τον αποκρυπτογραφήσουμε προκειμένου να ανακτήσουμε την αρχική αναπαράσταση με αναλλοίωτο το περιεχόμενό της. Με την ορολογία του κεφαλαίου 3, η κωδικοποίηση αποτελεί μια ειδική περίπτωση μετεγγραφής.

Αυτό που πρέπει να κρατήσουμε ως σημαντικό από το παράδειγμα των κωδικοποιημένων αναπαραστάσεων είναι ότι 1) οι κανόνες μετασχηματισμού που εφαρμόζονται τόσο κατά τη διαδικασία κρυπτογράφησης, όσο και κατά τη διαδικασία απόκρυπτογράφησης δεν μεταβάλλουν το περιεχόμενο των αναπαραστάσεων - μόνο τη μορφή τους και 2) ότι οι κωδικοποιημένες αναπαραστάσεις δεν αποτελούν κάποιο αυτόνομο είδος αναπαράστασης, αλλά υπάρχουν μόνο δευτερογενώς ("παρασιτικά", θα λέγαμε) από τις αναπαραστάσεις τις οποίες κωδικοποιούν.

Γι' αυτό οι αναπαραστάσεις εισόδου και εξόδου στα νευρωνικά δίκτυα θα πρέπει να νοηθούν ως κωδικοποιημένες αναπαραστάσεις -με την παραπάνω έννοια- γλωσσικών ή εικονιστικών αναπαραστάσεων.

Μία από τις μεθόδους κωδικοποίησης στους κόμβους εισόδου εικονιστικών και, γενικότερα, αναλογικών περιεχομένων είναι αυτή που παράγει αναπαραστάσεις με ισοδυναμικά χαρακτηριστικά, δηλαδή κάθε

³⁰⁹ Βλ. Nichols (1995).

³¹⁰ Μια γνωστή στο χώρο της φιλοσοφίας μέθοδος κρυπτογράφησης μπορεί να θεωρηθεί η αρίθμηση Gödel η οποία είναι ένας γενικός τρόπος για να αποδίδουμε σε κάθε έκφραση του προτασιακού λογισμού ένα διαφορετικό φυσικό αριθμό.

τιμήμα μιας τέτοιας αναπαράστασης διαθέτει πληροφορίες για το σύνολο του αναπαριστώμενου περιεχομένου. Υπάρχουν πολλές τεχνικές που οδηγούν σε αυτό το αποτέλεσμα. Η γνωστότερη ίσως από αυτές είναι ο μετασχηματισμός Fourier³¹¹. Κατά συνέπεια, η ισοδυναμικότητα που παρουσιάζουν οι ονομαζόμενες "αναπαραστάσεις εισόδου" στα νευρωνικά δίκτυα δεν θα πρέπει να θεωρηθεί ως ένα χαρακτηριστικό που προσδιορίζει το είδος του περιεχομένου που αναπαριστούν, αλλά το είδος της κωδικοποίησης που τους έχει γίνει. Οι αναπαραστάσεις εισόδου και εξόδου στα νευρωνικά δίκτυα, στο βαθμό που αποτελούν απλώς κωδικοποιημένες συμβολικές ή εικονιστικές αναπαραστάσεις παραλαμβάνουν όλα τα προβλήματα που προαναφέραμε για τις συμβολικές και τις εικονιστικές αναπαραστάσεις στα κεφάλαια 4 και 5.

6.8 Επίλογος

Με το κεφάλαιο αυτό ολοκληρώνουμε τη διερεύνηση των διαφόρων αναπαραστασιακών μορφών που χρησιμοποιούνται στην γνωσιακή επιστήμη. Είδαμε, ότι, όπως η κλασική γνωσιακή επιστήμη επιχείρησε να χειριστεί τις σημασιολογικές σχέσεις μεταξύ προτασιακών περιεχομένων κατοπτρίζοντάς τις στη σύνταξη των συμβολικών αναπαραστάσεων, έτσι και οι υποστηρικτές των νευρωνικών δικτύων μπορούν να θεωρηθούν ότι μελετούν μη προτασιακά περιεχόμενα, δηλαδή περιεχόμενα που αφορούν τη πρακτική γνώση δεξιοτήτων, κατοπτρίζοντάς τα στα βάρη των συνδέσεων μεταξύ των κόμβων του δικτύου. Ωστόσο, από τη διερεύνηση που κάναμε αποδείχθηκε ότι το ερώτημα για το τι ακριβώς είναι μια κατανεμημένη αναπαράσταση παραμένει ακόμα εν πολλοίς αναπάντητο. Η πρόταση στην οποία καταλήξαμε είναι ότι αν κανείς δεχθεί την αφετηριακή παραδοχή των υποστηρικτών της αναπαραστασιακής θεωρίας του νου ότι τα διαφορετικά είδη νοητικής αναπαράστασης διαφοροποιούνται από τον τρόπο κατοπτρισμού του αναπαραστασιακού περιεχομένου στην αναπαραστασιακή μορφή, τότε α) στα νευρωνικά δίκτυα η μορφή στην οποία "κατοπτρίζεται" το περιεχόμενο είναι τα βάρη των συνδέσεων και όχι ο βαθμός ενεργοποίησης των κόμβων, β) το "κατοπτριζόμενο" περιεχόμενο δεν είναι προτασιακό αλλά κάποιο είδος πρακτικής γνώσης δεξιοτήτων και γ) οι αναπαραστάσεις εισόδου και εξόδου των νευρωνικών δικτύων δεν

³¹¹ Η μαθηματική διατύπωση του μετασχηματισμού είναι η εξής:

$$g(u) = \int f(x)e^{-i2\pi ux} dx$$
 Η τιμή της συνάρτησης g σε οποιοδήποτε σημείο της μεταβάλλεται συναρτήσει της τιμής της συνάρτησης f σε κάθε σημείο της. Αντικαθιστώντας τη συνάρτηση $e^{-i2\pi ux}$ με άλλου τύπου συναρτήσεις παίρνουμε διαφορετικά είδη μετασχηματισμών που ικανοποιούν την ιδιότητα της ισοδυναμικότητας (βλ. van Gelder 1992 και Καραγιάννης 1989).

αποτελούν κατανεμημένες αναπαραστάσεις αλλά κωδικοποιημένες συμβολικές ή εικονιστικές αναπαραστάσεις.

ΚΕΦΑΛΑΙΟ ΕΒΔΟΜΟ

Κριτική των προγραμμάτων φυσικοποίησης του περιεχομένου των νοητικών αναπαραστάσεων

7.1 Εισαγωγή

Με το προηγούμενο κεφάλαιο ολοκληρώσαμε την εξέταση των διαφορετικών μορφών αναπαράστασης που χρησιμοποιούνται στο πλαίσιο της γνωσιακής επιστήμης. Ένα κοινό στοιχείο που προέκυψε είναι ότι τόσο οι υποστηρικτές των νοητικών ειδώλων όσο και οι υποστηρικτές των κατανεμημένων αναπαραστάσεων δεν εντοπίζουν τη διαφορά των εν λόγω αναπαραστάσεων από τις συμβολικές στο περιεχόμενό τους, αλλά στη σχέση του περιεχομένου με την αναπαραστασιακή μορφή, δηλαδή στον τρόπο που το αναπαραστασιακό περιεχόμενο κατοπτρίζεται στην αναπαραστασιακή μορφή. Έτσι, η διερεύνηση των διαφορών μεταξύ ειδών περιεχομένου περνά σε δεύτερη μοίρα και αντικαθίσταται από την μελέτη των διαφορών μεταξύ αναπαραστασιακών μορφών. Αυτή δε η παραδοχή οδηγεί τις περισσότερες φορές σε μια ακόμα πιο αμφισβητούμενη παραδοχή, ότι, δηλαδή, τα προς αναπαράσταση περιεχόμενα είναι του ίδιου τύπου, τουτέστιν προτασιακά (ή συστατικά αυτών, όπως αντικείμενα και ιδιότητες), και ότι οι διαφορές μεταξύ των τριών ειδών αναπαράστασης που χρησιμοποιούνται στο πλαίσιο της γνωσιακής επιστήμης έγκεινται απλά στο διαφορετικό τρόπο κατοπτρισμού αυτών των περιεχομένων στις αναπαραστασιακές μορφές.

Στο τέλος του κεφαλαίου 6 προτείνουμε μια διαφορετική ερμηνεία όσον αφορά τις αποκαλούμενες "κατανεμημένες αναπαραστάσεις", σύμφωνα με την οποία δεν πρόκειται καν για αναπαραστάσεις, δεδομένου ότι το περιεχόμενο που κατοπτρίζεται στους κόμβους του δικτύου, ήτοι η πρακτική γνώση δεξιοτήτων, δεν έχει αναφορική διάσταση. Επίσης, στο πέμπτο κεφάλαιο επισημάνουμε ότι αν το περιεχόμενο των εικονιστικών αναπαραστάσεων είναι προτασιακό, τότε οι συμβολικές αναπαραστάσεις είναι σίγουρα πολύ ισχυρότερες εκφραστικά από τις εικονιστικές. Γι' αυτό, προκειμένου να εξηγηθεί η εκφραστική δύναμη των εικονιστικών αναπαραστάσεων αντιπροτείνουμε ότι θα πρέπει να κατανοηθούν ως αναπαραστάσεις ενός είδους μη-προτασιακού περιεχομένου ή, τουλάχιστον, όχι αποκλειστικά προτασιακού περιεχομένου³¹². Ωστόσο, δεν προχωρήσαμε στην επεξεργασία αυτής της πρότασης γιατί μια τέτοια ευθύνη βαρύνει τους υποστηρικτές της αναπαραστασιακής θεωρίας του νου.

Ανεξάρτητα από αυτές τις προτάσεις όσον αφορά τις εικονιστικές αναπαραστάσεις και τις λεγόμενες "κατανεμημένες αναπαραστάσεις", αν κανείς παραμείνει στην κυρίαρχη αντίληψη που ανάγει τις διαφορές από τις συμβολικές αναπαραστάσεις στον διαφορετικό τρόπο κατοπτρισμού του αναπαραστασιακού περιεχομένου στην αναπαραστασιακή μορφή, τότε εξακολουθεί να εκκρεμεί ένα σημαντικότατο πρόβλημα. Η ιδέα του

³¹² Βλ., για παράδειγμα, την έννοια του αναλογικού περιεχομένου που εισήγαγε ο Peacocke 1986.

κατοπτρισμού (mirroring) πάνω στην οποία στηρίζεται η ΑΥΘΝ³¹³, αλλά και οι υπόλοιπες αναπαραστασιακές θεωρίες, δεν μπορεί να επιλύσει συνολικά το πρόβλημα του νοητικού περιεχομένου. Αν θυμηθούμε το παράδειγμα από την προτασιακή λογική που αναφέραμε στο πρώτο κεφάλαιο, αυτό που κατοπτρίζεται από το σημασιολογικό επίπεδο στο συντακτικό επίπεδο είναι το περιεχόμενο των λογικών συνδέσμων. Με άλλα λόγια, κατοπτρίζονται οι σημασιολογικές σχέσεις μεταξύ των προτάσεων - με κανένα τρόπο, όμως, δεν κατοπτρίζεται το περιεχόμενο των ιδίων των προτάσεων! Ο Fodor τα τελευταία χρόνια φαίνεται ορισμένες φορές να συνειδητοποιεί το συγκεκριμένο σημείο: "...η σύνταξη ενός τύπου κωδικοποιεί τη σημασία του" ειδικότερα εκείνη την πλευρά της σημασίας του που καθορίζει τον ρόλο του στην συναγωγή και στις σχέσεις μεταξύ σημασιών" (Fodor & Pylyshyn 1988, σελ. 303). Οι σημασιολογικές σχέσεις αποτελούν ουσιαστικά ένα άλλο όνομα για τους σημασιολογικούς κανόνες. Κατά συνέπεια, όλη η ιδέα περί κατοπτρισμού αφορά τον κατοπτρισμό των σημασιολογικών κανόνων στους συντακτικούς κανόνες (οι οποίοι με τη σειρά τους κατοπτρίζονται σε αιτιακούς νόμους).

Η ιδέα, λοιπόν, του κατοπτρισμού δεν απαντά στο πρόβλημα της αποβλεπτικότητας των νοητικών αναπαραστάσεων αλλά, τουλάχιστον όσον αφορά την ΑΥΘΝ, στο πρόβλημα της *ορθολογικότητας των νοητικών διαδικασιών*, ή, όπως λέει ο Fodor, στο "πώς η ορθολογικότητα είναι μηχανικά δυνατή" (Fodor 1987, σελ. 20). Συγκεκριμένα, απαντά στο ακόλουθο ερώτημα: δεδομένου ότι οι καταστάσεις ενός φυσικού συστήματος έχουν σημασιολογικές ιδιότητες, με ποιο τρόπο μεταβίνει το σύστημα από τη μια κατάσταση στην άλλη έτσι ώστε να διατηρείται η αλήθεια;³¹⁴ Συνεπώς, το πρόβλημα της αποβλεπτικότητας των νοητικών αναπαραστάσεων, δηλαδή το πρόβλημα του πώς μια υλική κατάσταση μπορεί να έχει σημασιολογικές ιδιότητες παραμένει.

Σε αυτό το κεφάλαιο θα εξετάσουμε το πρόβλημα της αποβλεπτικότητας και τις απαντήσεις που έχουν προταθεί κυρίως σε σχέση με την ΑΥΘΝ. Θα κάνουμε επίσης μια σύντομη αναφορά στους λόγους που η θεωρία νοητικών ειδώλων αδυνατεί να επιλύσει το πρόβλημα της αποβλεπτικότητας. Όσον αφορά, τέλος, τις υποτιθέμενες κατανεμημένες "αναπαραστάσεις" και τον τρόπο που αυτές αποκτούν περιεχόμενο, ελλείψει οποιασδήποτε επεξεργασμένης πρότασης, δεν θα προβούμε σε κάποια ιδιαίτερη διερεύνηση. Θα επισημάνουμε απλώς εκ των προτέρων ότι, δεδομένου ότι οι κατανεμημένες "αναπαραστάσεις" συνδέουν

³¹³ Βλ. ενότητες 1.3.4 και 2.4.2.

³¹⁴ Ένα αποτέλεσμα καθόλου ευκαταφρόνητο ωστόσο, αν σκεφτεί κανείς πόσο αποτυχημένη ήταν η μόνη εναλλακτική λύση, αυτή του Βρετανικού Εμπειρισμού, που προσπάθησε με τους νόμους του συνειρμού (αιτιότητα -γεννίαση) να εξηγήσει την διατήρηση της αλήθειας.

εισόδους με εξόδους οι οποίες στα νευρωνικά δίκτυα είναι, όπως είδαμε, κωδικοποιημένες εικονιστικές ή συμβολικές αναπαραστάσεις, οποιαδήποτε προβλήματα προκύψουν σχετικά με την απόκτηση περιεχομένου όσον αφορά τις συμβολικές και τις εικονιστικές αναπαραστάσεις κληρονομούνται και στις κατανεμημένες "αναπαραστάσεις". Τέλος, στον Επίλογο θα επανέλθουμε με ένα πρόσθετο επιχείρημα κατά της θέσης ότι τα νευρωνικά δίκτυα μπορούν να αποκτήσουν περιεχόμενο –έστω και αν αυτό θεωρηθεί ότι είναι μη προτασιακό και, συγκεκριμένα, κάποιο είδος πρακτικής γνώσης δεξιοτήτων, όπως υποστηρίξαμε στο τέλος του κεφαλαίου 6. Προαναγγέλλοντας αυτό το επιχείρημα –το οποίο για να αναπνυχθεί συστηματικά χρειάζονται τα αποτελέσματα του κεφαλαίου 8 – μπορούμε να πούμε ότι θα βασιστεί στη θέση ότι η επιδέξια δραστηριότητα (skillful activity) ανήκει πάντα σε κάποιον γι' αυτό και διακρίνεται από τα εξαρτημένα αναπνευστικά (βλ. κεφάλαιο 4). Η συναίσθηση της ικανότητας του πράττειν (agency) αποτελεί συγκροτητικό χαρακτηριστικό της επιδέξιας δραστηριότητας. Στα νευρωνικά δίκτυα, ωστόσο, δεν υπάρχει κανένας τρόπος μέχρι στιγμής να συλληφθεί αυτό το χαρακτηριστικό. Κατά συνέπεια, τα νευρωνικά δίκτυα δεν μπορούν να συλλάβουν το φαινόμενο της επιδέξιας δραστηριότητας στην ολότητα του.

7.2 *Ψυχοσημασιολογία: Το πρόβλημα της προέλευσης του περιεχομένου των αναπαραστάσεων*

Ο Fodor όπως και αρκετοί ερευνητές μαζί μ' αυτόν φτάνουν σταδιακά στην πλήρη συνειδητοποίηση του ότι η αποβλεπτικότητα δεν μπορεί να εξηγηθεί μόνο με βάση την ορθολογικότητα³¹⁵. Στις αρχές της δεκαετίας του '90 ο Fodor θέτει πλέον ξεκάθαρα το ζήτημα:

"Υπάρχουν, πιστεύω, τρία μεγάλα μεταφυσικά αινίγματα όσον αφορά το νου. Πώς μπορεί κάτι υλικό να έχει συνειδητές καταστάσεις; Πώς μπορεί κάτι υλικό να έχει σημασιολογικές ιδιότητες; Πώς μπορεί κάτι υλικό να είναι ορθολογικό; (όπου αυτό σημαίνει περίπου το εξής: πώς μπορούν οι μεταβάσεις ενός φυσικού συστήματος από τη μια κατάσταση στη άλλη να διατηρούν σημασιολογικές ιδιότητες;) Η ιστορία με τον παραλληλισμό [εννοείται, σύνταξης-σημασιολογίας] - εάν όντως δουλέψει - απαντά στο τρίτο ερώτημα" (Loewer & Rey 1991, σελ. 285)

³¹⁵ Αντίθετα κάποιοι άλλοι φιλόσοφοι, όπως ο Dennett, εξακολουθούν να πιστεύουν ότι η αποβλεπτικότητα μπορεί να εξηγηθεί με βάση την ορθολογικότητα. Όπως λέει χαρακτηριστικά ο Dennett (1978b, σελ. 19) η ορθολογικότητα είναι "η μητέρα της πρόθεσης (intention)". Βλ., επίσης, Sellars 1963a, Brandom 1994 οι οποίοι επιχειρούν με βάσει την έννοια της συναγωγής να εξηγήσουν το περιεχόμενο των αναπαραστάσεων.

Το πρόβλημα της αποβλεπτικότητας θεωρείται πλέον το κεντρικό πρόβλημα της γνωσιακής επιστήμης. Σε σχέση δε με το πρώτο ερώτημα - το πρόβλημα της συνείδησης- τόσο ο Fodor όσο και γενικότερα οι ερευνητές στον χώρο της κλασσικής γνωσιακής επιστήμης δεν έχουν σημειώσει την παραμικρή πρόοδο. Θα επανέλθουμε σε αυτό το ζήτημα στο όγδοο κεφάλαιο.

Με τους όρους της θεωρίας του Fodor, το πρόβλημα της αποβλεπτικότητας ισοδυναμεί με το εξής ερώτημα: Με ποιο τρόπο μια πρόταση της γλώσσας της σκέψης μπορεί να αναφέρεται σε μια συγκεκριμένη κατάσταση πραγμάτων; Αν δεχθεί κανείς τη θέση του Fodor ότι η γλώσσα της σκέψης έχει συνθετική σημασιολογία, τότε το ερώτημα ανάγεται στο πώς οι επιμέρους "λέξεις" αυτής της γλώσσας αποκτούν το περιεχόμενό τους. Δηλαδή, το ζητούμενο είναι πώς μια φυσική κατάσταση εγγενώς αναφέρεται σε μια άλλη φυσική κατάσταση - πώς μια φυσική κατάσταση καθίσταται εγγενώς σύμβολο/αναπαράσταση μιας άλλης φυσικής κατάστασης. "Εγγενώς", εδώ, σημαίνει ότι το σύμβολο παραπέμπει από μόνο του στο αναπαριστώμενο χωρίς να χρειάζεται κάποιον εξωτερικό ερμηνευτή-χρήστη³¹⁶ που θα του αποδώσει αυτή την λειτουργία - όπως, ενδεχομένως, χρειάζεται για τις αναπαραστάσεις της φυσικής γλώσσας ή της γλώσσας μηχανής του υπολογιστή³¹⁷. Κάτι τέτοιο είναι αναμενόμενο γιατί διαφορετικά οι αναπαραστάσεις θα δανείζονταν την αναφορική τους λειτουργία από τον ερμηνευτή-χρήστη, πράγμα που θα απαγόρευε τη δυνατότητα φυσιοκρατικής εξήγησης. Πιο συγκεκριμένα, σε μια τέτοια περίπτωση, θα χρησιμοποιούνταν οι προτασιακές στάσεις του ερμηνευτή-χρήστη για να εξηγηθεί η λειτουργία των νοητικών αναπαραστάσεων τη στιγμή που οι νοητικές αναπαραστάσεις έχουν εισαχθεί ακριβώς για την εξήγηση των προτασιακών στάσεων.

Από τις αρχές του '80 άρχισε να τίθεται επιτακτικά το αίτημα να επιλυθεί το πρόβλημα του περιεχομένου των νοητικών αναπαραστάσεων. Σημαντικό ρόλο σε αυτή την τροπή έπαιξε, ως ένα βαθμό, και το επιχείρημα του Searle (1980) για το κινέζικο δωμάτιο. Το επιχείρημα αυτό, μεταξύ άλλων, στρέφεται εναντίον της ΑΥΘΝ και καταλήγει σε μια πολύ ισχυρή θέση: με την αναγωγή (ή, εν προκειμένω, τον κατοπινισμό) των νοητικών καταστάσεων σε λειτουργικές - συντακτικές καταστάσεις δεν διαφεύγει απλάς κάποιο συστατικό του περιεχομένου τους αλλά το περιεχόμενο στο σύνολό του. Το επιχείρημα στηρίζεται σε μια ιδιοφυή αντιστροφή: αντί της προσομοίωσης της ανθρώπινης νόησης από τον υπολογιστή επιχειρείται, αντίθετα, η μίμηση των λειτουργιών του υπολογιστή από τον άνθρωπο. Πιο συγκεκριμένα, το επιχείρημα έχει ως εξής: ένας άνθρωπος αποκλεισμένος σε ένα δωμάτιο επικοινωνεί με τον

³¹⁶ Βλ. κεφάλαιο 3, ενότητα 3.3.1.4.

³¹⁷ Το ζήτημα όσον αφορά τη γλώσσα μηχανής του υπολογιστή γίνεται αποδεκτό και από τον ίδιο τον Fodor (1990b).

εξωτερικό κόσμο μέσω γραπτών μηνυμάτων. Τα μηνύματα που δέχεται είναι γραμμένα στα κινέζικα, το ίδιο και τα μηνύματα που στέλνει, χωρίς ωστόσο ο ίδιος να γνωρίζει κινέζικα. Αυτό το επιτυγχάνει χάρης σε ένα εγχειρίδιο-ευρετήριο στο οποίο ακολουθίες κινέζικων συμβόλων αντιστοιχίζονται με άλλες ακολουθίες κινέζικων συμβόλων. Ο άνθρωπος για να απαντήσει σε ένα μήνυμα βρίσκει στο ευρετήριο τη συγκεκριμένη ακολουθία συμβόλων και αντιγράφει σε ένα χαρτί την ακολουθία στην οποία αυτή αντιστοιχίζεται. Από τη σκοπιά ενός εξωτερικού παρατηρητή που γνωρίζει κινέζικα, τα μηνύματα που στέλνονται στο δωμάτιο είναι ερωτήματα, τα δε μηνύματα που εξέρχονται από αυτό είναι απαντήσεις στα συγκεκριμένα ερωτήματα. Στον εξωτερικό παρατηρητή, λοιπόν, δίδεται η εντύπωση ότι ο άνθρωπος μέσα στο δωμάτιο καταλαβαίνει κινέζικα, μια εντύπωση που είναι προφανώς ψευδής από την οπτική του ανθρώπου που βρίσκεται στο δωμάτιο. Συνεπώς, ο λειτουργισμός με τη αναγωγή των νοητικών καταστάσεων στη συντακτική επεξεργασία συμβόλων αποτυγχάνει πλήρως να συλλάβει το σημαντικότερο, ίσως, χαρακτηριστικό των νοητικών καταστάσεων, το περιεχόμενο - δηλαδή, την αποβλεπτικότητα τους. Το περιεχόμενο δεν μπορεί να συγκροτηθεί από συντακτικές³¹⁸ σχέσεις.

Ο Fodor αποδέχεται ως ένα βαθμό³¹⁹ αυτή την κριτική γιατί, όπως και ο ίδιος διατείνεται, η ΑΥΘΝ επιχειρεί να εξηγήσει τις νοητικές διαδικασίες προσφεύγοντας αποκλειστικά στις μη σημασιολογικές ιδιότητες των νοητικών αναπαραστάσεων. Σε κανένα σημείο της η ΑΥΘΝ δεν εξηγεί το πώς οι νοητικές αναπαραστάσεις αποκτούν σημασιολογικές ιδιότητες. Ο Fodor πιστεύει μάλιστα ότι η έλλειψη απάντησης σε αυτό το θεμελιακό ζήτημα δεν επηρεάζει την έρευνα που γίνεται στο πλαίσιο της γνωσιακής επιστήμης. Είναι χαρακτηριστικό ότι και ο ίδιος, κατά τα πρώτα χρόνια της ανάπτυξης της θεωρίας του, ενδιαφερόταν κυρίως για τους ψυχολογικούς νόμους και τις αρχές που διέπουν την συνθετική σημασιολογία της γλώσσας της σκέψης. Έτσι, έδινε ιδιαίτερη έμφαση στο πώς από τα ατομικά σύμβολα κατασκευάζονται σύνθετες νοητικές αναπαραστάσεις. Ωστόσο, πουθενά στις έρευνες αυτής της περιόδου δεν απαντά στο ερώτημα για την προέλευση του περιεχομένου, σε σημείο που ο Pylyshyn - στενός συνεργάτης του- να λέει μεταξύ σοβαρού και αστείου ότι αυτό που πιστεύει ο Fodor είναι ότι "it's syntax - all the way down!"³²⁰.

³¹⁸ Ο Searle στο μεταγενέστερο βιβλίο του *The Rediscovery of the Mind* σκληραίνει ακόμα περισσότερο τη θέση του υποστηρίζοντας ότι οι υπολογιστές δεν κάνουν ούτε καν συντακτική επεξεργασία. Η σύνταξη δεν είναι μια απλή αιτιακή κανονικότητα που μπορεί να υπάρχει στη φύση ανεξάρτητα από το νοούν υποκείμενο.

³¹⁹ Μια προφανής διαφοροποίησή του από τον Searle είναι ότι δεν αποδέχεται την οπτική του πρώτου προσώπου ως τρόπο προσέγγισης του περιεχομένου.

³²⁰ Αναφέρεται από τον Dennett (1991b, σελ. 90) σε ένα απολογιστικό κείμενό του για τον Fodor.

Την ίδια περίοδο, δηλαδή από τα μέσα δεκαετίας του '70 έως τις αρχές της δεκαετίας του '80, παρατηρείται μέσα στους κόλπους της γνωστικής επιστήμης μια σύγχυση όσον αφορά τους στόχους και τα επιτεύγματά της. Δεν είναι λίγοι οι ερευνητές που θεωρούν ότι με το έργο τους απαντούν και στο θεμελιακό ζήτημα του πώς αποκτούν οι νοητικές αναπαραστάσεις το περιεχόμενό τους. Όπως χαρακτηριστικά επισημαίνει ο Fodor "... η κατάσταση ήταν αρκετά ασαφής εξαιτίας τού ότι πολλοί ερευνητές δεν κατάφεραν να διακρίνουν τα ζητήματα λογικής σύνταξης από τα αυστηρώς εννοούμενα σημασιολογικά ζητήματα. Συνέπεια αυτής της σύγχυσης ήταν ότι πολλοί γνωστικοί επιστήμονες θεωρούσαν ότι έκαναν σημασιολογία ενώ στην πραγματικότητα δεν έκαναν" (Fodor 1990b, σελ. 315).

Ο ίδιος ο Fodor άρχισε να εργάζεται πάνω στο ζήτημα τού πώς τα ατομικά σύμβολα αποκτούν το περιεχόμενό τους στις αρχές του '80. Από τότε έχει μεταβάλλει αρκετές φορές τις απόψεις του επηρεάζοντας ταυτόχρονα σε μεγάλο βαθμό τις θέσεις των υπολοίπων ερευνητών (βλ. Loewer & Rey 1991, σελ. 299-304).

Ας δούμε όμως τώρα σε τι ακριβώς συνίσταται το πρόβλημα του περιεχομένου των νοητικών αναπαραστάσεων ή, με την ορολογία του Fodor (1987), το πρόβλημα της "ψυχοσημασιολογίας". Το ζητούμενο για την ψυχοσημασιολογία δεν είναι να διατυπωθεί μια θεωρία σημασιολογίας που να καθορίζει ποιο είναι το περιεχόμενο (κατανοημένο ως οι συνθήκες αλήθειας) των διάφορων όρων, όπως, για παράδειγμα, φιλοδοξεί να κάνει η σημασιολογία συνθηκών αλήθειας (truth-conditional semantics). Το ζητούμενο είναι να καθορισθεί το τι θεμελιώνει τη σχέση του συμβόλου με το συμβολιζόμενο. Με άλλα λόγια, το ερώτημα στο οποίο καλείται να απαντήσει η ψυχοσημασιολογία είναι: δυνάμει τίνος πράγματος μια εσωτερική κατάσταση αποκτά αναπαραστασιακό περιεχόμενο (βλ. Fodor 1990b, σελ. 318)³²¹. Μια πρώτη χαρτογράφηση και κριτική αξιολόγηση των υπαρχουσών απαντήσεων σε αυτό το ερώτημα έχει ήδη επιχειρηθεί στο τρίτο κεφάλαιο (βλ. ενότητες 3.3.1.3 και 3.3.1.4).

Ο Fodor δεν αρνείται τη χρησιμότητα μιας σημασιολογίας τύπου συνθηκών αλήθειας, απλώς επισημαίνει ότι αυτή απαντά σε διαφορετικό πρόβλημα. Το σημείο αυτό είναι αρκετά λεπτό και έχει προκαλέσει μεγάλη σύγχυση γι' αυτό θα επιμείνουμε κάπως παραπάνω για τη διασαφήνισή του. Στην φιλοσοφία του νου όταν κανείς μιλά γενικώς για θεωρίες σημασιολογίας νοητικών καταστάσεων, μπορεί να αναφέρεται σε θεωρίες που ανήκουν σε δύο εντελώς διαφορετικές κατηγορίες. Στη πρώτη κατηγορία ανήκουν θεωρίες που έχουν ως στόχο να περιγράψουν το περιεχόμενο των νοητικών καταστάσεων. Οι θεωρίες αυτές απλώς

³²¹ Βλ. επίσης Fodor & Lepore (1994) όπου διατυπώνεται το ερώτημα με δύο ακόμα τρόπους: "από που προέρχονται οι σημασιολογικές ιδιότητες;" (σελ. 142) και "σε τι επιγίνονται οι σημασιολογικές ιδιότητες των συμβόλων" (σελ. 143).

προσδιορίζουν ποιο είναι το περιεχόμενο των διαφόρων νοητικών καταστάσεων, χωρίς να ενδιαφέρονται να εξηγήσουν το σε τι οφείλεται το εκάστοτε περιγραφόμενο περιεχόμενο, δηλαδή από πού προέρχεται. Για παράδειγμα, τέτοιου τύπου θεωρίες όσον αφορά τη φυσική γλώσσα είναι οι διάφορες τυπικές σημασιολογίες όπως η φρεγκεανή σημασιολογία, η ταρσκιανή σημασιολογία και η σημασιολογία δυνατών κόσμων. Στη δεύτερη κατηγορία ανήκουν θεωρίες που έχουν ως στόχο να προσδιορίσουν το τι θεμελιώνει το περιεχόμενο των νοητικών αναπαραστάσεων. Με άλλα λόγια, το ερώτημα στο οποίο απαντούν είναι: δεδομένου του είδους ενός περιεχομένου -όπως αυτό έχει περιγραφεί από τις θεωρίες της πρώτης κατηγορίας - πώς μια νοητική κατάσταση αποκτά το εν λόγω περιεχόμενο³²²; Στην δεύτερη αυτή κατηγορία εντάσσονται, μεταξύ άλλων, η πληροφοριακή σημασιολογία (informational semantics) (Dretske 1981), η θεωρία της ασύμμετρης εξάρτησης (theory of asymmetric dependence) (Fodor 1987), η τελεοσημασιολογία (teleosemantics) (Millikan 1984), η σημασιολογία εννοιολογικού ρόλου (conceptual role semantics) (Block 1986) και η θεωρία της προθετικής στάσης (Dennett 1987). Όλες αυτές οι θεωρίες θα εξεταστούν στη συνέχεια του κεφαλαίου.

Για την αποφυγή σύγχυσης, από εδώ και στο εξής θα αναφερόμαστε στην πρώτη κατηγορία θεωριών με τον όρο "σημασιολογία συνθηκών αλήθειας" και στην δεύτερη με τον όρο του Fodor "ψυχοσημασιολογία"³²³.

Όπως επιστημαίνει ο Stalnaker "οι σημασιολογικές θεωρίες συνθηκών αλήθειας ... λένε ποιες είναι οι σημασιολογικές τιμές εκφράσεων διαφορετικών κατηγοριών και πώς οι σημασιολογικές τιμές σύνθετων εκφράσεων καθορίζονται από τις τιμές των μερών τους. ... Αλλά δηλώσεις τέτοιου είδους δεν λένε απολύτως τίποτα για το ποια χαρακτηριστικά των νοητικών καταστάσεων και της συμπεριφοράς των χρηστών αυτών των εκφράσεων καθιστούν αληθές το ότι οι λέξεις έχουν τις εκτάσεις (ή τις εντάσεις) που έχουν ... αυτό είναι το είδος των ερωτημάτων που η σημασιολογία εννοιολογικού ρόλου επιχειρεί να απαντήσει." (Loewer & Rey 1991, σελ. 233). Ο Fodor (Loewer & Rey 1991, σελ. 299) αποδέχεται πλήρως την διάκριση του Stalnaker εκτός από το τελευταίο τμήμα της, όπου αυτός, αντί της σημασιολογίας

³²² Ανοικτό παραμένει το ζήτημα για το εάν ένα τέτοιο ερώτημα αποτελεί φιλοσοφικό ερώτημα και όχι απλώς ένα επιστημονικό ερώτημα. Βλ. Cussins 1987, σελ. 126.

³²³ Βλέπε σχετικά με αυτή τη διάκριση von Eckardt 1993, σελ. 189-190 και Cummins 1989, σελ. 11. Επίσης ο Block (υπό έκδοση) αναφέρεται στις δύο κατηγορίες σημασιολογικών θεωριών αντίστοιχα με τους όρους "γλωσσολογική σημασιολογία" και "μεταφυσική σημασιολογία".

εννοιολογικού ρόλου, προτείνει μια άλλη λύση στο πρόβλημα της ψυχοσημασιολογίας - τη θεωρία της ασύμμετρης εξάρτησης.

Το πρόβλημα της ψυχοσημασιολογίας έχει επικρατήσει να αναφέρεται στην βιβλιογραφία και ως το πρόβλημα της θεμελίωσης των συμβόλων (The Symbol Grounding Problem) από το ομώνυμο άρθρο του S. Harnad (1990). Το συγκεκριμένο πρόβλημα αποτελεί πλέον ένα από τα μεγαλύτερα εμπόδια για την ολοκλήρωση της ΑΥΘΝ. Οι διάφορες θεωρίες που έχουν προταθεί για την επίλυσή του παρουσιάζουν όλες πολλά προβλήματα. Για το λόγο αυτό δεν υπάρχει κάποια συναίνεση μεταξύ των θεωρητικών της γνωσιακής επιστήμης όσον αφορά τον τρόπο επίλυσης του προβλήματος της ψυχοσημασιολογίας. Αυτό που γίνεται κοινά αποδεκτό, ωστόσο, είναι το ίδιο το πρόβλημα: όλοι σχεδόν³²⁴ οι υποστηρικτές της ΑΥΘΝ συμφωνούν ότι το πρόβλημα της θεμελίωσης των συμβόλων είναι πραγματικό και ότι χωρίς την επίλυσή του ολόκληρη η ΑΥΘΝ παραμένει μετέωρη.

Το ζητούμενο λοιπόν για την ΑΥΘΝ, αλλά και για κάθε άλλη τροποποιημένη αναπαραστασιακή θεωρία του νου, είναι η ανάπτυξη μιας φυσιοκρατικής θεωρίας για το πώς αποκτούν οι νοητικές αναπαραστάσεις το περιεχόμενό τους.

Γενικά, στη σύγχρονη φιλοσοφία του νου υπάρχουν δύο γενικές κατευθύνσεις για την αντιμετώπιση του συγκεκριμένου προβλήματος: ο ιντερναλισμός (internalism) ή ατομικισμός (individualism) και ο εξτερναλισμός (externalism) ή αντι-ατομικισμός (anti-individualism)³²⁵.

7.2.1 Ιντερναλισμός

Στο πλαίσιο της φιλοσοφίας του νου ο ιντερναλισμός - ατομικισμός αποτελεί μια θέση για τον τρόπο εξατομίκευσης και ταξινόμησης των νοητικών καταστάσεων και των νοητικών διαδικασιών. Σύμφωνα με αυτή τη θέση, οι νοητικές καταστάσεις και διαδικασίες ενός υποκειμένου πρέπει να προσδιορίζονται χωρίς αναφορά στα αντικείμενα και τις ιδιότητες του κόσμου μέσα στον οποίο το υποκείμενο είναι ενταγμένο. Ο ατομικιστής υποστηρίζει ότι το περιεχόμενο των νοητικών καταστάσεων μπορεί να περιγραφεί ανεξάρτητα από το εάν υπάρχει ή όχι στο περιβάλλον κάποιο συγκεκριμένο αντικείμενο ή ιδιότητα. Ο Tyler Burge, βασικός πολέμιος του ατομικισμού, προσδιορίζει αυτή τη θέση ως εξής: "οι νοητικές φύσεις όλων των νοητικών καταστάσεων (και συμβάντων) είναι τέτοιες ώστε δεν υπάρχει καμιά βαθιά εξατομικευτική σχέση μεταξύ της ύπαρξης ενός ατόμου σε καταστάσεις ορισμένων ειδών

³²⁴ Ένας σημαντικός θεωρητικός της γνωσιακής επιστήμης που διαφωνεί με αυτή την τοποθέτηση είναι ο Robert Cummins (1996).

³²⁵ Η διάκριση ατομικισμού / αντι-ατομικισμού δεν θα πρέπει να συγχέεται με τη διάκριση ατομισμού (atomism) / ολισμού. Όπως θα δούμε στην ενότητα 7.2.5 μια ατομικιστική θεωρία μπορεί να είναι ολιστική.

και της φύσης των φυσικών ή κοινωνικών περιβαλλόντων του ατόμου" (Burge 1986, σελ. 3-4). Σύμφωνα με αυτόν τον αρνητικό προσδιορισμό, ο ατομικισμός θέτει μια ριζική διάκριση ανάμεσα στις νοητικές καταστάσεις ενός ατόμου και το φυσικό και κοινωνικό περιβάλλον μέσα στο οποίο βρίσκεται το άτομο. Οι νοητικές καταστάσεις, σύμφωνα με μια προσφιλή έκφραση, είναι "μέσα στο κεφάλι". Αυτός είναι και ο λόγος που ο ατομικισμός ονομάζεται ορισμένες φορές και ιντερναλισμός αντιπαρατιθέμενος προς τον εξτερναλισμό (externalism). Ας δούμε, όμως, για ποιους λόγους ορισμένοι ερευνητές υποστηρίζουν τον ατομικισμό.

Καταρχάς, επιστημαίνεται ότι για τη θεωρία του νοητικού περιεχομένου αυτό που έχει σημασία είναι το πώς το υποκείμενο *αναπαριστά* τον κόσμο και όχι το πώς πραγματικά είναι ο κόσμος. Η θέση αυτή μοιάζει, εκ πρώτης όψεως, εύλογη αν σκεφτούμε ότι ορισμένες φορές έχουμε ψευδαισθήσεις, ή ότι πολλές φορές οι νοητικές μας καταστάσεις αφορούν καταστάσεις πραγμάτων που δεν είναι ενεργεία υπαρκτικές (actual). Σε αυτές τις περιπτώσεις, η συμπεριφορά μας φαίνεται να μην καθορίζεται από το πώς είναι ο κόσμος, αλλά από το πώς αναπαριστούμε τον κόσμο³²⁶.

Επίσης, για έναν υλιστή, όπως είναι κάποιος που ταυτίζει (είτε ως τύπο, είτε ως έκτυπο) τα νοητικά φαινόμενα με τα εγκεφαλικά φαινόμενα, ή, έστω, θεωρεί ότι τα νοητικά φαινόμενα επιγίνονται στα εγκεφαλικά φαινόμενα, ο ατομικισμός / ιντερναλισμός εμφανίζεται ως η μόνη συμβατή θέση: από τη στιγμή που ο εγκέφαλος είναι μέσα στο κεφάλι, ο νους δεν θα μπορούσε να βρίσκεται κάπου αλλού³²⁷.

Προς υποστήριξη αυτής της θέσης οι ιντερναλιστές επικαλούνται το ρόλο που παίζουν οι αιτιακές δυνάμεις στις επιστημονικές εξηγήσεις. Το επιχείρημα μπορεί να συνοψισθεί ως εξής: οι επιστήμες εξατομικεύουν τις εξηγητικές κατηγορίες που χρησιμοποιούν βασιζόμενες στις αιτιακές δυνάμεις των φυσικών ειδών που οι κατηγορίες ονομάζουν³²⁸. Όμως, οι αιτιακές δυνάμεις οποιουδήποτε πράγματος επιγίνονται στις φυσικές του ιδιότητες. Κατά συνέπεια, εάν η ψυχολογία θέλει να είναι επιστημονική θα πρέπει οι νοητικές καταστάσεις ενός ατόμου να επιγίνονται στις εγγενείς φυσικές του ιδιότητες. Όπως επιστημαίνει ο Fodor: "Οι αιτιακές δυνάμεις επιγίνονται στην τοπική μικροδομή. Στην ψυχολογική περίπτωση, επιγίνονται στην τοπική νευρωνική δομή" (Fodor 1987, σελ. 45). Το

³²⁶ Θα εξετάσουμε αναλυτικά αυτή τη θέση στο όγδοο κεφάλαιο, ενότητα 8.2.2.

³²⁷ Όπως επιστημαίνει και ο Fodor (1987, σελ. 30): "[κάθε] επιστημονικά χρησιμη έννοια περί ψυχολογικής κατάστασης οφείλει να σέβεται την επιγένεση: η επιγένεση νου/εγκεφάλου ... αποτελεί, εξάλλου, την καλύτερη ιδέα που βρέθηκε μέχρι σήμερα για το πώς η νοητική αιτιότητα είναι δυνατή".

³²⁸ Ο Fodor (1987, σελ. 44) λέει σχετικά: "Η κατηγοριοποίηση στην επιστήμη είναι χαρακτηριστικά ταξινομήση μέσω αιτιακών δυνάμεων. Ταυτότητα αιτιακών δυνάμεων αποτελεί ταυτότητα αιτιακών συνεπειών επί νομολογικά δυνατών πλαισίων."

επιχείρημα αυτό μπορεί να θεωρηθεί, από μια άποψη, ως μια απλή εφαρμογή μιας γενικής ταξινομικής αρχής που υποτίθεται ότι διέπει εν γένει την επιστημονική εξήγηση: "Ο ατομικισμός είναι μια εντελώς γενική μεθοδολογική αρχή στην επιστήμη, μια αρχή που απορρέει απλά από το στόχο του επιστήμονα για αιτιακές εξηγήσεις και με την οποία, κατά συνέπεια, όλες οι επιστημονικές ταξινομήσεις οφείλουν να συμμορφώνονται" (Fodor 1987, σελ. 42-43)³²⁹.

Τέλος, ένας άλλος λόγος για την υιοθέτηση του ατομικισμού είναι η υποτιθέμενη αυθεντία του πρώτου προσώπου: αν πιστεύω ότι υπάρχει το *x* αντικείμενο στο οπτικό μου πεδίο μπορεί κάλλιστα να κάνω λάθος, ωστόσο αυτό για το οποίο δεν μπορώ να κάνω λάθος είναι ότι έχω αυτήν την πεποίθηση. Επίσης, πολλές φορές οι νοητικές μας καταστάσεις παραμένουν αυστηρά ιδιωτικές. Για παράδειγμα, μπορώ να ανακαλέσω στην μνήμη μου κάποιο συμβάν χωρίς να το αποκαλύψω σε κανέναν άλλο. Σε αυτή την περίπτωση μοιάζει φυσικό να υποθέσει κανείς ότι κάτι συμβαίνει εντός μου στο οποίο και οφείλεται η ανάμνησή μου.

Δύο συγκεκριμένες εκφράσεις του ατομικισμού αποτελούν ο μεθοδολογικός σολιψισμός του Fodor και η αρχή της αυτονομίας του Stich.

7.2.1.1 Μεθοδολογικός Σολιψισμός

Σύμφωνα με τον Fodor (1980, σελ. 244), η ψυχολογία πρέπει να ασχολείται μόνο με τις "στενές" (narrow) ψυχολογικές καταστάσεις, δηλαδή καταστάσεις που δεν προϋποθέτουν "την ύπαρξη κανενός άλλου ατόμου πέραν του υποκειμένου στο οποίο αποδίδονται αυτές οι καταστάσεις"³³⁰. Η ταξινόμηση των ψυχολογικών καταστάσεων μέσω των περιεχομένων τους πρέπει να είναι συμβατή με την ταξινόμηση των ψυχολογικών καταστάσεων μέσω των μορφικών (formal) τους ιδιοτήτων: "οι νοητικές καταστάσεις είναι διακριτές κατά περιεχόμενο μόνο εάν μπορούν να ταυτισθούν με σχέσεις μορφικά διακριτών αναπαραστάσεων" (Fodor 1980, σελ. 227). Η έμφαση που δίνει ο Fodor στην έννοια των μορφικών ιδιοτήτων των αναπαραστάσεων έχει να κάνει με το γεγονός ότι τέτοιου τύπου ιδιότητες είναι τοπικές, δηλαδή εσωτερικές στον οργανισμό.

³²⁹ Ο ίδιος ο Fodor (1987) κρατά μια συμβιβαστική στάση μεταξύ εξτερναλισμού και ιντερναλισμού και γι' αυτό επιχειρεί να δείξει ότι η ιδέα ότι οι νοητικές καταστάσεις επιγίνονται στις εγκεφαλικές δεν αντιφάσκει με την ιδέα ότι η αναφορά μιας νοητικής κατάστασης (δηλαδή, κάτι εκτός του κεφαλιού) καθορίζει το περιεχόμενό της. Η στάση αυτή του Fodor δημιουργεί ανυπέρβλητα προβλήματα που, όμως, δεν θα μας απασχολήσουν σε αυτή την εργασία (βλ. Baker 1995, σελ. 42-56 και Wilson 1995, σελ. 31-58).

³³⁰ Ο Fodor ακολουθεί εδώ τον Putnam 1975, σελ. 220. Ο Putnam, όπως θα δούμε στην ενότητα 7.2.3, είναι ουσιαστικά ο πρώτος που εισήγαγε την διάκριση ιντερναλισμού - εξτερναλισμού και στενού - ευρέως περιεχομένου.

7.2.1.2 Η Αρχή της Αυτονομίας

Η αρχή της αυτονομίας του Stich βασίζεται σε μια παρεμφερή ιδέα: "Η βασική ιδέα σε αυτή την αρχή είναι ότι οι καταστάσεις και οι διαδικασίες που θα πρέπει να αφορούν τον ψυχολόγο είναι εκείνες που επιγίνονται στην τρέχουσα, εσωτερική, φυσική κατάσταση του οργανισμού ... [Ο]ποιεσδήποτε διαφορές μεταξύ οργανισμών οι οποίες δεν εκδηλώνονται ως διαφορές στις τρέχουσες, εσωτερικές, φυσικές καταστάσεις θα πρέπει να αγνοούνται από μια ψυχολογική θεωρία ... [Ι]στορικά και περιβαλλοντικά γεγονότα θα είναι σημαντικά από ψυχολογική άποψη μόνο εάν επηρεάζουν την τρέχουσα, εσωτερική, φυσική κατάσταση ενός οργανισμού" (Stich, 1983, σελ. 164-165)³³¹. Ο Stich με την αρχή της αυτονομίας, ουσιαστικά, συνδέει και αυτός τον ατομικισμό με την (τοπική) επιγένεση (supervenience): οι νοητικές καταστάσεις ενός ατόμου οφείλουν να επιγίνονται στις εσωτερικές φυσικές καταστάσεις του³³².

³³¹ Ο Stich προκειμένου να υπερασπισθεί την αρχή της αυτονομίας επινοεί ένα νοητικό πείραμα που είναι γνωστό και ως επίχειρημα της αντικατάστασης -φυσικά δεν είναι καθόλου προφανές το πώς ένα νοητικό πείραμα "ανακηρύσσεται" σε επίχειρημα. Το εν λόγω νοητικό πείραμα έχει ως εξής: έστω ότι ο Α απαγάγεται ενόσω κοιμάται και αντικαθίσταται από μια ρέπλικα πανομοιότυπη με αυτόν από φυσική άποψη. Σε αυτήν την περίπτωση διατείνεται ο Stich (1983, σελ. 165-166) κανείς πέραν του Α και του απαγωγέα δεν θα καταλάβει την διαφορά. Ούτε καν η ίδια η ρέπλικα δεν θα καταλάβει την διαφορά. Ο Stich δεν υποστηρίζει εδώ ότι ο Α και η ρέπλικα είναι πανομοιότυποι από νοητική άποψη, διότι δέχεται ότι διαφέρουν από ιστορική άποψη. Αυτό που υποστηρίζει είναι ότι ο Α και η ρέπλικα θα συμπεριφέρονται κατά τον ίδιο τρόπο, ή, ακριβέστερα, ότι θα κάνουν τις ίδιες σωματικές κινήσεις. Εφόσον, μάλιστα, η ψυχολογία είναι η επιστήμη που στοχεύει στην εξήγηση της συμπεριφοράς, κάθε κατάσταση, διαδικασία ή ιδιότητα που δεν χαρακτηρίζει τόσο τη ρέπλικα όσο και τον Α δεν αφορά την ψυχολογία.

Δεν θεωρούμε ότι το παραπάνω "επίχειρημα" αποδεικνύει οτιδήποτε. Στην καλύτερη περίπτωση απλώς εκκαλεί τις διαισθήσεις όσων είναι ήδη πεπεισμένοι για την ορθότητα της αρχής της αυτονομίας. Κάποιος που δεν αποδέχεται την αρχή της αυτονομίας θα μπορούσε, αντίθετα, να υποστηρίξει ότι από τη στιγμή που ο Α έχει διαφορετική ιστορία από την ρέπλικα έπεται ότι έχει και διαφορετικές μνήμες και οι διαφορετικές μνήμες προκαλούν διαφορετικές πράξεις.

³³² Η έννοια της επιγένεσης εισήχθη στη φιλοσοφία του νου γύρω στο 1970 από τον Davidson (1980a): "...τα νοητικά χαρακτηριστικά είναι υπό μια έννοια εξαρτώμενα, ή επιγινόμενα, από τα φυσικά χαρακτηριστικά. Μια τέτοια επιγένεση θα μπορούσε να εκληφθεί ότι σημαίνει ότι δεν μπορούν να υπάρχουν δύο συμβάντα όμοια από κάθε φυσική άποψη, αλλά διαφορετικά από κάποια νοητική άποψη, ή ότι ένα αντικείμενο δεν μπορεί να μεταβληθεί από κάποια νοητική άποψη χωρίς να μεταβληθεί από κάποια φυσική άποψη". Στη συνέχεια η έννοια της επιγένεσης αναλύθηκε διεξοδικά από πολλούς φιλοσόφους και προτάθηκαν διάφορες παραλλαγές της (βλ. Kim 1993).

7.2.2 Εξτερναλισμός

Για να κατανοήσουμε καλύτερα τι ακριβώς προσβέδει ο ατομικισμός είναι σκόπιμο να εξετάσουμε την αντίπαλη θέση, τον αντι-ατομικισμό ή εξτερναλισμό. Η θέση αυτή στην γενικότερη διατύπωση της είναι ότι το περιεχόμενο των νοητικών καταστάσεων δεν εξαρτάται μόνο από τα εγγενή (intrinsic) χαρακτηριστικά του υποκειμένου αλλά εξαρτάται επίσης από την ιστορία του υποκειμένου και το φυσικό και κοινωνικό του περιβάλλον. Εξ' ου και το σύνθημα του Putnam "οι σημασίες δεν είναι μέσα στο κεφάλι".

Μια πρώτη διατύπωση του εξτερναλισμού μπορεί να ανιχνευτεί στο πλαίσιο της γνωσιολογίας και συγκεκριμένα στην προσέγγιση που ονομάζεται θεωρία αξιοπιστίας (reliabilism) (Armstrong 1973, Dretske 1969, Nozick 1981, Goldman 1986). Σύμφωνα με αυτή την προσέγγιση, η γνώση (και η δικαιολόγηση των πεποιθήσεων) δεν είναι ζήτημα θεμελίωσης σε βασικές πεποιθήσεις ούτε συνεκτικότητα με άλλες πεποιθήσεις. Η γνώση εξαρτάται από το εάν οι πεποιθήσεις συνδέονται με το σωστό τρόπο με τα γεγονότα στα οποία αναφέρονται - μια σύνδεση που είναι εξωτερική ως προς το σύστημα πεποιθήσεων του υποκειμένου και ως τέτοια μη προσβάσιμη γνωσιακά από το υποκείμενο (Dretske 1969). Σύμφωνα λοιπόν με τη θεωρία αξιοπιστίας, μια πεποίθηση είναι δικαιολογημένη εάν παράγεται από μια ακολουθία αξιόπιστων ψυχολογικών διαδικασιών. Ο Shoemaker (1994, p.268) συνοψίζει την θεωρία αξιοπιστίας ως εξής: "Ο νους μας είναι έτσι συγκροτημένος, ή ο εγκέφαλός μας είναι έτσι "καλωδιωμένος", ώστε για ένα μεγάλο εύρος νοητικών καταστάσεων, η ύπαρξη κάποιου σε μια συγκεκριμένη νοητική κατάσταση του προκαλεί, κάτω από ορισμένες συνθήκες, την πεποίθηση ότι βρίσκεται σε αυτή τη νοητική κατάσταση. ... Οι πεποιθήσεις που παράγονται κατ' αυτόν τον τρόπο θα θεωρούνται γνώση, όχι εξαιτίας της ποσότητας ή της ποιότητας των αποδεικτικών στοιχείων (evidence) στα οποία αυτές στηρίζονται (γιατί δεν στηρίζονται σε κανένα αποδεικτικό στοιχείο), αλλά εξαιτίας της αξιοπιστίας του μηχανισμού μέσω του οποίου δημιουργούνται"

Η επίκληση των ψυχολογικών διαδικασιών φανερώνει ότι η θεωρία αξιοπιστίας συνδέεται άμεσα με το εγχείρημα της γνωσιακής επιστήμης. Πάντως, θα πρέπει να επισημάνουμε ότι η γνωσιολογική αυτή έκφανση του εξτερναλισμού αφορά ένα μικρότερο εύρος φαινομένων, δηλαδή αποκλειστικά τις δικαιολογημένες πεποιθήσεις. Ενώ, ο εξτερναλισμός ως προς την προέλευση του περιεχομένου των νοητικών καταστάσεων αφορά, προγραμματικά τουλάχιστον, όχι μόνο όλες τις πεποιθήσεις (δικαιολογημένες ή μη), αλλά, γενικότερα, όλες τις προτασιακές στάσεις³³³.

³³³ Όπως θα δούμε, τελικά, στην ενότητα 7.3, η παραδοχή ότι το υποκείμενο δεν έχει γνωσιακή πρόσβαση στις διαδικασίες μέσω των οποίων αποκτά τη γνώση δεν

Έχοντας σχηματίσει μια πρώτη εικόνα για το τι πρεσβεύει ο ιντερναλισμός και ο εξτερναλισμός, μπορούμε τώρα να εξετάσουμε τι είδους θεωρίες για την προέλευση του περιεχομένου έχουν αναπτυχθεί με αφετηρία αυτές τις θέσεις. Αυτή η εξέταση από μεθοδολογική άποψη γίνεται για δύο αλληλοδιαπλεκόμενους λόγους:

1. Θα μας επιτρέψει να παρακολουθήσουμε εκ του σύνεγγυς την εσωτερική δυναμική των διαφόρων εξειδικευμένων θεωριών για την φυσικοποίηση του νοητικού περιεχομένου και να εντοπίσουμε τα προβλήματα γύρω από τα οποία αυτές συγκροτούνται. Μια τέτοια εξέταση θα μπορούσε να χαρακτηριστεί "εσωτερική", γιατί δεν επερωτά τα βασικά μεταφυσικά θεμέλια πάνω στα οποία στηρίζονται αυτές οι θεωρίες, αλλά ελέγχει το πώς απαντούν στα προβλήματα που οι ίδιες οι θεωρίες εντοπίζουν και επιχειρούν να επιλύσουν. Αυτή η "εσωτερική" εξέταση υιοθετεί μια ευμενή (charitable) στάση έναντι των θεωριών με στόχο την καλύτερη κατανόηση αυτών υπό τους δικούς τους όρους. Ταυτόχρονα, η διερεύνηση της επάρκειας των απαντήσεων που δίδουν οι θεωρίες στα προβλήματα γύρω από τα οποία οι ίδιες ρητά αναπτύσσονται, προσφέρει ένα στέρεο έδαφος για μια "εσωτερική" κριτική.
2. Η "εσωτερική" κριτική εξέταση των διαφόρων εξειδικευμένων θεωριών για την φυσικοποίηση του νοητικού περιεχομένου δημιουργεί μια ισχυρότερη νομιμοποίηση οποιασδήποτε επακόλουθης "εξωτερικής" κριτικής όλων αυτών των εγχειρημάτων, δηλαδή οποιασδήποτε κριτικής που δεν στέκεται στις λεπτομέρειες των θεωριών αλλά επερωτά τις μεταφυσικές προϋποθέσεις στις οποίες αυτές στηρίζονται. Έτσι, ο συνήθως αφοριστικός και σχηματικός χαρακτήρας της "εξωτερικής" κριτικής γίνεται περισσότερο αποδεκτός.

7.2.3. Η αιτιακή - ιστορική θεωρία της αναφοράς ως αφετηριακό σημείο της διαμάχης ιντερναλισμού - εξτερναλισμού.

Πρωτοπόροι στην ανάπτυξη μιας εξτερναλιστικής σημασιολογίας είναι ο Kripke (1972) και ο Putnam (1975b). Οι θεωρίες τους, παρότι είχαν ως στόχο να εξηγήσουν την προέλευση του περιεχομένου όρων της φυσικής γλώσσας, τελικά άσκησαν σημαντική επιρροή και σε όσους επεχείρησαν να αναπτύξουν ανάλογες προσεγγίσεις για το περιεχόμενο των νοητικών αναπαραστάσεων. Γι' αυτό αξίζει να εξετάσουμε τα βασικά στοιχεία αυτών των θεωριών.

Ο Kripke ανέπτυξε μια νέα θεωρία για τη σημασία των κυρίων ονομάτων απορρίπτοντας την θεωρία ανάλυσης των ονομάτων μέσω περιγραφών (Description Theory of Names) που απέδιδε στον Frege και

θα πρέπει να θεωρηθεί κριτήριο διάκρισης των εξτερναλιστικών από τις ιντερναλιστικές θεωρίες όσον αφορά την προέλευση του νοητικού περιεχομένου.

στον Russell. Αργότερα, ο Putnam ακολούθησε την ίδια στρατηγική για τους γενικούς όρους που αναφέρονται σε φυσικά είδη. Στα επόμενα θα εστιάσουμε την προσοχή μας στη θεωρία του Putnam όπως αυτή εισάγεται κυρίως στο ιστορικό του άρθρο "The Meaning of 'Meaning'" (Putnam 1975b). Προκειμένου να γίνει κατανοητή αυτή η θεωρία θα πρέπει να εξεταστεί σε σχέση με τις προηγούμενες θεωρίες σημασίας (ή, με την ορολογία του κεφαλαίου 3, γλωσσικού περιεχομένου) στις οποίες επιτίθεται. Τα βασικά στοιχεία που χαρακτηρίζουν, σύμφωνα με τον Putnam (1975b, σελ. 216-219), τις παραδοσιακές θεωρίες είναι τα ακόλουθα:

1. Τουλάχιστον από το Μεσαίωνα και μετά, οι παραδοσιακές θεωρίες σημασίας βασίζονται σε μια αμφισημία που υποτίθεται ότι ανακάλυψαν όσον αφορά την έννοια της σημασίας. Η αμφισημία αυτή έχει επιχειρηθεί να διευκρινισθεί με διάφορες διακρίσεις, γνωστότερες από τις οποίες είναι οι διακρίσεις έκτασης / έντασης και αναφοράς / νοήματος. Η έκταση ενός όρου αποτελείται ουσιαστικά από το σύνολο των νοημάτων για τις οποίες ο όρος αληθεύει. Η έννοια της 'έντασης', από την άλλη μεριά, παρέμεινε αρκετά σκοτεινή καθότι επιχειρήθηκε να προσδιορισθεί μέσω της εξίσου ασαφούς έννοιας της 'έννοιας'.
2. Θεωρήθηκε ως αυτονόητο ότι, ενώ μπορούν να υπάρξουν όροι με ίδια έκταση και διαφορετική ένταση, το αντίθετο δεν είναι δυνατόν να συμβεί: "δύο όροι δεν μπορούν να διαφέρουν στην έκταση και να έχουν την ίδια ένταση" (Putnam 1975b, σελ. 218). Αυτό, πιθανολογεί ο Putnam, οφείλεται στο ότι η ένταση ή έννοια ενός όρου θεωρήθηκε ως μια σύζευξη κατηγορημάτων (δηλαδή, ως ένα σύνολο περιγραφών) και ως εκ τούτου, ως παρέχουσα τις ικανές και αναγκαίες συνθήκες προκειμένου μια οντότητα να ανήκει στην έκταση του όρου. Έτσι, η ένταση θεωρήθηκε ότι έχει ένα πρωτείο έναντι της έκτασης, υπό την έννοια ότι την καθορίζει.
3. Η σύνδεση της σημασίας (ως έντασης) με την έννοια οδήγησε στο να θεωρείται η σημασία ως κάτι νοητικό. Ακόμα και οι θεωρίες του Frege ή του Carnap, που εξεγέρθηκαν κατά του ψυχολογισμού, δεν κατορθώνουν, σύμφωνα με τον Putnam, να αποφύγουν αυτή την σύνδεση. Διότι ναι μεν εκλαμβάνουν τις έννοιες ως αφηρημένες οντότητες αντί για νοητικές οντότητες³³⁴, ωστόσο "η 'σύλληψη' αυτών των αφηρημένων οντοτήτων εξακολουθεί να αποτελεί ένα ατομικό ψυχολογικό ενέργημα."³³⁵

Τα βασικά αυτά στοιχεία που χαρακτηρίζουν τις παραδοσιακές θεωρίες σημασίας, σύμφωνα με τον Putnam (1975b, σελ. 219)

³³⁴ Βλ. Κεφάλαιο 3, ενότητα 3.3.1.2.

³³⁵ Putnam 1975, σελ. 218 και σελ. 222.

αποκαλύπτουν δύο κεντρικές παραδοχές που συνοπτικά μπορούν να διατυπωθούν ως εξής:

1. η ψυχολογική κατάσταση καθορίζει την ένταση³³⁶
2. η ένταση καθορίζει την έκταση

Από τις δύο αυτές παραδοχές απορρέει ως λογική συνέπεια η παραδοχή ότι η ψυχολογική κατάσταση καθορίζει την έκταση η θέτει ως στόχο του να αναφέρει την τελευταία παραδοχή όσον αφορά τους όρους που αναφέρονται σε φυσικά είδη. Έτσι, λοιπόν, επιχειρεί να δείξει ότι η ψυχολογική κατάσταση ενός υποκειμένου δεν αποτελεί ούτε επαρκή ούτε αναγκαία συνθήκη για τον καθορισμό της έκτασης ούτε επαρκή εννοει ψυχολογική κατάσταση με τη στενή έννοια, όπως δηλαδή αυτή προσδιορίζεται αν κανείς υιοθετήσει τη στάση του μεθοδολογικού σολιτισμού. Με άλλα λόγια, στο συγκεκριμένο πλαίσιο η ψυχολογική κατάσταση κατανοείται υπερναλιστικά.

Η παραδοχή ότι η ψυχολογική κατάσταση καθορίζει την έκταση δεν είναι αναγκαία διότι όταν ένα υποκείμενο αναφέρεται σε ένα μη τυπικό έκτυπο ενός φυσικού είδους ενδέχεται να βρίσκεται σε ένα μη τυπικό κατάσταση τελείως διαφορετική από εκείνη στην οποία θα βρισκόταν αν αναφερόταν σε ένα τυπικό έκτυπο του ίδιου φυσικού είδους. Στην πρώτη περίπτωση, με άλλα λόγια, το υποκείμενο ενδέχεται να σκέφτεται το μη τυπικό έκτυπο του φυσικού είδους υπό μια περιγραφή πολύ διαφορετική από εκείνη υπό την οποία σκέφτεται ένα τυπικό έκτυπο του φυσικού είδους. Ωστόσο, το γεγονός ότι το υποκείμενο βρίσκεται κάθε φυσικού διαφορετική ψυχολογική κατάσταση δεν σημαίνει ότι αναφέρεται και σε διαφορετικά φυσικά είδη. Όπως λέει και ο Putnam (1975b, σελ. 250):

³³⁶ Αν διατυπωθεί με την φρεγκεανή ορολογία, αυτή η παραδοχή ισοδυναμεί με την παραδοχή ότι η ψυχολογική κατάσταση καθορίζει το νόημα. Όπως λέει και ο Putnam (1975, σελ. 222): "η ψυχολογική κατάσταση μοναδικά καθορίζει με αντιστρέφει τη 'φορά' καθορισμού". Ωστόσο, αυτό είναι μια περιρμηγία του Frege διότι είναι εκείνο που προηγείται λογικά και "καθορίζει" οντότητα, δηλαδή το νόημα, κατά τη σύλληψή του. Σε ένα μεταγενέστερο έργο του Putnam (1988, σελ. 129), μπορούμε να ανιχνεύσουμε μια ελαφρώς διαφορετική προσέγγιση όσον αφορά τη σχέση 'ελατωνικών' οντοτήτων και ψυχολογική κατάσταση. Ο Putnam (1988, σελ. 129) περιεργάζεται με τη θεωρία του Frege: "...οι ομιλητές υποτίθεται ότι είναι περισσότερο συμβατή με τη θεωρία του Frege". ...οι ομιλητές υποτίθεται ότι είναι ικανοί να κατανοήσουν τη νοητική προσοχή τους σε έννοιες μέσω κάποιου πράγματος που μοιάζει με την αντλήνη και εάν η Α και η Β είναι διαφορετικές έννοιες, τότε εστιάζοντας την προσοχή στην Α και εστιάζοντας την προσοχή στην Β είναι διαφορετικές νοητικές καταστάσεις. Έτσι, ακόμα και σε αυτές τις θεωρίες, η νοητική κατάσταση του ομιλητή καθορίζει σε ποια έννοια εστιάζει την προσοχή του και, ως εκ τούτου, καθορίζει σε τι αναφέρεται."

ΚΕΦΑΛΑΙΟ 7: ΚΡΙΤΙΚΗ ΤΩΝ ΠΡΟΓΡΑΜΜΑΤΩΝ ΦΥΣΙΚΟΠΟΙΗΣΗΣ

"Τρίποδες τρίρες και λεικές τρίρες δεν είναι λογικά αντιφατικές οντότητες ... Εάν οι τρίρες είχαν τις τρίες τους δεν θα έπαυαν ως εκ αν έχαναν τα φτερά τους."

Ο Putnam προκειμένου να δείξει ότι η ψυχολογική κατάσταση ενός υποκειμένου δεν αποτελεί επαρκή συνθήκη για τον καθορισμό της έκτασης όρων που αναφέρονται σε φυσικά είδη επισημαίνει ότι η ψυχολογική κατάσταση καθορίζει την έκταση η θέτει ως στόχο του να αναφέρει την τελευταία παραδοχή όσον αφορά τους όρους που αναφέρονται σε φυσικά είδη. Έτσι, λοιπόν, επιχειρεί να δείξει ότι η ψυχολογική κατάσταση ενός υποκειμένου δεν αποτελεί ούτε επαρκή ούτε αναγκαία συνθήκη για τον καθορισμό της έκτασης ούτε επαρκή εννοει ψυχολογική κατάσταση με τη στενή έννοια, όπως δηλαδή αυτή προσδιορίζεται αν κανείς υιοθετήσει τη στάση του μεθοδολογικού σολιτισμού. Με άλλα λόγια, στο συγκεκριμένο πλαίσιο η ψυχολογική κατάσταση κατανοείται υπερναλιστικά.

Η έννοια του δυνατού κόσμου έχει μια μακρά ιστορία στη φιλοσοφία³³⁷. Ιδιαίτερα, στον αιώνα μας χρησιμοποιήθηκε κυρίως στο πλαίσιο της τροπικής λογικής, όπου και απέκτησε μια σχετικά σαφήνεια. Πιο συγκεκριμένα, η έννοια του δυνατού κόσμου κατανοήθηκε στο πλαίσιο της τροπικής λογικής ως ένα σύνολο συνεπών μεταξύ τους προτάσεων. Ωστόσο, στο νοητικό πεδίο που προτείνει ο Putnam στο έργο του "The Meaning of 'Meaning'" δεν είναι ιδιαίτερα σαφής η έννοια του δυνατού κόσμου. Το κάνουμε αυτό γιατί αισθανόμαστε ότι από αρκετές απόψεις η έννοια έχει νόημα και είναι ιδιαίτερα σημαντική ακόμα κι αν χρειάζεται να γίνει περισσότερο ακριβώς υποθέσουμε, επίσης, ότι τουλάχιστον σε ορισμένες περιπτώσεις είναι δυνατό να μιλάμε για το ίδιο άτομο ως υπάρχον σε περισσότερους από έναν δυνατούς κόσμους" (Putnam 1975b, σελ. 230). Ωστόσο, όλα αυτά είναι δυνατούς κόσμους" (Putnam 1975b, σελ. 230). Ωστόσο, όλα αυτά στο πλαίσιο της τροπικής λογικής έχει νόημα να μιλάμε για δυνατούς κόσμους ως σύνολο μη-αντιφατικών προτάσεων, πολύ δύσκολα θα γινόταν αποδεκτό ότι οι δυνατοί κόσμοι που μπορεί κανείς να φανταστεί όταν κάνει φιλοσοφικά νοητικά περάσματα αποτελούν σύνολα μη αντιφατικών

³³⁷ Βλ. Leibniz 1985.

μεταξύ τους προτάσεων. Πολύ λιγότερο, δε, θα γινόταν αποδεκτό ότι ο ενεργεια υπαρκτός κόσμος αποτελεί, ή έστω μπορεί να περιγραφεί ως, ένα σύνολο μη αντιφατικών μεταξύ τους προτάσεων. Εξάλλου, αν ο ενεργεια υπαρκτός κόσμος και οι διάφοροι δυνατοί κόσμοι μπορούσαν να εκφραστούν ως ένα σύνολο μη αντιφατικών μεταξύ τους προτάσεων, τότε δεν θα υπήρχε η ανάγκη προσφυγής στη διαίσθηση την οποία συστηματικά επικαλούνται όσοι χρησιμοποιούν τα νοητικά πειράματα στη φιλοσοφία. Επίσης, ένα άλλο σημαντικό ερώτημα που παραμένει αναπάντητο είναι τι διαφοροποιεί τον ενεργεια υπαρκτό κόσμο από έναν δυνατό κόσμο. Δεν θα επεκταθούμε, ωστόσο, περισσότερο στο πρόβλημα των δυνατών κόσμων διότι αυτό το ζήτημα υπερβαίνει το πλαίσιο της παρούσας εργασίας. Ως εκ τούτου, στα επόμενα δεν θα αμφισβητήσουμε τη βασιμότητα της έννοιας του δυνατού κόσμου.

Ας δούμε, λοιπόν, εν συντομία το νοητικό πείραμα με την δίδυμη Γη. Ο Α και ο Β που είναι πανομοιότυποι μόριο προς μόριο, βρίσκονται ο ένας στην Γη και ο άλλος σε έναν πλανήτη ενός δυνατού κόσμου -την δίδυμη Γη- ο οποίος είναι φαινομενικά πανομοιότυπος με την Γη. Επίσης, τόσο ο Α όσο και ο Β, οι οποίοι, σημειωτέον, δεν γνωρίζουν τη χημική σύσταση του νερού, όταν θέλουν να αναφερθούν στο νερό χρησιμοποιούν τη λέξη "νερό". Ωστόσο, οι δύο πλανήτες, παρότι φαινομενικά (δηλαδή, μακροσκοπικά) πανομοιότυποι, διαφέρουν ως προς την μοριακή σύσταση του νερού: στην δίδυμη Γη το νερό δεν έχει την χημική σύσταση H_2O αλλά XYZ . Παρότι, δηλαδή, το νερό έχει τις ίδιες μακροσκοπικές ιδιότητες - είναι ένα άχρωμο και άοσμο υγρό που ξεδιψά τους οργανισμούς, σβήνει τη φωτιά κ.λπ. - και, κατά συνέπεια, προσδιορίζεται με την ίδια περιγραφή (δηλαδή, την ίδια ένταση), η πραγματική του υπόσταση διαφέρει στους δύο πλανήτες. Άρα, συμπεραίνει ο Putnam, η αναφορά της λέξης "νερό" όταν εκφέρεται από τον Β δεν είναι το H_2O αλλά το XYZ και άρα διαφέρει από την αναφορά της ίδιας λέξης όταν αυτή χρησιμοποιείται από τον Α. Έτσι, λοιπόν, ενώ ο Α και ο Β βρίσκονται σε πανομοιότυπες ψυχολογικές καταστάσεις³³⁸ όταν προφέρουν την λέξη "νερό" τελικά αναφέρονται σε διαφορετικά πράγματα³³⁹.

³³⁸ Μια ατυχής παρενέργεια της επιλογής του νερού ως παραδείγματος φυσικού είδους είναι ότι ο ανθρώπινος οργανισμός αποτελείται και από νερό, κατά συνέπεια δεν είναι δυνατόν ένας γήινος να είναι πανομοιότυπος με τον δίδυμο του. Ωστόσο, αυτή η παρενέργεια είναι εντελώς συμπτωματική και δεν επηρεάζει την ουσία του νοητικού πειράματος του Putnam. Για παράδειγμα, θα μπορούσαμε να χρησιμοποιήσουμε αντί για νερό ένα οποιοδήποτε άλλο φυσικό είδος που δεν περιέχεται στον ανθρώπινο οργανισμό.

³³⁹ Ένα εύλογο ερώτημα είναι για ποιον ακριβώς λόγο υποστηρίζει ο Putnam ότι ο Α και ο Β βρίσκονται σε πανομοιότυπες (στενές) ψυχολογικές καταστάσεις; Το γεγονός ότι ο Α και ο Β είναι πανομοιότυποι από μόνο του σίγουρα δεν αρκεί, διότι ακόμα και ο Α από μόνος του μπορεί να βρεθεί σε πολλές διαφορετικές

Ο Putnam υποστηρίζει ότι η λέξη "νερό" θα αναφερόταν σε διαφορετικά πράγματα ακόμα και αν (σε κάποια προηγούμενη εποχή πχ. γύρω στο 1750) κανένα μέλος των δύο κοινοτήτων δεν γνώριζε την χημική σύσταση του νερού της Γης και της διδυμής Γης³⁴⁰. Επίσης, αν ο Α μεταφερόταν στην διδυμη Γη, παράτι θα ονόμαζε, όπως και ο Β, το νερό της διδυμής Γης "νερό", αυτός, σε αντίθεση με τον Β, θα έκανε λάθος, διότι θα πίστευε ότι το νερό της διδυμής Γης θα ήταν η ίδια υπόσταση που γνώριζε με το όνομα "νερό" στην Γη. Έτσι λοιπόν, σύμφωνα πάντα με τον Putnam, η λέξη "νερό" όταν εκφέρεται από το στόμα του Α έχει διαφορετική έκταση από ό,τι όταν εκφέρεται από το στόμα του Β, παρά το γεγονός ότι οι ψυχολογικές καταστάσεις των Α και Β είναι πανομοιότυπες. Αυτό συμβαίνει επειδή οι υποστάσεις στις οποίες εκάστοτε η λέξη "νερό" αναφέρεται είναι διαφορετικές (Putnam 1988, σελ. 31-32). Για όλους αυτούς τους λόγους, αν μπορεί κανείς να τους θεωρήσει ως λόγους, ο Putnam συμπεραίνει ότι η 'σημασία' λέξεων που αναφέρονται σε φυσικά είδη δεν καθορίζεται από την ψυχολογική κατάσταση του χρήστη της λέξης: "meanings' just ain't in the head!"³⁴¹ (Putnam 1975b, σελ. 227). Λέγοντας 'σημασία' (με εισαγωγικά) ο Putnam δεν εννοεί αποκλειστικά την έκταση ενός όρου αλλά κάτι ευρύτερο. Για την ακρίβεια, εννοεί την "διαισθητική προαναλυτική χρήση" (Putnam 1975b, σελ. 224) του όρου, δηλαδή μια χρήση που δεν προϋποθέτει τη διάκριση έκτασης / έντασης ή αναφοράς / νοήματος.

Η μεθοδολογία με την οποία ο Putnam δείχνει ότι η ψυχολογική κατάσταση δεν αποτελεί επαρκή συνθήκη για τον καθορισμό της έκτασης παρουσιάζει αρκετά προβλήματα. Υποδείξαμε ορισμένα από αυτά με την αναφορά που κάναμε στην έννοια των δυνατών κόσμων και στις υποσημειώσεις 336, 339, 340, 341. Ωστόσο, δεν θα επιμείνουμε στην κριτική του επιχειρήματος του Putnam καθότι αυτό αφορά τον τρόπο που

ψυχολογικές καταστάσεις. Για παράδειγμα, όταν ο Α αναφέρεται στις τίγρεις βρίσκεται σε διαφορετική ψυχολογική κατάσταση από εκείνη στην οποία βρίσκεται όταν αναφέρεται στο νερό. Για αυτό το λόγο ο Putnam προσθέτει τον περιορισμό ότι το νερό στην Γη και τη διδυμη Γη έχει τις ίδιες μακροσκοπικές ιδιότητες, δηλαδή εμφανίζεται με το ίδιο τρόπο στον Α και το Β. Με την ορολογία της ΑΥΘΝ αυτό θα σήμαινε ότι ο Α και ο Β έχουν τον ίδιο τύπο νοητικής αναπαράστασης όταν αναφέρονται στο νερό.

³⁴⁰ Ο Putnam (1975, σελ. 224) υποστηρίζει συγκεκριμένα ότι ο Α και ο Β θα "κατανοούσαν τον όρο 'νερό' διαφορετικά το 1750 παρά [θα] βρίσκονταν στην ίδια ψυχολογική κατάσταση". Ωστόσο, εύλογα αναρωπιέται κανείς τι είδους κατανόηση μπορεί να είναι αυτή στην οποία αναφέρεται ο Putnam. Πώς είναι δυνατόν να *κατανοεί* κανείς κάτι που δεν το γνωρίζει καν; Θα επανέλθουμε σε αυτό το ζήτημα στην ενότητα 7.3.

³⁴¹ Αυτό το σλόγκαν από μόνο του είναι, βέβαια, παραπλανητικό διότι αναφέρεται σε όλα τα είδη σημασίας, ενώ το επιχειρήμα του Putnam αφορά μόνο τη σημασία όρων που αναφέρονται σε φυσικά είδη.

αποκτούν περιεχόμενο ορισμένα είδη αναπαραστάσεων της φυσικής γλώσσας και όχι οι νοητικές αναπαραστάσεις. Όπως προεπιταμε, η προσέγγιση του Putnam μας ενδιαφέρει απλώς για να καταδείξουμε το αφετηριακό σημείο των ιντερναλιστικών και εξτερναλιστικών προσεγγίσεων στο πρόβλημα της προέλευσης του νοητικού περιεχομένου.

Μέχρι στιγμής εξετάσαμε την κριτική του Putnam στην ιντερναλιστική προσέγγιση, δηλαδή στη θέση ότι η στενή ψυχολογική κατάσταση καθορίζει την έκταση των όρων που αναφέρονται σε φυσικά είδη. Ας δούμε τώρα τη θετική πρόταση του Putnam όσον αφορά το τι καθορίζει την αναφορά των όρων που ονομάζουν φυσικά είδη. Η πρόταση αυτή αποτέλεσε το κύριο πρότυπο για όλες σχεδόν τις εξτερναλιστικές προσεγγίσεις που αναπτύχθηκαν στη φιλοσοφία του νου.

Αν δεχτούμε ότι η ψυχολογική κατάσταση δεν καθορίζει την έκταση του όρου "νερό", τότε το πρόβλημα που τίθεται είναι με ποιον τρόπο η λέξη "νερό" συνδέεται με την έκταση H_2O . Σίγουρα ο τρόπος που επιτυγχάνεται αυτό δεν εξαρτάται από τη λέξη "νερό" αυτή καθαυτή, διότι η ίδια ακριβώς λέξη στην δίδυμη Γη ονομάζει την υπόσταση XYZ. Τι είναι, λοιπόν, αυτό που καθορίζει την αναφορά της λέξης νερό; Ο Putnam επιχειρεί να δώσει μια απάντηση σε αυτό το ερώτημα επεξεργαζόμενος μια έννοια αιτιακής - ιστορικής σύνδεσης μεταξύ της λέξης και της έκτασής της. Η βασική ιδέα είναι ότι η αναφορά των γενικών όρων που αναφέρονται σε φυσικά είδη προσδιορίζεται από την αιτιακή αλυσίδα που συνδέει την παρουσία χρήση των συγκεκριμένων όρων με προηγούμενες χρήσεις τους φτάνοντας έως την αρχική πράξη "βαπτίσματος" των όρων. Η θεωρία του Putnam ονομάστηκε για αυτό το λόγο "αιτιακή - ιστορική θεωρία της αναφοράς". Ας δούμε ένα παράδειγμα. Ο όρος "νερό" αναφέρεται στο H_2O χάρις σε μια αιτιακή αλυσίδα που εκτείνεται από τις παρούσες χρήσεις του όρου σε προηγούμενες, οι οποίες τελικά φτάνουν στο αρχικό "εισαγωγικό συμβάν" κατά το οποίο αποδόθηκε μια προσεγγιστική περιγραφή - ένα στερεότυπο- στον όρο "νερό". Ο Putnam στο "Explanation and Reference" εξηγεί αυτή την διαφορετικού τύπου σύνδεση των λέξεων με την αναφορά τους, εξετάζοντας τον όρο "ηλεκτρισμός".

"Δεν μπορώ ... να σκεφτώ κάτι που κάθε χρήστης του όρου 'ηλεκτρισμός' πρέπει να γνωρίζει εκτός τού ότι ο ηλεκτρισμός είναι ... κάποιου είδους φυσικό μέγεθος, και, ενδεχομένως, ότι ο 'ηλεκτρισμός' (ή το ηλεκτρικό φορτίο ή φορτία) έχει την ικανότητα να ρέει ή να κινείται. Ο Benjamin Franklin γνώριζε ότι ο ηλεκτρισμός εκδηλώνεται υπό την μορφή σπινθήρων ή κεραυνών· κάποιος άλλος ίσως να γνωρίζει περί ρευμάτων και ηλεκτρομαγνητών· κάποιος άλλος για άτομα που αποτελούνται από θετικά και αρνητικά φορτισμένα σωματίδια. Θα μπορούσαν όλοι να χρησιμοποιούν τον όρο 'ηλεκτρισμός' χωρίς να υπάρχει κάποια διακρίσιμη 'ένταση' (intension) που όλοι να έχουν κοινή. Θέλω να προτείνω ότι αυτό που πράγματι έχουν κοινό είναι το εξής: ότι καθένας τους συνδέεται μέσω ενός συγκεκριμένου είδους αιτιακής αλυσίδας με μια κατάσταση στην οποία δίδεται μια περιγραφή

του ηλεκτρισμού, γενικά μια αιτιακή περιγραφή - δηλαδή, μια περιγραφή που ξεχωρίζει τον ηλεκτρισμό ως το φυσικό μέγεθος που είναι υπεύθυνο για ορισμένα αποτελέσματα με ένα συγκεκριμένο τρόπο" (Putnam 1975a, σελ. 199-200)

Ο Putnam κάνει λόγο εδώ για δύο τύπους αιτιακών συνδέσεων. Ο πρώτος τύπος σύνδεσης καθορίζει το πώς μια "κατά προσέγγιση αληθής περιγραφή" (Putnam 1975a, σελ. 201)- ένα στερεότυπο - αποδίδεται για πρώτη φορά σε έναν όρο. Η περιγραφή αυτή δεν είναι ούτε επαρκής ούτε αναγκαία³⁴². Πρόκειται για ένα "εισαγωγικό συμβάν" (ibid. σελ. 201) -όχι κατ' ανάγκη μοναδικό- κατά το οποίο κάποιος ομιλητής μέσω ενός καταδεικτικού ορισμού αποκτά την ικανότητα να χρησιμοποιεί έναν όρο σύμφωνα με μια κατά προσέγγιση αληθή περιγραφή του. Ο Putnam σχετίζει αυτόν τον τύπο αιτιακής σύνδεσης με την συνεισφορά του φυσικού περιβάλλοντος.

Ο δεύτερος τύπος αιτιακής σύνδεσης σχετίζεται με την συνεισφορά του κοινωνικού περιβάλλοντος και αφορά το πώς η σημασία ενός όρου "μεταδίδεται" από ομιλητή σε ομιλητή. Ο Putnam λέει χαρακτηριστικά:

"Εάν διδάξω την λέξη [‘ηλεκτρισμός’] σε κάποιον άλλον λέγοντάς του ότι η λέξη ‘ηλεκτρισμός’ είναι το όνομα ενός φυσικού μεγέθους, και λέγοντας του ορισμένα γεγονότα για αυτό τα οποία δεν συνιστούν αιτιακή περιγραφή . . . -ακόμα και αν τα γεγονότα που του λέω δεν συνιστούν μια καθοριστική περιγραφή οποιουδήποτε είδους, πολύ περισσότερο μια αιτιακή περιγραφή- παρόλα αυτά, η ύπαρξη της λέξης στο λεξιλόγιό του θα συνδέεται αιτιακά με την ύπαρξη της στο δικό μου λεξιλόγιο, και έτσι, τελικά, με ένα εισαγωγικό συμβάν" (ibid., σελ 201)

Το ενδεχόμενο, στο συγκεκριμένο παράδειγμα, ο μαθητής να αποκτήσει μια πολύ ελλιπότερη περιγραφή του όρου από εκείνη που διαθέτει ο δάσκαλος δεν επηρεάζει, σύμφωνα με τον Putnam, το γεγονός ότι και οι δύο ομιλητές, όταν χρησιμοποιούν τον συγκεκριμένο όρο, αναφέρονται στο ίδιο είδος. Ο μαθητής παρά τις ελλείψεις γνώσεως του δεν έχει πρόβλημα να χρησιμοποιεί τον όρο. Απλώς όταν συναντά αμφίβολες περιπτώσεις μπορεί πάντα να συμβουλευτεί τον δάσκαλο ή άλλους ειδικούς. Έτσι λοιπόν, σύμφωνα με τον Putnam, υπάρχει ένας "καταμερισμός της γλωσσικής εργασίας" σύμφωνα με τον οποίο κάποιοι ειδικοί επιφορτίζονται το έργο τού να γνωρίζουν τις διακρίσεις ανάμεσα στους διάφορους όρους, ενώ οι υπόλοιποι ομιλητές δεν απαιτείται να διαθέτουν τέτοιες γνώσεις, χωρίς αυτό ταυτόχρονα να τους καθιστά μη ικανούς χρήστες των όρων. Η γλώσσα δεν αποτελεί μια ουσιασώς

³⁴² Στην προκειμένη περίπτωση μπορούμε να παρατηρήσουμε ότι αποτελεί μια αιτιακή περιγραφή, δηλαδή μια περιγραφή που καθορίζει την αναφορά του όρου με βάση την αιτία που την προκάλεσε.

ατομικιστική δραστηριότητα, αλλά μια μορφή συνεργατικής δραστηριότητας. Με άλλα λόγια, δεν χρειάζεται καθέτι που είναι αναγκαίο για την χρήση της γλώσσας να αποθηκεύεται μέσα στον νου του εκάστοτε ατόμου (Putnam 1988, σελ. 25). Οι μη ειδικοί ομιλητές όταν χρησιμοποιούν έναν όρο αναφέρονται σε ό,τι αναφέρονται και οι ειδικοί τούς οποίους και αποδέχονται ως έσχατους κριτές.

Συνοψίζοντας μπορούμε να πούμε ότι η θετική πρόταση του Putnam απαρτίζεται από δύο υποθεωρίες. Η πρώτη θεωρία μιλά για τον αρχικό καθορισμό της σημασίας ενός όρου μέσω της εγκαθίδρυσης μιας αιτιακής σύνδεσης μεταξύ του όρου και του αναφερόμενου, ενώ η δεύτερη για τον τρόπο διάδοσης της σημασίας του όρου από ομιλητή σε ομιλητή. Υπάρχουν πολλά προβλήματα που παρουσιάζουν οι δύο αυτές υποθεωρίες η κάθε μια από μόνη της. Ωστόσο, θα αναφερθούμε καταρχάς σε ένα πρόβλημα συμβατότητας που ενδεχομένως υφίσταται μεταξύ των δύο υποθεωριών. Σχηματικά μπορούμε να διατυπώσουμε το πρόβλημα ως εξής: αν οι δύο θεωρίες αναδεικνύουν, η πρώτη τον ρόλο του φυσικού περιβάλλοντος και η δεύτερη τον ρόλο του κοινωνικού περιβάλλοντος στον καθορισμό της σημασίας ενός όρου, τότε σίγουρα δεν θα πρέπει κάθε ένας από τους δύο αυτούς παράγοντες να επαρκεί από μόνος του για τον καθορισμό της σημασίας ενός όρου. Διότι, σε μια τέτοια περίπτωση, θα υπήρχαν δύο ανεξάρτητοι τρόποι πλήρους καθορισμού της σημασίας ενός όρου, οπότε θα ήταν δυνατό για τον ίδιο όρο να καθορίζονται δύο διαφορετικές σημασίες.

Ωστόσο, τα ίδια τα κείμενα του Putnam (1975a) δεν αποκλείουν μια τέτοια ερμηνεία. Εφόσον, η σημασία ενός όρου μπορεί να καθορισθεί με ένα εισαγωγικό συμβάν αυτό σημαίνει ότι η "συνεισφορά του περιβάλλοντος" αποτελεί επαρκή συνθήκη για την παγίωση της σημασίας. Από την άλλη μεριά, ο Putnam μιλώντας για τον καταμερισμό της γλωσσικής εργασίας δηλώνει ότι: "είναι μόνο η κοινωνιολογιστική κατάσταση του συλλογικού γλωσσικού σώματος στο οποίο ανήκει ο ομιλητής αυτή που παγώνει την έκταση". Αν δεχθεί κανείς αυτή την ερμηνεία μπορεί να διακρίνει δύο είδη εξτερναλισμού: τον κοινωνικό και τον μη κοινωνικό εξτερναλισμό (βλ. Bilgrami 1992, σελ. 22). Κύριος εκπρόσωπος του κοινωνικού εξτερναλισμού θεωρείται ο Tyler Burge. Στα επόμενα θα βασιστούμε αποκλειστικά στον μη κοινωνικό εξτερναλισμό όπως προκύπτει από τη θεωρία του Putnam, διότι είναι ακριβώς αυτός ο τύπος εξτερναλισμού που άσκησε τη σημαντικότερη επιρροή στην ανάπτυξη των διαφόρων εξτερναλιστικών θεωριών του νοητικού περιεχομένου.

Προτού στραφούμε όμως σε αυτές τις θεωρίες αξίζει να εξετάσουμε ένα σημαντικό πρόβλημα που αντιμετωπίζει ο μη κοινωνικός εξτερναλισμός του γλωσσικού περιεχομένου και το οποίο δικαιολογεί την ανάπτυξη μη κοινωνικών εξτερναλισμών του νοητικού περιεχομένου.

Ο Putnam, μιλώντας για τη συνεισφορά του φυσικού περιβάλλοντος στον καθορισμό της σημασίας ενός γλωσσικού όρου, επιμελώς προσπαθεί να αφήσει εκτός της θεωρίας του κάθε αναφορά στην

φιλοσοφία του νου γιατί έτσι θα αναγκαζόταν να επανεισαγάγει τις νοητικές καταστάσεις των υποκειμένων που χρησιμοποιούν τους γλωσσικούς όρους. Αυτή η επιλογή, όμως, κάνει τον τρόπο που ο γλωσσικός όρος συνδέεται με (ή κληλατεί) την αναφορά του να φαίνεται μαγικός. Διότι τι άλλο θα μπορούσε να είναι αυτό που εγκαθιδρύει και συντηρεί τη σχέση ενός όρου με την αναφορά του αν όχι τα υποκείμενα που χρησιμοποιούν τον όρο; Πώς, για παράδειγμα, θα μπορούσε να εγκαθιδρυθεί η σχέση ενός όρου με την αναφορά του χωρίς κάποιον ομιλητή που επιτελεί τη πράξη βάπτισης;

Καθίσταται λοιπόν φανερή η ανάγκη συμπλήρωσης της θεωρίας του Putnam με κάποια εξήγηση του τρόπου που οι νοητικές καταστάσεις αποκτούν το περιεχόμενό τους. Έτσι, η θεωρία του Putnam μπορεί να θεωρηθεί ότι παρείχε τόσο τη μεθοδολογία όσο και το κίνητρο για την ανάπτυξη της εξτερναλιστικής σημασιολογίας του νοητικού περιεχομένου. Το πέρασμα από την εξτερναλιστική σημασιολογία της γλώσσας στην εξτερναλιστική σημασιολογία του νου έγινε σχεδόν μηχανικά: ό,τι ισχύει για τις λέξεις της φυσικής γλώσσας θεωρήθηκε ότι μπορεί να δειχθεί και για τις "λέξεις" της νοητικής γλώσσας, δηλαδή για τις νοητικές αναπαραστάσεις. Από αυτήν την κίνηση προέκυψε και το νέο σύνθημα: "οι νοητικές καταστάσεις δεν είναι μέσα στο κεφάλι" (McGinn 1989, σελ. 31).

Ο ίδιος ο Putnam, όπως επισημαίνει ο Burge (1979), δεν κάνει το βήμα από τη εξτερναλιστική σημασιολογία της γλώσσας στην εξτερναλιστική σημασιολογία του νου διότι θεωρεί ότι όταν τα μη-αποβλεπτικά χαρακτηριστικά ενός ατόμου (όπως η χημική σύσταση, η φυσιολογική δομή, οι αποκρίσεις και τα ερεθίσματα κτλ) παραμένουν σταθερά, τότε και το περιεχόμενο των νοητικών του αναπαραστάσεων που αναφέρονται σε φυσικά είδη παραμένει σταθερό. Μάλιστα, ο Putnam ονομάζει αυτό το περιεχόμενο "στενό" θέλοντας να το αντιδιαστείλει από το "ευρύ" περιεχόμενο, δηλαδή το περιεχόμενο που αποδίδουμε σε ένα υποκείμενο αν λάβουμε υπόψη μας το περιβάλλον μέσα στο οποίο αυτό είναι ενταγμένο. Όπως αναφέραμε στην ενότητα 7.2.1.1 η έννοια του στενού περιεχομένου αποτέλεσε αντίστοιχα το σημείο εκκίνησης του ιντερναλιστικού προγράμματος, σύμφωνα με το οποίο οι ψυχολογικές εξηγήσεις πρέπει να βασίζονται μόνο στο στενό περιεχόμενο.

7.2.4 Εξτερναλιστικές προσεγγίσεις

Οι εξτερναλιστικές θεωρίες για την προέλευση του νοητικού περιεχομένου άρχισαν να αναπτύσσονται στα τέλη της δεκαετίας του 1970 και βρίσκονται σε πλήρη άνθηση από τα μέσα της δεκαετίας του 1980. Ο Fodor διαδραμάτισε και σε αυτήν την εξέλιξη ένα ρόλο κλειδί. Ενώ μέχρι τις αρχές της δεκαετίας του 1980 με την υποστήριξη του μεθοδολογικού σολιψισμού ήταν ένας από τους βασικούς υπέρμαχους του ιντερναλισμού, με τα χρόνια άρχισε σιγά-σιγά να μεταστρέφει τη θέση του δοκιμάζοντας διάφορες εξτερναλιστικές λύσεις στο πρόβλημα της προέλευσης του νοητικού περιεχομένου. Οι λύσεις που πρότεινε ο Fodor δεν είναι, ίσως, οι

πιο επιτυχείς, ωστόσο με την κριτική που άσκησε σε αντίπαλες θέσεις και με την ικανότητά του να εκτοπίζει και να ταξινομεί τις ανηθέςεις του, έδωσε μια ουσιαστική ώθηση στην διαμόρφωση της διαμάχης. Για τους λόγους αυτούς θα ανασυγκροτήσουμε το όλο πρόβλημα, τουλάχιστον στην αφητηρία του, με τους όρους του Fodor.

7.2.4.1 Σημασιολογία ενδεικτών

Η σχέση νοητικής αναπαράστασης - αναπαριστώμενου, σύμφωνα με τον Fodor, δεν μπορεί να βασίζεται ούτε στη σύμβαση ούτε στην ομοιότητα³⁴³. Η σύμβαση απορρίπτεται γιατί τα νοητικά σύμβολα είναι έμφυτα, ενώ η ομοιότητα δεν γίνεται δεκτή διότι, σε αντίθεση με την αναπαράσταση, είναι μια συμμετρική σχέση. Για παράδειγμα, αν το α ομοιάζει με το β, τότε και το β ομοιάζει με το α. Αντίθετα, αν το α αναπαριστά το β, τότε το β δεν αναπαριστά το α³⁴⁴. Το συγκεκριμένο επιχείρημα έχει διατυπωθεί από τον Goodman και αναλύεται διεξοδικά στο παράρτημα 1. Ο Goodman, όπως έχουμε ήδη προαναφέρει στο τρίτο κεφάλαιο, δείχνει γενικότερα ότι η ομοιότητα δεν αποτελεί ούτε επαρκή ούτε αναγκαία συνθήκη για να καταστεί ένα πράγμα εικονιστική αναπαράσταση, δηλαδή να αναφέρεται στο απεικονιζόμενο. Η σχέση ομοιότητας από μόνη της επαρκεί μόνο για να καταστήσει ένα αντικείμενο *εγγραφή* και όχι αναπαράσταση (βλ. κεφάλαιο 3). Αν δεχθούμε, λοιπόν, αυτή την κριτική, τα νοητικά είδωλα για τα οποία μιλά ο Kosslyn (βλ. κεφάλαιο 5) δεν αποτελούν απεικονιστικό είδος αναπαράστασης, αλλά κάποιο είδος εγγραφής.

Μια τρίτη δυνατότητα για τη σχέση νοητικής αναπαράστασης - αναπαριστώμενου είναι να εξηγηθεί η αναπαράσταση με βάση την έννοια του *ενδείκτη*³⁴⁵. Ένας ενδείκτης συνδέεται αιτιακά με το ενδεικνυόμενο: ο

³⁴³ Ο Fodor θεωρεί κυριότερο υποστηρικτή της θέσης ότι η σχέση νοητικής αναπαράστασης - αναπαριστώμενου είναι σχέση ομοιότητας τον Hume. Βλ. Fodor & Lepore 1994, σελ. 143.

³⁴⁴ Βλέπε Fodor (1990), σελ. 33. Επίσης, ένας άλλος λόγος που ο Fodor απορρίπτει την ομοιότητα είναι ότι δεν μπορεί να συλλάβει τη διάκριση καθόλου και επιμέρους (Fodor, 1985a).

³⁴⁵ Ο Fodor εδώ κάνει χρήση του τριμερούς χωρισμού των σημείων από τον Peirce σε εικόνες, σύμβολα και ενδείκτες. Παραθέτουμε τους αντίστοιχους ορισμούς του Peirce: "Οτιδήποτε, είτε είναι ποιότητα είτε υπάρχον άτομο είτε νόμος, αποτελεί μια Εικόνα-Ομοίωμα κάποιου πράγματος, στο βαθμό που ομοιάζει με αυτό το πράγμα και χρησιμοποιείται ως σημείο αυτού." (Peirce 1932, 2.247). "Ένα Σύμβολο είναι ένα σημείο που αναφέρεται στο Αντικείμενο στο οποίο παραπέμπει δυνάμει ενός νόμου, συνήθως ένα συνειρμό γενικών ιδεών, ο οποίος λειτουργεί για να προκαλέσει την ερμηνεία του Συμβόλου ως αναφερόμενου σε αυτό το αντικείμενο." (Peirce 1932, 2.247). "Ένας Ενδείκτης είναι ένα σημείο που αναφέρεται στο Αντικείμενο στο οποίο παραπέμπει δυνάμει απλώς του ότι επηρεάζεται πραγματικά από αυτό το Αντικείμενο." (Peirce 1932, 2.247).

καπνός είναι ενδείκτης φωτιάς γιατί παράγεται αιτιακά από αυτή. Ο Fodor, όπως και αρκετοί άλλοι ερευνητές, βασίζόμενοι στην έννοια του ενδείκτη επιχειρούν να αναπτύξουν μια αιτιακή θεωρία για το νοητικό περιεχόμενο³⁴⁶. Αυτός είναι ο λόγος που η συγκεκριμένη θεωρία ονομάζεται, συχνά, σημασιολογία ενδεικτών (indicator semantics). Στο τρίτο κεφάλαιο ασκήσαμε μια προκαταρκτική κριτική στη θέση ότι η αιτιακή σχέση ενδεικνυόμενου – ενδείκτη μπορεί να εξηγήσει την αναφορική σχέση που χαρακτηρίζει τις αναπαραστάσεις. Σε ό,τι ακολουθεί θα επαναλάβουμε αυτή την κριτική πολύ πιο αναλυτικά και συστηματικά.

7.2.4.1.1 Χονδροειδής αιτιακή θεωρία

Σε μια πρώτη προσέγγιση, η εξήγηση που χρησιμοποιεί η σημασιολογία ενδεικτών έχει την εξής μορφή:

To x αναπαριστά το y αν και μόνον αν το y είναι η αιτία του x.

Ο Fodor ονομάζει τη θεωρία που βασίζεται σε μια τέτοιου τύπου εξήγηση "χονδροειδή αιτιακή θεωρία" (crude causal theory)³⁴⁷. Μια εξήγηση τέτοιου τύπου είναι καθαρά φυσιοκρατική, εφόσον το δεύτερο μέρος της ισοδυναμίας βασίζεται αποκλειστικά στην έννοια της αιτιότητας. Η αναπαραστασιακή σχέση ανάγεται, κατ' ουσίαν, σε μια αιτιακή σχέση. Ο αναπαραστασιακός χαρακτήρας των νοητικών καταστάσεων επιχειρείται να εξηγηθεί μέσω της ύπαρξης κάποιων αιτιακών συσχετίσεων μεταξύ των αναπαραστάσεων και του αναπαριστώμενου. Η χονδροειδής αιτιακή θεωρία δεν πρέπει να θεωρηθεί ότι αφορά αποκλειστικά κάποιους αποτραπημένους από τα εγκόσμια αναλυτικούς φιλοσόφους. Αντίθετα, πρόκειται για μια θεωρία που, σε μια πρώτη μορφή της, προϋποτίθεται από μια ολόκληρη σειρά νέων τεχνολογιών που στόχο έχουν την απεικόνιση - χαρτογράφηση των λειτουργιών του εγκεφάλου³⁴⁸.

Ας εξετάσουμε, όμως τώρα, λίγο πιο συστηματικά την χονδροειδή αιτιακή θεωρία. Καταρχάς, η αιτιακή συσχέτιση της εκάστοτε αναπαραστάσεως με ό,τι αναπαριστά πρέπει να μην είναι συμπτωματική. Με άλλα λόγια, η σύνδεση ενδεικνυόμενου - ενδείκτη οφείλει να είναι αξιόπιστη. Για παράδειγμα, η νοητική αναπαραστάση ενός αλόγου πρέπει να προκαλείται από κάθε αλόγο. Αυτό σημαίνει ότι προκειμένου να υπάρχει μια αξιόπιστη αιτιακή σχέση μεταξύ αναπαριστώμενου και

³⁴⁶ Ένα πρόβλημα για τον Fodor, το οποίο ωστόσο δεν θα μας απασχολήσει στο παρόν κεφάλαιο, είναι πώς συμβιβάζεται ο εξτερναλισμός του ως προς το περιεχόμενο με τον ατομικισμό του ως προς τις ψυχολογικές γενικεύσεις και τους αποβλεπτικούς νόμους.

³⁴⁷ Fodor 1987, σελ. 99.

³⁴⁸ Τέτοιες τεχνικές είναι η τομογραφία εκπομπής ποζιτρονίων (PET), απεικόνιση λειτουργικού μαγνητικού συντονισμού (fMRI), η MEG κλπ. Βλ. επίσης, την σύγχρονη συζήτηση περί γνωσιακών χαρτών (Dretske 1988, σελ. 62).

νοητικής αναπαράστασης πρέπει η εν λόγω σχέση να είναι νομολογική. Ο Fodor αναφέρεται σε αυτό τον τύπο αιτιακής σχέσης με τον όρο "αξιόπιστη συσχέτιση" (reliable correlation).

Σύμφωνα, λοιπόν, με την χονδροειδή αιτιακή θεωρία, μια νοητική αναπαράσταση αναφέρεται στο x , αν η νοητική αναπαράσταση παράγεται νομολογικά από το x . Το αναφερόμενο της αναπαράστασης ταυτίζεται με την αιτία της αναπαράστασης. Ακριβέστερα, κάθε έκτυπο νοητικού συμβόλου αναφέρεται στην αιτία που το προκάλεσε και κάθε τύπος νοητικού συμβόλου εκφράζει την ιδιότητα της οποίας η παρουσία αξιόπιστα προκαλεί τα έκτυπα του εκάστοτε τύπου νοητικού συμβόλου³⁴⁹.

Ωστόσο, σε αυτήν την απλή μορφή, η βασική ιδέα της σημασιολογίας ενδεικτικά παρουσιάζει προφανώς σημαντικότερα προβλήματα. Θα αναφερθούμε ενδεικτικά σε τρία από αυτά. Ένα πρώτο πρόβλημα είναι η αναφορά σε μη-υπάρχοντα αντικείμενα: σε αυτή την περίπτωση δεν υπάρχει αιτιακή σχέση μεταξύ νοητικής αναπαράστασης και αναπαριστώμενου, καθότι τα μη-υπάρχοντα αντικείμενα δεν έχουν αιτιακές δυνάμεις. Ένα δεύτερο πρόβλημα είναι η αναφορά σε μελλοντικά αντικείμενα. Και σε αυτή την περίπτωση είναι αδύνατον κάτι που δεν έχει ακόμα επέλθει να αποτελέσει την αιτία μιας παροντικής νοητικής κατάστασης. Τέλος, ένα τρίτο πρόβλημα δημιουργείται με τις νοητικές καταστάσεις που αναφέρονται σε αφηρημένες οντότητες. Για παράδειγμα, οι μαθηματικές οντότητες δεν υφίστανται εν χώρῳ και χρόνῳ, κατά συνέπεια δεν μπορούν να επενεργήσουν αιτιακά.

Το κοινό χαρακτηριστικό των τριών προβλημάτων που επισημάναμε είναι ότι οι νοητικές καταστάσεις αναφέρονται, μεταξύ άλλων, και σε οντότητες που είναι αδύνατο να έχουν αιτιακές δυνάμεις. Κατά συνέπεια, σε αυτές τις περιπτώσεις κάθε απόπειρα εξήγησης της αποβλεπτικότητας των νοητικών καταστάσεων με βάση την αιτιακή συσχέτιση αναπαράστασης - αναπαριστώμενου φαίνεται εκ προοιμίου καταδικασμένη.

Πώς απαντούν σε μια τέτοια συντριπτική κριτική οι υποστηρικτές της σημασιολογίας ενδεικτών; Η τακτική τους είναι μάλλον απλή: δέχονται ότι οι συγκεκριμένες περιπτώσεις είναι όντως προβληματικές, ωστόσο σπεύδουν να επισημάνουν ότι, ταυτόχρονα, είναι και εξαιρετικά σύνθετες. Πρωταρχικός τους στόχος είναι να αναπτύξουν μια θεωρία που να εξηγεί την αποβλεπτικότητα πολύ απλών από εξελικτική άποψη καταστάσεων, όπως η αντίληψη ορισμένων αντικειμένων³⁵⁰ ή επιθυμίες που είναι βασικές για την επιβίωση και την αναπαραγωγή. Αν η θεωρία αποδειχθεί επιτυχής για αυτές τις απλές αναπαραστασιακές καταστάσεις,

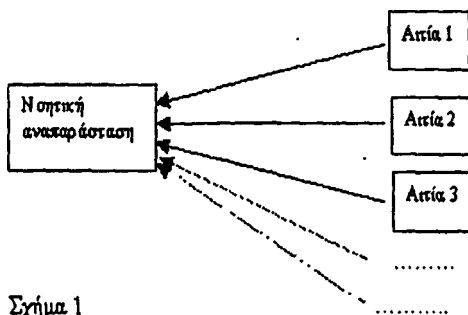
³⁴⁹ Fodor 1987, σελ. 99.

³⁵⁰ Βλ. Fodor 1991, σελ. 256.

τότε και μόνο τότε έχει νόημα να προχωρήσει κανείς στην εξέταση των παραπάνω πιο σύνθετων αναπαραστασιακών καταστάσεων³⁵¹.

Ωστόσο, ακόμα και αν περιορίσουμε το εύρος εφαρμογής της χονδροειδούς αιτιακής θεωρίας σε βασικές από εξελικτική άποψη αντιληπτικές λειτουργίες πάλι ανακύπτουν πολύ σημαντικά προβλήματα.

Ας ξεκινήσουμε με ένα παράδειγμα. Το ηλιοτρόπιο είναι ένα φυτό που έχει την ικανότητα να στρέφει τα άνθη του προς την κατεύθυνση του ηλίου. Όπως οι αυτόματοι ανιχνευτές καπνού που χρησιμοποιούνται για την εκκίνηση του συστήματος πυρόσβεσης σε κτίρια, έτσι και το ηλιοτρόπιο μπορεί να θεωρηθεί ως ένας ανιχνευτής του ήλιου. Η παρουσία του ήλιου προκαλεί την αλλαγή της κατεύθυνσης του άνθους και η κατεύθυνση του άνθους, μετά από κάθε αλλαγή της, ενδεικνύει την παρουσία του ήλιου. Ακολουθώντας την χονδροειδή αιτιακή θεωρία, θα μπορούσαμε να πούμε ότι η κατεύθυνση του άνθους αναπαριστά την αιτία που την επέφερε, δηλαδή τον ήλιο. Ωστόσο, είναι φανερό ότι τα πράγματα δεν είναι έτσι, τουλάχιστον όσον αφορά το παράδειγμα με το ηλιοτρόπιο. Καταρχάς, το ηλιοτρόπιο δεν στρέφει τα άνθη του μόνο όταν εκτεθεί στον ήλιο. Για παράδειγμα, αν κατά τη διάρκεια της νύχτας το φωτίσουμε με κάποιον προβολέα, τότε αυτό θα στρέψει τα άνθη του προς τον προβολέα. Η μεταβολή της κατεύθυνσης του άνθους μπορεί να προέλθει όχι μόνο από τον ήλιο αλλά και από άλλα σώματα που ακτινοβολούν. Πρόκειται, ουσιαστικά, για το πρόβλημα της άπειρης διάζευξης³⁵²: η αλλαγή της κατεύθυνσης του άνθους δεν ενδεικνύει τον ήλιο, αλλά μια άπειρη διάζευξη του τύπου ήλιος ή προβολέας ή (βλ. σχήμα 1).



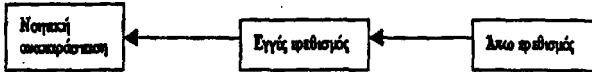
Σχήμα 1

Το πρόβλημα, λοιπόν, είναι πώς μπορούν να διακριθούν από όλες τις δυνατές αιτίες μιας αναπαραστασιακής κατάστασης εκείνες που αποτελούν μέρος των συνθηκών αληθείας της και άρα προσδιορίζουν την

³⁵¹ Σύμφωνα με τον Sterelny (1990) δεν είναι όλοι οι συγγραφείς τέτοιων θεωριών πάντα τόσο "μετριοπαθείς".

³⁵² Βλ. Fodor 1990f.

αναφορά και το περιεχόμενό της. Αν το μόνο κριτήριο στο οποίο μπορούμε να ανατρέξουμε προκειμένου να αποδώσουμε περιεχόμενο είναι η ύπαρξη κάποιας αιτιακής σχέσης, τότε μοιάζει να μην υπάρχει τρόπος να επιλέξουμε ορισμένες μόνο από τις αιτίες μιας κατάστασης ως τις συνθήκες αληθείας της.



Σχήμα 2

Ένα συγγενές πρόβλημα είναι το πρόβλημα του "βάθους"³⁵³ (βλέπε σχήμα 2) και προκύπτει από τη διάκριση ανάμεσα σε εγγύς και άνω ερεθίσματα (βλ. κεφάλαιο 2, ενότητα 2.6.1): το ηλιοτρόπιο, στην πραγματικότητα, δεν ανιχνεύει κάποια άνω ερεθίσματα όπως ο ήλιος, αλλά ανιχνεύει εγγύς ερεθίσματα, συγκεκριμένα την ύπαρξη ακτινοβολίας ορισμένου εύρους μήκους κύματος στην επιφάνειά του. Ωστόσο, δεν συμβαίνει το ίδιο και με τα αισθητηριακά μας όργανα των οποίων η εξήγησή της λειτουργίας αποτελεί, όπως επισημάναμε, τον πρωταρχικό στόχο όσων υποστηρίζουν την αιτιακή θεωρία περιεχομένου. Κατά την οπτική αντίληψη, για παράδειγμα, αυτό που βλέπουμε δεν είναι τα ερεθίσματα στον αμφιβληστροειδή, ούτε η ενεργοποίηση του οπτικού φλοιού, αλλά κάποιο αντικείμενο με τις ιδιότητές του. Αυτό καταδεικνύεται στις περιπτώσεις όπου, ενώ μεταβάλλονται τα εγγύς ερεθίσματα, η αντίληψη διατηρείται σταθερή. Για παράδειγμα, ένα αντικείμενο που στο φως της μέρας φαίνεται λευκό, εξακολουθεί να φαίνεται λευκό ακόμα και κάτω από κακές συνθήκες φωτισμού. Επίσης, ακίνητα αντικείμενα δεν εμφανίζονται να μετακινούνται όποτε μετακινούμε το βλέμμα μας ή όταν τα πλησιάζουμε τρέχοντας. Παρότι, δηλαδή, το είδωλο στον αμφιβληστροειδή μεταβάλλεται, δεν συμβαίνει το ίδιο και για το αντικείμενο που αντιλαμβανόμαστε. Άλλα παραδείγματα αποτελούν η σταθερότητα σχήματος και μεγέθους που εμφανίζει η αντίληψη μετακινούμενων αντικειμένων (βλ. Dretske 1981, σελ. 162-164). Άρα, λοιπόν, ένα σημαντικό χαρακτηριστικό της οπτικής αντίληψης (και όχι μόνο) είναι ότι δεν αναπαρίστανται τα εγγύς συμβάντα από τα οποία αιτιακά προκαλείται η εκάστοτε αντίληψη, αλλά αναπαρίστανται τα άνω συμβάντα. Ανεξάρτητα από τους λόγους για τους οποίους συμβαίνει αυτό, οποιαδήποτε θεωρία του περιεχομένου θα πρέπει να λάβει υπόψη της τον συγκεκριμένο περιορισμό. Αν, λοιπόν, συνεχίσουμε να χρησιμοποιούμε το μοντέλο του ανιχνευτή για την κατανόηση της αιτιακής θεωρίας του

³⁵³ Βλ. Sterelny (1990), σελ. 120, Jacob (1997), σελ. 66.

περιεχομένου θα πρέπει να λάβουμε υπόψη μας ότι ο απαιτούμενος ανιχνευτής οφείλει να ενδεικνύει άπω και όχι εγγύς ερεθίσματα.

Τέλος, ένα άλλο πολύ σημαντικό πρόβλημα που αντιμετωπίζει η χονδροειδής αιτιακή θεωρία είναι το πρόβλημα της λανθασμένης αναπαράστασης³⁵⁴: εφόσον η κατεύθυνση του άνθους αναπαριστά κάθε φορά την αιτία της μετακίνησης του άνθους στην παρούσα κατεύθυνση, δεν είναι δυνατή η ύπαρξη της αναπαράστασης χωρίς την ύπαρξη κάποιας αιτίας που την προκαλεί. Κατά συνέπεια, δεν είναι δυνατή η ύπαρξη κάποιας λανθασμένης αναπαράστασης, διότι σε αυτή την περίπτωση δεν υφίσταται το αναπαριστώμενο που θα λειτουργήσει ως αιτία της αναπαράστασης.

Το πρόβλημα της λανθασμένης αναπαράστασης είναι στενά συνδεδεμένο με το πρόβλημα της άπειρης διάζευξης. Το πρώτο πρόβλημα δείχνει ότι η ύπαρξη αιτιακής συσχέτισης μεταξύ του x και του y δεν αποτελεί αναγκαία συνθήκη για να είναι το y αναπαράσταση του x , εφόσον συστατικό γνώρισμα των αναπαράστασεων είναι ότι ορισμένες φορές είναι λανθασμένες, οπότε σε αυτή την περίπτωση δεν υπάρχει αιτία για να τις προκαλέσει. Το δεύτερο πρόβλημα δείχνει ότι η ύπαρξη αιτιακής συσχέτισης μεταξύ του x και του y δεν αποτελεί επαρκή συνθήκη για να είναι το y αναπαράσταση του x . Αυτό συμβαίνει επειδή η ύπαρξη της αιτιακής σχέσης δεν επαρκεί για να προσδιορίσει ποιο ακριβώς μέλος μιας άπειρης διάζευξης αναπαριστά η αναπαράσταση.

Τα δύο αυτά προβλήματα ανάγκασαν τους υποστηρικτές της σημασιολογίας ενδεικτών να απορρίψουν την χονδροειδή αιτιακή θεωρία. Οι διάφορες απόπειρες για την επίλυση των συγκεκριμένων προβλημάτων οδήγησαν στην ανάπτυξη πιο εκλεπτυσμένων θεωριών.

7.2.4.1.2 Η πληροφοριακή σημασιολογία του Dretske

Ο Dretske (1981) ήταν ο πρώτος που ασχολήθηκε με έναν συστηματικό τρόπο με την ανάπτυξη της σημασιολογίας ενδεικτών³⁵⁵. Με το βιβλίο του *Knowledge and the Flow of Information* επεχείρησε να εξηγήσει την προέλευση του νοητικού περιεχομένου βασίζόμενος στην έννοια της πληροφορίας. Συγκεκριμένα, υποστήριξε ότι μια εγκεφαλική κατάσταση καθίσταται νοητική κατάσταση αν διαθέτει μια πληροφοριακή λειτουργία. Ο Dretske θεωρεί ότι η έννοια της πληροφορίας δεν πρέπει να συγγέται με την έννοια της σημασίας (meaning), διότι η πρώτη, σε αντίθεση με τη δεύτερη, δεν προϋποθέτει την ύπαρξη κάποιου ερμηνευτή. Η πληροφορία είναι μια αντικειμενική ποσότητα που βρίσκεται παντού

³⁵⁴ Βλ. Dretske 1986.

³⁵⁵ Θα πρέπει να επισημάσουμε ότι είχε προηγηθεί ένα άρθρο του Stampel (1979) προς την ίδια κατεύθυνση. Ωστόσο το βιβλίο του Dretske, πέρα από το ότι ήταν πολύ πιο επεξεργασμένο και αναλυτικό, είχε ήδη ολοκληρωθεί το 1979.

στη φύση³⁵⁶. Ο Dretske κατανοεί την έννοια της πληροφορίας με τον ίδιο τρόπο, ουσιαστικά, που η χονδροειδής αιτιακή θεωρία κατανοεί την έννοια της αξιόπιστης συσχέτισης³⁵⁷.

Κάθε δομή που μεταφέρει πληροφορίες, μεταφέρει πάντα ένα τεράστιο αριθμό πληροφοριών. Ο Dretske υποστηρίζει ότι η διαδικασία με την οποία το πληροφοριακό περιεχόμενο μετατρέπεται σε σημασιολογικό περιεχόμενο είναι μια διαδικασία μάθησης κατά την οποία από το σύνολο του πληροφοριακού περιεχομένου μιας δομής επιλέγεται κάποιο επιμέρους τμήμα της για συγκεκριμένη χρήση: "Όπως ακριβώς εμείς συμβατικά αποδίδουμε λειτουργίες παροχής πληροφοριών σε κατασκευάσματα και όργανα, καθιστώντας έτσι τις δραστηριότητες και καταστάσεις τους - ενδείξεις οργάνων, αναβοσβήνοντα φώτα κ.λπ. - αναπαραστάσεις των συνθηκών του κόσμου για τις οποίες ενδιαφερόμαστε, έτσι και η μάθηση μετατρέπει νευρωνικές καταστάσεις που μεταφέρουν πληροφορίες - "ενδείξεις οργάνων" μέσα στο κεφάλι, θα μπορούσαμε να πούμε- σε δομές που έχουν τη λειτουργία να παρέχουν κάποιο ζωτικό τμήμα των πληροφοριών που μεταφέρουν." (Dretske 1994, σελ. 261).

Ο Dretske ασχολήθηκε ειδικότερα με δύο ειδών νοητικές καταστάσεις, τις αισθητηριακές εμπειρίες και τις πεποιθήσεις: "Οι εμπειρίες και οι πεποιθήσεις είναι απλά εκείνες οι εσωτερικές, πιθανώς φυσικές, καταστάσεις ενός συστήματος που έχουν τη λειτουργία να παρέχουν πληροφορίες (στην περίπτωση των εμπειριών) και να τις επιστρατεύουν (στην περίπτωση των πεποιθήσεων) προς χρήση στον έλεγχο της συμπεριφοράς" (Dretske 1994, σελ. 260). Οι μεν πρώτες αποκτούν το σημασιολογικό τους περιεχόμενο μέσω της φυλογενετικής επιλογής, οι δε δεύτερες μέσω της οντογενετικής μάθησης³⁵⁸. Δυο συνέπειες αυτής της θέσης - οι οποίες παρουσιάζουν ανεξάρτητο ενδιαφέρον - είναι ότι αναγκαία συνθήκη για να διαθέτει ένα σύστημα, πρώτον, εμπειρία είναι να έχει εξελιχθεί μέσω φυσικής επιλογής και, δεύτερον, πεποιθήσεις είναι να έχει τη δυνατότητα μάθησης. Ειδικότερα, η δεύτερη συνέπεια είναι ιδιαίτερα ενδιαφέρουσα γιατί αποκλείει την ύπαρξη έμφυτων πεποιθήσεων και εννοιών. Όπως εύστοχα επισημαίνει ο Dretske:

³⁵⁶ Βλ. Dretske 1981, σελ. vii: "Εν αρχή ην η πληροφορία. Η λέξη ήρθε αργότερα. Η μετάβαση επατεύθηκε μέσω της ανάπτυξης οργανισμών με την ικανότητα να εκμεταλλεθούν επιλεκτικά αυτή την πληροφορία προκειμένου να επιβιώσουν και να διαωρίσουν το είδος τους... [Ας] νοήσουμε την πληροφορία (αλλά όχι τη σημασία) ως μια αντικειμενική ποσότητα, κάτι του οποίου η παραγωγή, η μετάδοση και η λήψη δεν απαιτεί ή προϋποθέτει με οποιοδήποτε τρόπο ερμηνευτικές διαδικασίες." Το πρόγραμμα αποκωδικοποίησης του ανθρώπινου γονιδιώματος αποτελεί ίσως το καλύτερο σύγχρονο παράδειγμα αυτής της αντίληψης περί της έννοιας της πληροφορίας.

³⁵⁷ Βλ. Fodor 1987, σελ. 103.

³⁵⁸ Ibid. σελ. 261.

"δεν χρειαζόμαστε πεποιθήσεις, ούτε στην πραγματικότητα τις επικαλούμαστε, για να εξηγήσουμε αντανάκλαστικά και άλλες γενετικά καθοριζόμενες συμπεριφορές. Ακόμα και αν έχω λόγους για να κλείσω τα μάτια όταν κάποιος μου ρίχνει γροθιά, δεν τα κλείνω για αυτούς τους λόγους.... Η φυσική επιλογή μπορεί... να εξηγήσει πολλά για τη συμπεριφορά των ζώων, συμπεριλαμβανομένων και των ανθρώπων, αλλά οι συμπεριφορές που βοηθά να εξηγηθούν δεν είναι οι συμπεριφορές που εξηγούν οι πεποιθήσεις και επιθυμίες." (Dretske 1994, σελ. 262)

Μετά από αυτή τη πολύ σύντομη παρουσίαση της θεωρίας του Dretske ας εξετάσουμε το εάν επιλύει το πρόβλημα της λανθασμένης αναπαράστασης. Το συγκεκριμένο πρόβλημα είναι ένα πρόβλημα κανονιστικότητας: μια αναπαράσταση είναι λανθασμένη όταν αυτό που αναπαριστά είναι διαφορετικό από αυτό που θα *όφειλε* να αναπαριστά. Ο Dretske επεχείρησε να λύσει το πρόβλημα της λανθασμένης αναπαράστασης διακρίνοντας την περίοδο εκμάθησης μιας έννοιας από την περίοδο χρήσης της έννοιας. Σύμφωνα λοιπόν με αυτή τη διάκριση, το περιεχόμενο μιας νοητικής αναπαράστασης και, εν προκειμένω, μιας έννοιας συγκροτείται κατά την περίοδο της εκμάθησής της. Σε αυτή την περίοδο επικρατούν, σύμφωνα με τον Dretske, βέλτιστες συνθήκες που επιτρέπουν στον νέο οργανισμό να μάθει την έννοια, συχνά με την βοήθεια κάποιου διδασκάλου. Για παράδειγμα, όταν διδάσκουμε σε ένα παιδί να ξεχωρίζει τα σπουργίτια από τα υπόλοιπα πουλιά, του δείχνουμε πολλά διαφορετικά σπουργίτια από κοντινή απόσταση και κατά τέτοιο τρόπο ώστε να βλέπει τα διακριτικά τους χαρακτηριστικά. Επίσης, του δείχνουμε διαφορετικά είδη πουλιών και το ενθαρρύνουμε, υποστηρίζει ο Dretske, όταν βλέπει σπουργίτια να λέει "σπουργίτια" και όταν δεν βλέπει σπουργίτια να λέει "όχι σπουργίτια"³⁵⁹. Μετά το πέρας της εκπαίδευσης παύουν να επικρατούν ιδανικές συνθήκες και είναι δυνατόν ένας οργανισμός να δεχθεί ερεθίσματα που να προκαλέσουν την εμφάνιση μιας λανθασμένης αναπαράστασης. Αυτός είναι ο τρόπος με τον οποίο καθίσταται δυνατή η λανθασμένη αναπαράσταση. Ο Dretske συνοψίζει το επιχειρήμα του ως εξής:

"Στην κατάσταση εκμάθησης λαμβάνεται ειδική μέριμνα προκειμένου τα εισερχόμενα σήματα να έχουν μια ένταση, μια ισχύ, επαρκή για να αποδώσουν την απαιτούμενη πληροφορία στο εκπαιδευόμενο υποκείμενο.... Τέτοιες προφυλάξεις λαμβάνονται κατά την περίοδο εκμάθησης ... για να εξασφαλισθεί ότι αναπτύσσεται μια εσωτερική δομή με το κατάλληλο σημασιολογικό περιεχόμενο, μια εσωτερική δομή ... [με] την πληροφορία ότι το s είναι F. ... Αλλά από τη στιγμή που έχουμε σημασία, από τη στιγμή που το υποκείμενο έχει

³⁵⁹ Βλ. Dretske 1981, σελ 195.

αρθρώσει μια δομή που είναι επιλεκτικά ευαίσθητη στην πληροφορία ότι τα πράγματα είναι F... έκτυπα αυτού του τύπου μπορούν να προκληθούν από σήματα που σπερδύνται την κατάλληλη πληροφορία. Όταν συμβαίνει αυτό, το υποκείμενο πιστεύει ότι το s είναι F αλλά ... δεν γνωρίζει ότι το s είναι F.... Έχουμε μια περίπτωση λανθασμένης αναπαράστασης - ένα έκτυπο μιας δομής με λανθασμένο περιεχόμενο. Έχουμε, με μια λέξη, σημασία χωρίς αλήθεια." (Dretske 1981, σελ. 195)

Η απάντηση που προτείνει ο Dretske στο ερώτημα πώς είναι δυνατή η λανθασμένη αναπαράσταση παρουσιάζει ανυπέρβλητα προβλήματα γι' αυτό έχει πλέον εγκαταλειφθεί ακόμα και από τον ίδιο. Κατ' αρχάς, δεν είναι σαφές ότι υπάρχει κάποιο σημείο από το οποίο και μετά σταματά η μάθηση. Ο Dretske μοιάζει να προϋποθέτει ότι ο οργανισμός μετά το πέρας της μάθησης έχει επιτύχει μια τέλεια συσχέτιση της νοητικής αναπαράστασης με ό,τι αναπαριστά. Αυτός φαίνεται να είναι ο μόνος τρόπος να διαχωριστεί η περίοδος πριν και μετά το πέρας της μάθησης. Ωστόσο, η υπόθεση ότι η περίοδος μάθησης τερματίζεται όταν επιτευχθεί τέλεια συσχέτιση νοητικής αναπαράστασης και αναπαριστώμενου είναι μάλλον ανυπόστατη από εμπειρική άποψη.

Επίσης, σε μια τέτοια περίπτωση, σύμφωνα με τον Fodor (1987, σελ. 104), δεν λύνεται το πρόβλημα της λανθασμένης αναπαράστασης. Ας δούμε όμως το επιχειρήμα του αναλυτικά. Έστω ότι ένας οργανισμός έχει εκπαιδευτεί να δημιουργεί νοητικά έκτυπα "A" όταν εμφανίζονται συνθήκες A και ότι μετά το πέρας της εκπαίδευσης δημιουργείται κάποια στιγμή το έκτυπο νοητικής κατάστασης "A" κατά την εμφάνιση συνθηκών B. Ο Fodor θεωρεί ότι αυτό σημαίνει ότι αν εμφανίζονταν συνθήκες B πριν το πέρας της εκπαίδευσης θα προκαλούσαν και πάλι κάποιο έκτυπο του "A". Άρα, λοιπόν, η συσχέτιση που η εκπαίδευση επέβαλε δεν είναι τελικά μεταξύ του A και του "A", αλλά μεταξύ του A v B και του "A". Γι' αυτό ο Fodor (1987, σελ. 104) καταλήγει στο εξής συμπέρασμα: "Αν τα "A" συσχετίζονται με τα (A v B), τότε το περιεχόμενο ενός εκτύπου "A" είναι ότι A v B. Έτσι, ένα έκτυπο "A" που προκαλείται από κάποιο B δεν είναι λανθασμένο". Με άλλα λόγια, επανερχόμαστε στο πρόβλημα της διάζευξης.

7.2.4.1.3 Η θεωρία της ασύμμετρης εξάρτησης

Μια δεύτερη απόπειρα να εξηγηθεί η προέλευση του περιεχομένου των νοητικών αναπαραστάσεων μέσα στο πλαίσιο της σημασιολογίας ενδεικτών αποτελεί η θεωρία της ασύμμετρης εξάρτησης που ανέπτυξε ο Fodor. Η θεωρία αυτή είναι ειδικά σχεδιασμένη για να απαντά στο πρόβλημα της άπειρης διάζευξης. Αυτό που κάνει, ουσιαστικά, ο Fodor είναι να επεξεργαστεί ένα ειδικό τύπο σχέσης που τον ονομάζει ασύμμετρη εξάρτηση. Η ασύμμετρη εξάρτηση προσδιορίζεται με τη βοήθεια αντιγεγονικών συνθηκών. Ας δούμε αυτή τη σχέση με ένα παράδειγμα. Έστω ότι από ένα σύνολο άπω ερεθισμάτων μόνο όσα είναι του τύπου α αποτελούν τις συνθήκες αληθείας μιας νοητικής

αναπαράστασης #. Αυτό δεν αποκλείει το ενδεχόμενο ότι άνω ερεθίσματα του τύπου β μπορούν να προκαλέσουν την εμφάνιση της νοητικής αναπαράστασης #. Ωστόσο, σύμφωνα με την θεωρία του Fodor, τα άνω ερεθίσματα του τύπου β δεν αποτελούν μέρος των συνθηκών αληθείας της νοητικής αναπαράστασης #, διότι στην αντιγεγονική περίπτωση που τα άνω ερεθίσματα τύπου α δεν παρήγαν νοητικές αναπαραστάσεις του τύπου #, ούτε και τα άνω ερεθίσματα τύπου β θα παρήγαν αυτόν τον τύπο νοητικής αναπαράστασης. Για αυτόν το λόγο οι νοητικές αναπαραστάσεις # που παράγονται από άνω ερεθίσματα τύπου β εξαρτώνται ασύμμετρα από τις νοητικές αναπαραστάσεις # που παράγονται από άνω ερεθίσματα τύπου α³⁶⁰. Με τυπικούς όρους, η σχέση της ασύμμετρης εξάρτησης μπορεί να διατυπωθεί ως εξής:

Το σύμβολο s αναπαριστά την ιδιότητα P (και όχι την ιδιότητα Q) εάν

1. υπάρχει κάποιος νόμος που καθορίζει ότι η ύπαρξη της ιδιότητας P προκαλεί τη δημιουργία εκτύπων του συμβόλου s
2. έκτοπα του συμβόλου s δημιουργούνται ορισμένες φορές και από την ύπαρξη της ιδιότητας Q (διαφορετικής από την P)
3. όταν η ύπαρξη της ιδιότητας Q προκαλεί την δημιουργία εκτύπων του συμβόλου s , τότε η Q εξαρτάται ασύμμετρα από τον νόμο που ισχύει μεταξύ P και s . Δηλαδή, το γεγονός ότι η Q προκαλεί ορισμένες φορές το σύμβολο s εξαρτάται από το γεγονός ότι η P προκαλεί το s , ενώ, αντίθετα, το γεγονός ότι η P προκαλεί το s δεν εξαρτάται από το γεγονός ότι ορισμένες φορές το Q προκαλεί το s .

Σύμφωνα, λοιπόν, με τη σχέση ασύμμετρης εξάρτησης, μια νοητική αναπαράσταση s παριστά ένα σκύλο και όχι ένα λύκο αν

- η εμφάνιση ενός σκύλου θα προκαλούσε την αναπαράσταση s ακόμα και όταν η εμφάνιση ενός λύκου δεν προκαλούσε την s , και
- αντίστροφα, η εμφάνιση ενός λύκου δεν θα προκαλούσε την αναπαράσταση s αν η εμφάνιση ενός σκύλου δεν προκαλούσε την s .

Με άλλα λόγια, η σύνδεση λύκος - s εξαρτάται ασύμμετρα από τη σχέση σκύλος - s .

Γύρω από τη θεωρία της ασύμμετρης εξάρτησης έχει αναπτυχθεί μια αρκετά τεχνική συζήτηση με την οποία, ωστόσο, δεν θα ασχοληθούμε διότι θεωρούμε ότι το βασικό πρόβλημα που αντιμετωπίζει η συγκεκριμένη θεωρία μπορεί να εντοπιστεί με μια πολύ απλή παρατήρηση: η σχέση ασύμμετρης εξάρτησης δεν μπορεί να ορισθεί με καθαρά νομολογικούς όρους. Αν υπάρχει κάποιος νόμος που καθορίζει ότι η ύπαρξη της ιδιότητας P προκαλεί τη δημιουργία εκτύπων του συμβόλου s και κάποια στιγμή ανακαλύψουμε ότι η ύπαρξη της ιδιότητας Q προκαλεί και αυτή τη δημιουργία εκτύπων του συμβόλου s , τότε αυτό που κάνουμε

³⁶⁰ Βλ. Fodor 1991g, σελ. 91.

είναι να τροποποιούμε το νόμο, ούτως ώστε να συμπεριλάβει και τη σχέση τύπου Q - s. Κανείς δεν θα έλεγε αντίθετα, σε μια τέτοια περίπτωση, ότι ο αρχικός νόμος είναι ορθός και ότι η φύση κάνει λάθος όταν παράγει σχέσεις τύπου Q - s. Έτσι, λοιπόν, το τι είναι ορθό δεν μπορεί να οριστεί με αποκλειστικά αιτιακούς όρους. Χρειάζεται να αναφερθούμε σε κάποιο άλλο επίπεδο για να εξηγήσουμε γιατί η νοητική αναπαράσταση s είναι λανθασμένη όταν παράγεται από την Q. Το επίπεδο αυτό καθορίζει το τι θα όφειλε να αναπαριστά η s ή, αν παραμείνουμε στο πλαίσιο του Fodor, το τι θα όφειλε να αποτελεί την αιτία της s.

Εφόσον, λοιπόν, η χρήση αποκλειστικά αιτιακών σχέσεων δεν επαρκεί για να προσδιοριστεί το αν υφίσταται μια σχέση ασύμμετρης εξάρτησης, ο ορισμός που δίνει ο Fodor για την έννοια της ασύμμετρης εξάρτησης δεν είναι καθαρά αιτιακός. Το εάν η αιτιακή σχέση Q-s εξαρτάται ασύμμετρα από τη αιτιακή σχέση P-s δεν είναι κάτι που μπορεί να διαπιστωθεί ανατρέχοντας σε κάποια άλλη αιτιακή σχέση.

Ωστόσο, ο Fodor μη μπορώντας να ανατρέξει σε κάποια περαιτέρω αιτιακή σχέση για να εξηγήσει το πώς εγκαθιδρύεται η σχέση ασύμμετρης εξάρτησης, αποτυγχάνει να δώσει κάποια εξήγηση της εν λόγω σχέσης. Υπό αυτή την έννοια, η θεωρία του Fodor αποτελεί απλώς μια τυπική συνθήκη η οποία καθορίζει τις ελάχιστες προϋποθέσεις που πρέπει να πληροί μια θεωρία της προέλευσης του νοητικού περιεχομένου προκειμένου να αποφεύγει το πρόβλημα της άπειρης διάλυσης.

Στις επόμενες ενότητες θα εξετάσουμε φυσιοκρατικές θεωρίες που ανατρέχουν σε μη αιτιακές σχέσεις για να εξηγήσουν την προέλευση του νοητικού περιεχομένου. Οι σχέσεις αυτές παρότι μη αιτιακές εξακολουθούν να είναι φυσιοκρατικές υπό την έννοια ότι δεν προϋποθέτουν την έννοια της αποβλεπτικότητας.

7.2.4.2 Σημασιολογία της επιτυχίας (Success semantics)

Όλες οι θεωρίες που έχουμε μέχρι στιγμής εξετάσει ταυτίζουν το περιεχόμενο μιας νοητικής αναπαράστασης με την αιτία που την προκάλεσε. Έτσι, ως περιεχόμενο θεωρήθηκε μια κατάσταση που προηγείται της αναπαράστασης.

Μια εντελώς διαφορετική προσέγγιση είναι να θεωρήσουμε ότι το περιεχόμενο μιας νοητικής αναπαράστασης σχετίζεται με ό,τι έπεται της αναπαράστασης και, συγκεκριμένα, με τη συμπεριφορά που η παρουσία της νοητικής αναπαράστασης προκαλεί. Σύμφωνα με αυτή την πρόταση, το περιεχόμενο μιας νοητικής κατάστασης μπορεί να θεωρηθεί, σε μια πρώτη προσέγγιση, ότι σχετίζεται με τα αποτελέσματα που η παρουσία της επιφέρει, ή, με γενικότερους όρους, με την λειτουργία που επιτελεί.

Αν προσπαθήσουμε να εξειδικεύσουμε λίγο περισσότερο το ζήτημα, βλέπουμε ότι δεν μπορούμε πλέον να μιλάμε γενικώς και αορίστως για νοητικές αναπαράστασεις. Εφόσον πρόκειται για νοητικές καταστάσεις που προκαλούν συμπεριφορά, αυτές θα πρέπει να έχουν τον χαρακτήρα επιθυμιών. Αν επιθυμώ να πω νερό πηγαίνω στην βρύση και πίνω νερό. Η επιθυμία μου για νερό είναι η αιτία της συμπεριφοράς μου.

Ωστόσο, δεν αρκεί η επιθυμία και μόνο για την παραγωγή τής συμπεριφοράς. Μια πρόσθετη συνθήκη είναι το τι πιστεύω όσον αφορά το νερό. Αν, για παράδειγμα, πιστεύω ότι η βρύση δεν έχει νερό τότε δεν θα κινηθώ μέχρι τη βρύση για να ικανοποιήσω την επιθυμία μου. Μια πράξη, λοιπόν, εμπλέκει τουλάχιστον δύο ειδών νοητικές καταστάσεις: επιθυμίες και πεποιθήσεις. Πώς μπορεί, όμως, κανείς να προσδιορίσει το περιεχόμενο των δύο αυτών καταστάσεων στηριζόμενος στις πράξεις που αυτές προκαλούν; Ο J. Whyte, με δύο άρθρα του στις αρχές της δεκαετίας του '90, επεχείρησε να απαντήσει στο συγκεκριμένο ερώτημα βασιζόμενος στην έννοια της επιτυχίας ή, για την ακρίβεια, στην έννοια της επιτυχημένης πράξης³⁶¹. Γι' αυτό και ονόμασε την σημασιολογία που ανέπτυξε, σημασιολογία της επιτυχίας (success semantics). Το γεγονός ότι ο Whyte επιχειρεί να εξηγήσει το πώς αποκτούν περιεχόμενο οι πεποιθήσεις και οι επιθυμίες καθιστά την θεωρία του "μη-μετριωπαθή"³⁶². Δεν πρόκειται, δηλαδή, για κάποια θεωρία που στόχο έχει να εξηγήσει έμφυτες νοητικές καταστάσεις που απαντώνται σε απλούς οργανισμούς, αλλά προτασιακές στάσεις.

Η θεωρία του Whyte στο βαθμό που ανατρέχει σε συμβάντα εξωτερικά του οργανισμού για να προσδιορίσει το νοητικό περιεχόμενο εντάσσεται και αυτή στις εξτερναλιστικές θεωρίες.

Πότε μια πράξη μπορεί να θεωρηθεί επιτυχημένη; Ο Whyte όρισε την επιτυχία μιας πράξης με βάση την έννοια της εκπλήρωσης μιας επιθυμίας που συνδυαζόμενη με κάποια πεποίθηση προκαλεί την πράξη. Έτσι, βασιζόμενος σε μια ιδέα του Ramsey (1927), όρισε το περιεχόμενο (ήτοι, τις συνθήκες αλήθειας) μιας πεποίθησης ως εξής:

"(R) η συνθήκη αλήθειας μιας πεποίθησης είναι εκείνη η συνθήκη που εγγυάται την εκπλήρωση κάθε επιθυμίας μέσω της πράξης που θα προκαλέσει η πεποίθηση συνδυαζόμενη με την επιθυμία." (Whyte 1991, σελ. 65)

Για παράδειγμα, αν πιστεύω ότι η βρύση έχει νερό και έχω την επιθυμία να πω νερό, τότε η μετακίνησή μου μέχρι την βρύση (που προκλήθηκε από αυτές τις δύο προτασιακές στάσεις) εγγυάται την εκπλήρωση της επιθυμίας μου. Ο Whyte επιχειρεί, με άλλα λόγια, να προσδιορίσει το περιεχόμενο μιας πεποίθησης στηριζόμενος στην έννοια της εκπλήρωσης μιας επιθυμίας, δηλαδή χρησιμοποιώντας το περιεχόμενο

³⁶¹ Ο Whyte δεν ήταν ο πρώτος που ανέπτυξε αυτή την ιδέα. Ο ίδιος μνημονεύει τον Ramsey (1927) ως εμπνευστή της ιδέας. Ο Godfrey-Smith (1994) συνδέει ρητά αυτή την ιδέα με την πραγματιστική έννοια της αλήθειας. Επίσης, ο Parineau (1987) εξετάζει και απορρίπτει μια απλούστερη μορφή σημασιολογίας με βάση την έννοια της επιτυχίας. Ο Parineau αναφέρει ως ενδεχόμενους υποστηρικτές μιας τέτοιας σημασιολογίας τους Putnam (1978), Appiah (1986) και Stalnaker (1986). Βλ. επίσης Mellor (1988).

³⁶² Βλ. υποσημείωση 351.

μιας επιθυμίας. Ωστόσο, η εξήγηση του περιεχομένου μιας αποβλεπτικής κατάστασης με βάση το περιεχόμενο μιας άλλης σίγουρα δεν μας βγάζει έξω από τη σφαίρα του αποβλεπτικού. Ούτε θα αποτελούσε λύση το να ορίσουμε το περιεχόμενο μιας επιθυμίας ως εξής:

"(F) η συνθήκη εκπλήρωσης μιας επιθυμίας είναι εκείνη η συνθήκη που εγγυημένα απορρέει από κάθε πράξη που προκαλείται από αυτή την επιθυμία, εάν οι πεποιθήσεις με τις οποίες συνδυάζεται για να προκαλέσει την πράξη είναι αληθείς"
(Whyte 1991, σελ. 65)

Προφανώς, στον ορισμό (F) το περιεχόμενο μιας επιθυμίας επιχειρείται να εξηγηθεί με βάση το περιεχόμενο κάποιων πεποιθήσεων. Αν παραμείνουμε, λοιπόν, στους ορισμούς (R) και (F) είναι αδύνατον να δώσουμε μια μη αποβλεπτική εξήγηση στο περιεχόμενο είτε των πεποιθήσεων είτε των επιθυμιών. Στόχος του Whyte είναι να δείξει ότι αυτός ο "αποβλεπτικός κύκλος" μπορεί να σπάσει και, συγκεκριμένα, ότι το περιεχόμενο ορισμένων επιθυμιών μπορεί να ορισθεί κατά ένα τρόπο που δεν εξαρτάται από το περιεχόμενο των πεποιθήσεων. Αν αυτό επιτευχθεί έστω και για λίγες επιθυμίες³⁶³ τότε, σύμφωνα με τον Whyte, όλες οι υπόλοιπες επιθυμίες καθώς και οι πεποιθήσεις μπορούν να παραχθούν με διαδοχικές εφαρμογές των ορισμών (R) και (F).

Προτού, όμως, εξετάσουμε την εν λόγω πρόταση είναι σκόπιμο να εξετάσουμε ορισμένα προβλήματα που δημιουργεί ο ορισμός (R). Σύμφωνα με αυτόν, το περιεχόμενο μιας πεποίθησης μπορεί να ορισθεί με βάση το περιεχόμενο μιας επιθυμίας. Ωστόσο, δεν είναι σαφές αν ο ορισμός (R) ισχύει για όλα τα είδη πεποιθήσεων. Καταρχάς, υπάρχουν πολλές πεποιθήσεις που δεν έχουν αποτελέσει ποτέ αιτία κάποιας πράξης μας. Για παράδειγμα, η πεποίθησή μου ότι ο Σωκράτης ήταν φαλακρός, ποτέ δεν προκάλεσε κάποια πράξη μου. Μια λύση σε αυτό το πρόβλημα προτείνει ο Crane (1995, σελ. 183): "Αυτό που έχει σημασία είναι ότι είναι πάντα δυνατό να σκεφτούμε κάποια κατάσταση όπου μια πεποίθηση μπορεί να προκαλέσει μια πράξη". Για παράδειγμα, αν συμμετείχα σε ένα διαγωνισμό γνώσεων και μου ετίθετο η ερώτηση "ονομάστε έναν φαλακρό φιλόσοφο", τότε η σχετική πεποίθησή μου, σε συνδυασμό με την επιθυμία μου να νικήσω, θα αποτελούσε αιτία της πράξης μου να ονομάσω τον Σωκράτη.

Ένα δεύτερο πρόβλημα, πολύ σημαντικότερο κατά τη γνώμη μας, σχετίζεται με τη δυνατότητα εξατομίκευσης των πεποιθήσεων στην εκάστοτε πράξη. Ας επανέλθουμε στο αρχικό παράδειγμα: αν πιστεύω ότι η βρύση έχει νερό και έχω την επιθυμία να πω νερό, τότε η μετακίνησή μου μέχρι την βρύση (που προκλήθηκε από αυτές τις δύο προτασιακές στάσεις) εγγυάται την εκπλήρωση της επιθυμίας μου. Είναι, όμως, ακριβές

³⁶³ Ο Whyte τις ονομάζει "βασικές".

ότι η συγκεκριμένη πράξη προκλήθηκε από το συνδυασμό της επιθυμίας μου για νερό με μια μεμονωμένη πεποίθηση; Σίγουρα όχι! Ακόμα και αν πιστεύω, για παράδειγμα, ότι η βρύση έχει νερό αλλά, ταυτόχρονα, πιστεύω ότι η διαδρομή μέχρι τη βρύση είναι εξαιρετικά επικίνδυνη, δεν θα κινηθώ προς τη βρύση. Γενικότερα, υπάρχει μια απειρία πεποιθήσεων που ενδεχομένως εμπλέκονται στην απόφασή μου να κινηθώ προς την βρύση για να πω νερό. Κατά συνέπεια, η ικανοποίηση μιας επιθυμίας μέσω μιας επιτυχούς πράξης δεν ορίζει το περιεχόμενο κάποιας μεμονωμένης πεποίθησης, αλλά μιας σύζευξης πεποιθήσεων. Πώς είναι, λοιπόν, δυνατή η εξατομίκευση του περιεχομένου μιας μεμονωμένης πεποίθησης;

Ο Whyte θεωρεί ότι απ' όλες αυτές τις πεποιθήσεις ένα υποσύνολό τους αφορά οποιαδήποτε πράξη μας, δηλαδή ότι υπάρχουν κάποιες συνθήκες του κόσμου οι οποίες προϋποτίθενται για την επιτυχία οποιασδήποτε πράξης. Μία τέτοια συνθήκη είναι, σύμφωνα με τον Whyte, η ύπαρξη βαρύτητας. Έτσι, η ιδέα που προτείνει είναι ότι, επειδή ακριβώς οι συγκεκριμένες συνθήκες του κόσμου απαιτούνται για την επιτυχία οποιασδήποτε πράξης, γι' αυτό και μπορούν να θεωρηθούν ως δεδομένες. Κατά συνέπεια, αρκεί να εστιάσουμε την προσοχή μας στις επιλέξιμες συνθήκες που εκάστοτε πρέπει να πληρούνται προκειμένου μια επιθυμία να ικανοποιείται.

Ωστόσο, αυτή η πρόταση του Whyte δεν υποστηρίζεται από κάποιο επιχείρημα. Ακόμα και το παράδειγμα που ο ίδιος δίνει δεν είναι ορθό. Η ύπαρξη βαρύτητας, σίγουρα, δεν αποτελεί μια συνθήκη που προϋποτίθεται για κάθε επιτυχή πράξη: οι αστροναύτες του διαστημικού σταθμού MIR δεν βασίζονται τις πράξεις τους (ή τις περισσότερες από αυτές) στην πεποίθηση ότι υπάρχει βαρύτητα.

Όμως, ακόμα και αν υπήρχε ένα σύνολο πεποιθήσεων που εμπλέκονταν σε κάθε επιτυχή πράξη, πάλι θα απέμενε ένα μεγάλο σύνολο πεποιθήσεων, διαφορετικών κάθε φορά, που θα εμπλέκονταν στην εκάστοτε πράξη. Οπότε το πρόβλημα παραμένει: η ικανοποίηση μιας επιθυμίας μέσω μιας επιτυχούς πράξης δεν ορίζει το περιεχόμενο κάποιας μεμονωμένης πεποίθησης, αλλά το περιεχόμενο ενός συνόλου πεποιθήσεων. Έτσι, εφόσον σε μια πράξη εμπλέκεται όχι μια μεμονωμένη πεποίθηση αλλά ένα σύνολο πεποιθήσεων, η επιτυχία της πράξης, αν ακολουθήσουμε τη λογική του Whyte, σημαίνει ότι αληθεύει το σύνολο των πεποιθήσεων. Κατά συνέπεια, το πρόγραμμα του Whyte αποδεικνύεται ότι δεν μπορεί καν να ξεκινήσει. Ο αναγωγικός ορισμός που προτείνει δεν είναι δυνατόν να εφαρμοστεί γιατί δεν υπάρχει τρόπος εξατομίκευσης των πεποιθήσεων.

Ωστόσο, θα μπορούσε κανείς να αντιπτείνει ότι το πρόγραμμα του Whyte επιτυγχάνει, τουλάχιστον, να ορίσει το περιεχόμενο συνόλων πεποιθήσεων με βάση το περιεχόμενο κάποιων βασικών επιθυμιών. Αν, λοιπόν, καταστεί τελικά δυνατό οι βασικές επιθυμίες να ορισθούν χωρίς αναφορά σε πεποιθήσεις, θα έχει τουλάχιστον δειχθεί ότι ο αποβλεπτικός κύκλος μπορεί να σπάσει. Βέβαια, ένα τέτοιο επίτευγμα δεν θα εξηγήσει

το πώς αποκτούν περιεχόμενο μεμονωμένες πεποιθήσεις, ούτε μη βασικές επιθυμίες, ωστόσο αυτό δεν θα αναιρούσε την σπουδαιότητά του.

Το ζητούμενο, λοιπόν, είναι αν η εκπλήρωση μιας (βασικής) επιθυμίας μπορεί να ορισθεί κατά έναν μη αποβλεπτικό τρόπο. Τι θα μπορούσε, όμως, να συνιστά εκπλήρωση μιας επιθυμίας; Όταν εκπληρώνεται μια επιθυμία αυτό επιφέρει ένα συναίσθημα ικανοποίησης. Στόχος του Whyte είναι να ορίσει την έννοια της ικανοποίησης μιας επιθυμίας με μη αποβλεπτικούς όρους.

Πότε μια επιθυμία ικανοποιείται; Μια πρώτη ιδέα είναι ότι όταν μια επιθυμία ικανοποιείται εκλείπει (Whyte 1991, σελ. 66). Ωστόσο, υπάρχουν πολλοί τρόποι για να εκλείψει μια επιθυμία χωρίς να ικανοποιηθεί. Για παράδειγμα, αν κάποιος με κλωστήσει στο στομάχι τότε, ενδεχομένως, η επιθυμία μου για νερό να εκλείψει. Κανείς, όμως, δεν θα έλεγε σε αυτή την περίπτωση ότι η επιθυμία μου ικανοποιήθηκε. Μια άλλη ιδέα είναι ότι μια επιθυμία ικανοποιείται όταν δημιουργείται μια πεποίθηση ότι ικανοποιήθηκε. Για παράδειγμα, η επιθυμία μου να βρω νερό δεν θα ικανοποιηθεί παρά μόνο όταν οι αισθήσεις μου με κάνουν να πιστέψω ότι βρήκα νερό (Whyte 1991, σελ. 69). Ωστόσο, μια τέτοια εξήγηση δεν μπορεί να γίνει αποδεκτή από τον Whyte, διότι δεν καταφέρνει να ξεφύγει από τον αποβλεπτικό κύκλο: το περιεχόμενο μιας επιθυμίας επιχειρείται να εξηγηθεί με βάση το περιεχόμενο κάποιας πεποίθησης.

Ο Whyte προκειμένου να απαντήσει στο ερώτημα πότε μια επιθυμία ικανοποιείται υιοθετεί μια ιδέα του Dretske (1988). Σύμφωνα με αυτή, η εκπλήρωση μιας επιθυμίας ενισχύει, γενικά, την προδιάθεση να ενεργεί κανείς κατά τον τρόπο που οδήγησε στην εκπλήρωση της επιθυμίας³⁶⁴. Για παράδειγμα, αν ήθελα νερό και ικανοποίησα την επιθυμία μου πηγαίνοντας στη βρύση, τότε είναι πολύ πιθανό όποτε θελήσω και πάλι νερό να ξαναπάω στη βρύση. Ο Whyte θεωρεί ότι αν δεν υπήρχε η ικανοποίηση που παράγεται από την εκπλήρωση μιας επιθυμίας, δεν θα ενισχόνταν η συμπεριφοριστική προδιάθεση για την επανάληψη της πράξης που οδήγησε στην εκπλήρωση της επιθυμίας. Η ικανοποίηση μιας επιθυμίας είναι, με άλλα λόγια, ένα αποτέλεσμα που επιφέρει η εκπλήρωση της επιθυμίας³⁶⁵.

Ο Whyte ανατρέχει εδώ σε μια ιδέα που προέρχεται από την εφαρμογή της εξελικτικής εξήγησης στην οντογενετική μάθηση. Η χρήση της εξελικτικής θεωρίας για την εξήγηση της προέλευσης του νοητικού περιεχομένου αποτελεί μια αυτόνομη προσέγγιση που θα εξετασθεί στην αμέσως επόμενη ενότητα. Πάντως ακόμα και αν αποδειχθεί ορθή αυτή η προσέγγιση, δεν θα διασώσει τη σημασιολογία της επιτυχίας γιατί όπως είδαμε ο αναγωγικός ορισμός που προτείνεται από τον Whyte δεν είναι

³⁶⁴ Βλ. Whyte 1991, σελ. 67.

³⁶⁵ Βλ. Whyte 1991, σελ. 68.

δυνατόν να εφαρμοστεί καθότι δεν υπάρχει τρόπος εξατομίκευσης των πεποιθήσεων.

Η σημασιολογία της επιτυχίας από μόνη της είναι μια ημιτελής θεωρία. Κατά συνέπεια, δεν έχει νόημα να εξετάσουμε το εάν κατορθώνει να απαντήσει στα προβλήματα της λανθασμένης αναπαράστασης και της άπειρης διάζευξης. Γιατί, τότε, κάναμε τον κόπο να εξετάσουμε μια ημιτελή θεωρία; Ο λόγος είναι ότι υπάρχουν αρκετοί θεωρητικοί που αποδέχονται ως ένα επίπεδο το τρόπο που ο Whyte θέτει το πρόβλημα και διαφανούν απλώς στο τρόπο επίλυσής του. Πιο συγκεκριμένα, το πρόβλημα, όπως τίθεται, είναι η εξήγηση του περιεχομένου των νοητικών καταστάσεων με βάση τα αποτελέσματα που αυτές προκαλούν. Αυτό που γίνεται αποδεκτό από ορισμένους θεωρητικούς είναι η γενική κατεύθυνση της εξήγησης, δηλαδή ότι το περιεχόμενο των νοητικών καταστάσεων πρέπει να εξηγηθεί με βάση τα αποτελέσματα που αυτές προκαλούν, ή γενικότερα, τη λειτουργία που επιτελούν και όχι με βάση τις αιτίες που τις προκάλεσαν.

Υπάρχουν, τουλάχιστον, δύο σημεία στα οποία διαφοροποιούνται όσοι αποδέχονται αυτή τη γενική κατεύθυνση. Το πρώτο είναι ότι δεν ανατρέχουν στα τρέχοντα αποτελέσματα μιας νοητικής κατάστασης του οργανισμού, αλλά σε αποτελέσματα που είχαν παρελθούσες εμφανίσεις νοητικών καταστάσεων του ίδιου τύπου. Αυτές οι παρελθούσες εμφανίσεις μπορεί να συνβέθισαν είτε σε προγόνους του οργανισμού, είτε σε προηγούμενα στάδια της ανάπτυξής του. Όσοι δέχονται την πρώτη περίπτωση εντάσσονται στο πρόγραμμα της τελεολογικής σημασιολογίας, ενώ όσοι δέχονται τη δεύτερη συντάσσονται με μια θεωρία που υποστηρίζει ο Dretske σε πρόσφατο έργο του³⁶⁶. Το δεύτερο σημείο στο οποίο διαφοροποιούνται όσοι δέχονται την γενική κατεύθυνση ότι το περιεχόμενο μιας νοητικής κατάστασης πρέπει να εξηγηθεί με βάση τα αποτελέσματα της και όχι την αιτία της, αφορά το εύρος των φαινομένων που επιχειρούν να εξηγήσουν. Υπάρχουν, αφενός, οι μη-μετριοπαθείς θεωρίες που επιχειρούν, όπως και ο Whyte, να εξηγήσουν το σύνολο των νοητικών καταστάσεων και κυρίως τις προτασιακές στάσεις και, αφετέρου, υπάρχουν οι μετριοπαθείς θεωρίες που περιορίζονται στην εξήγηση στοιχειωδών και έμφυτων αναπαραστασιακών καταστάσεων, όπως, για παράδειγμα, η αναπαραστασιακή κατάσταση στην οποία βρίσκεται ο εγκέφαλος ενός βατράχου κατά τη διάρκεια αντίχενωσης και σύλληψης μυγών.

Ακολουθώς θα ασχοληθούμε με την τελεολογική σημασιολογία και μάλιστα στην μετριοπαθή της εκδοχή γιατί η συγκεκριμένη θεωρία παρουσιάζει τη καλύτερη συγκρότηση και τα ισχυρότερα επιχειρήματα σε σχέση με τις υπόλοιπες.

³⁶⁶ Βλ. Dretske 1988.

7.2.4.3 Τελεολογική σημασιολογία

7.2.4.3.1 Η έννοια της ίδιας λειτουργίας

Καταρχάς, θα πρέπει να διευκρινισθεί ότι ο όρος "τελεολογικός", όπως χρησιμοποιείται σε αυτό το πλαίσιο, δεν σχετίζεται με τον αντίστοιχο όρο του Αριστοτέλη. "Τελεολογικός", εν προκειμένω, δεν σημαίνει "αυτός που τείνει προς ένα σκοπό, ένα τέλος". Στην τελεολογική σημασιολογία, ο όρος "τελεολογικός" χρησιμοποιείται όπως ορίζεται στην δαρβινική θεωρία της εξέλιξης. Μέσα σε αυτό το πλαίσιο μπορεί κανείς να μιλήσει για την *τελεολογική λειτουργία* ενός οργάνου, μιας ενστικτώδους συμπεριφοράς ή ενός χαρακτηριστικού κάποιου οργανισμού. Επειδή η έννοια του τελεολογικού παραπέμπει έντονα στην αριστοτελική θεωρία είναι προτιμότερο να υιοθετήσουμε τον λιγότερο φορτισμένο όρο της "βιολογικής λειτουργίας".

Πώς μπορεί να ορισθεί, λοιπόν, η έννοια της βιολογικής λειτουργίας; Για να απαντήσουμε σε αυτό το ερώτημα θα ακολουθήσουμε την Millikan (1984, 1993a) η οποία είναι η πρώτη που εφάρμοσε τελεολογικά επιχειρήματα για την εξήγηση του περιεχομένου των νοητικών αναπαραστάσεων και έχει, ίσως, το σημαντικότερο έργο σχετικά με αυτό το ζήτημα. Η Millikan ανέλυσε την έννοια της βιολογικής λειτουργίας με βάση την έννοια της "*ιδίας λειτουργίας*"³⁶⁷ (proper function). Η "*ίδια λειτουργία*" είναι ένας τεχνικός όρος που εισήγαγε η Millikan προκειμένου να μιλήσει για "αναλογίες και δυσαναλογίες μεταξύ πραγμάτων που ανήκουν σε εντελώς διαφορετικές κατηγορίες - σωματικά όργανα, εργαλεία, σκόπιμες (purposive) συμπεριφορές, γλωσσικά στοιχεία, εσωτερικές αναπαραστάσεις, σήματα ζώων, έθιμα κλπ." (Millikan 1984, σελ. 38).

Η έννοια της ίδιας λειτουργίας μπορεί να καθοριστεί μέσω μιας ατιοκρατικής θεώρησης, όπως αυτή προκύπτει από τη δαρβινική θεωρία της εξέλιξης³⁶⁸. Σύμφωνα με αυτή, το Z είναι η ίδια λειτουργία ενός

³⁶⁷ Η ίδια λέει χαρακτηριστικά: "Σκόπευα (όπως προτάθηκε στο Millikan 1984 σελ. 2) την πρώτη σημασία που δίνει το Webster's για τον όρο 'proper' η οποία συμπίπτει με το λατινικό proprium που σημαίνει one's own." (Millikan 1998)

³⁶⁸ Βέβαια, θα μπορούσε κανείς να μιλήσει και για την ίδια λειτουργία ενός εργαλείου, εκεί όμως η πρόβλεψη της λειτουργίας οφείλεται στις προθέσεις του κατασκευαστή. Γενικότερα, θα πρέπει να τονιστεί ότι το γεγονός ότι η Millikan αναλύει την έννοια της ίδιας λειτουργίας μέσω της εξελικτικής θεωρίας δεν σημαίνει ότι δέχεται ότι κάθε ίδια λειτουργία έχει μια βιολογική καταγωγή. Η ίδια σχολιάζει χαρακτηριστικά: "Όταν μιλούσα πρόχειρα συνήθιζα να ονομάζω τις ίδιες λειτουργίες 'βιολογικές λειτουργίες', αλλά αυτό οδήγησε σε δύο ειδών παρανοήσεις. Θεωρήθηκε [πρώτον] ότι η ανάλυσή μου ήταν μια απόπειρα να αναλύσω τον τρόπο που οι βιολόγοι χρησιμοποιούν τον όρο 'λειτουργία' και [δεύτερον] θεωρήθηκε ότι όλες οι ίδιες λειτουργίες παράγονται μέσω γενετικής επιλογής. Καμιά από αυτές τις απόψεις δεν είναι σωστή." (Millikan 1999) Μια

χαρακτηριστικού X στο σύστημα S εάν και μόνο αν τα X επιλέχθηκαν για να κάνουν το Z, επειδή με το να κάνουν τα X το Z συνεισέφεραν στον πολλαπλασιασμό των προγόνων του S (Jacob 1999)³⁶⁹.

Ας δούμε ένα παράδειγμα. Η ίδια λειτουργία ενός ανθρώπινου σπερματοζωαρίου δεν ορίζεται απ' όσα θα κάνει το σπερματοζωάριο -ή, αν το λέγαμε με αποβλεπτικό λεξιλόγιο, απ' όσα *σκαπεί* να κάνει το σπερματοζωάριο - αλλά απ' όσα έκαναν σπερματοζωάρια των προγόνων μας και συνέβαλλαν στην επιβίωση και την αναπαραγωγή του ανθρώπινου είδους. Με άλλα λόγια, για τον εντοπισμό της ίδιας λειτουργίας των σπερματοζωαρίων, σύμφωνα με τη θεωρία της εξέλιξης, δεν θα πρέπει να ερευνήσουμε το τι κάνει ένα σπερματοζωάριο από τη στιγμή της δημιουργίας του και πέρα, ούτε το τι έκαναν άλλα σπερματοζωάρια που παρήχθησαν προγενέστερα κατά τη διάρκεια της ζωής ενός ανθρώπου. Για να εντοπίσουμε την ίδια λειτουργία των σπερματοζωαρίων πρέπει να εντοπίσουμε το ρόλο που αυτά έπαιζαν στους πρόγονούς μας, δηλαδή πρέπει να διερευνήσουμε γεγονότα που έλαβαν χώρα πριν από τη γένεσή μας. Γενικότερα, λουπόν, μπορούμε να παρατηρήσουμε ότι η ίδια λειτουργία καθορίζεται από την εξελικτική ιστορία και όχι από το τι συμβαίνει στο παρόν: "*Το ότι ένα όργανο ή σύστημα έχει συγκεκριμένες ίδιες λειτουργίες καθορίζεται από την ΙΣΤΟΡΙΑ του. Δεν καθορίζεται από τις παρούσες ιδιότητες, την παρούσα δομή, τις τρέχουσες προδιαθέσεις ή τις τρέχουσες λειτουργίες του.*" (Millikan 1993d, σελ. 242). Η ίδια λειτουργία της ανθρώπινης καρδιάς είναι να αντλεί το αίμα, ανεξάρτητα από το εάν κάποια συγκεκριμένη καρδιά κατορθώνει ή όχι να επιτελέσει αυτή τη λειτουργία. Ακόμα και οι δυσλειτουργικές καρδιές έχουν ως ίδια λειτουργία την άντληση του αίματος.

Ωστόσο, η έννοια της ίδιας λειτουργίας δεν πρέπει να θεωρηθεί ως μια στατιστική έννοια. Η ίδια λειτουργία των σπερματοζωαρίων δεν καθορίζεται από το τι αυτά έκαναν κατά μέσο όρο στη διάρκεια της εξελικτικής τους ιστορίας, διότι, όπως χαρακτηριστικά επισημαίνει και η Millikan, τα σπερματοζωάρια *σπανίως* επιτυγχάνουν να γονιμοποιήσουν κάποιο ωάριο. Θα μπορούσε, βέβαια, κανείς να αντιτείνει ότι αυτή η απασύνδεση της ίδιας λειτουργίας από το στατιστικό μέσο όρο ισχύει για το συγκεκριμένο παράδειγμα, αλλά σίγουρα δεν ισχύει για άλλα βιολογικά χαρακτηριστικά ή όργανα. Αν οι καρδιές δεν επιτελούσαν *ως επί το πλείστον* τη βιολογική τους λειτουργία, τότε σίγουρα δεν θα επιλέγονταν. Ωστόσο, μια τέτοια αντίρρηση δεν αναιρεί τη θέση της Millikan. Διότι αν

θέση με την οποία θα συμφωνούσε η Millikan είναι ότι η έννοια της ίδιας λειτουργίας προκύπτει ως μια γενίκευση της έννοιας της βιολογικής λειτουργίας.

³⁶⁹ Βλ. επίσης Millikan 1993d, σελ. 242: "...οι ίδιες λειτουργίες κάθε σωματικού οργάνου ή συστήματος είναι εκείνες οι λειτουργίες που βοήθησαν να εξηγηθεί η επιβίωση ή ο πολλαπλασιασμός κατά τη διάρκεια της εξελικτικής ιστορίας των ειδών που περιείχαν το όργανο ή σύστημα."

βρούμε έστω και μία περίπτωση κατά την οποία η ίδια λειτουργία δεν ταυτίζεται με τη κατά μέσο όρο "λειτουργία", η συγκεκριμένη θέση επιβεβαιώνεται.

Η ίδια λειτουργία ενός πράγματος ορίζει το τι αυτό κάνει σε "κανονικές (normal) συνθήκες", δηλαδή σε "ιστορικά βέλτιστες συνθήκες" (Millikan 1993, σελ. 87). Αυτές οι κανονικές συνθήκες δεν είναι άλλες από τις συνθήκες που επιτρέπουν την επιτυχή λειτουργία ενός πράγματος, δηλαδή τη λειτουργία που συντελεί στην επιβίωση και αναπαραγωγή του οργανισμού. Η έννοια της ίδιας λειτουργίας είναι, λοιπόν, μια κανονιστική (normative) έννοια. Η ίδια λειτουργία ενός πράγματος ορίζεται ως αυτό που "οφείλει να κάνει, και όχι [ως] αυτό που επί του παρόντος κάνει ή είναι προδιατεθειμένο να κάνει" (Rowlands 1997, σελ. 281)³⁷⁰.

Αν κανείς εξετάσει μόνο την τρέχουσα λειτουργία ενός βιολογικού οργάνου, είναι πολύ πιθανό να καταλήξει σε μια πληθώρα διαφορετικών λειτουργιών. Για παράδειγμα, εξετάζοντας την καρδιά ενός ζώου αυτή εμφανίζεται να επιτελεί, μεταξύ άλλων, τις ακόλουθες λειτουργίες: α) παράγει ρυθμικούς ήχους, β) μετατοπίζει ελαφρά το δέρμα πάνω από αυτή ανά τακτά χρονικά διαστήματα, γ) αντλεί το αίμα. Αν όμως λάβουμε υπόψη μας την εξελικτική ιστορία του ζώου στο οποίο ανήκει η καρδιά, τότε μπορούμε να περιορίσουμε το εύρος των αποδεκτών απαντήσεων. Η ίδια λειτουργία της καρδιάς είναι να αντλεί το αίμα και όχι να παράγει ρυθμικούς ήχους ή να μετατοπίζει ελαφρά το δέρμα ανά τακτά χρονικά διαστήματα, γιατί πρόγονοι του ζώου επέζησαν και αναπαρήχθησαν, μεταξύ άλλων, επειδή οι καρδιές τους αντλούσαν αίμα και όχι επειδή παρήγαν ρυθμικούς ήχους, ή μετατόπιζαν ελαφρά το δέρμα ανά τακτά χρονικά διαστήματα. Η καρδιά έχει τη συγκεκριμένη ίδια λειτουργία επειδή παρήγε πλεονεκτικά από εξελικτική άποψη αποτελέσματα. Η δαρβινική θεωρία, λοιπόν, επιτρέπει την επιλογή ορισμένων μόνο από τις δυνατές λειτουργίες ενός οργάνου ως ίδιες λειτουργίες.

Μιλώντας κάπως αυστηρότερα, στην σύγχρονη αναλυτική φιλοσοφία έχουν διακριθεί δύο έννοιες λειτουργίας. Η πρώτη είναι αυτή που μας ενδιαφέρει στην παρούσα ενότητα, δηλαδή η έννοια της ίδιας λειτουργίας. Η δεύτερη είναι η έννοια της συστημικής λειτουργίας³⁷¹. Ο Cummins (1975 σελ. 756) επισημαίνει μιλώντας για αυτού του είδους τις λειτουργίες "...η λειτουργική ανάλυση μπορεί κανονικά να διεξαχθεί στη βιολογία τελείως ανεξάρτητα από εξελικτικές θεωρήσεις: μια σύνθετη ικανότητα ενός οργανισμού (ή ενός τμήματος ή συστήματος του οργανισμού) θα μπορούσε να εξηγηθεί εάν προσφύγουμε στη λειτουργική ανάλυση ανεξάρτητα από το πώς αυτή σχετίζεται με την ικανότητα του

³⁷⁰ Βλ. επίσης McGinn (1989), σελ. 144-145.

³⁷¹ Βλ. Cummins (1975). Ο όρος "συστημική λειτουργία" παρότι δεν χρησιμοποιείται από τον ίδιο τον Cummins είναι ιδιαίτερα εύστοχος και εισάγεται από την Preston (1998).

οργανισμού να συντηρεί το είδος." Η έννοια της συστημικής λειτουργίας αφορά κάποιο υποσύστημα ενός συστήματος και ορίζεται *συγχρονικά*, δηλαδή από την τρέχουσα σχέση του υποσυστήματος με το υπόλοιπο σύστημα. Πρόκειται για την έννοια της λειτουργίας όπως κατανοείται στο πλαίσιο του λειτουργισμού. Θα επανέλθουμε στην έννοια της συστημικής λειτουργίας στην ενότητα 7.2.5.1 όπου θα εξετάσουμε την ιντερναλιστική απάντηση στο πρόβλημα της προέλευσης του νοητικού περιεχομένου.

Συνοψίζοντας τη θέση της Millikan μπορούμε να πούμε ότι η ίδια λειτουργία δεν εξετομικεύεται ανατρέχοντας στην τρέχουσα λειτουργία ούτε στις τρέχουσες προδιαθέσεις, αλλά στην εξελικτική ιστορία: "Οι βιολογικές κατηγορίες δεν περιχαρακώνονται εάν εξετάσουμε την τρέχουσα δομή, τις τρέχουσες προδιαθέσεις ή τις τρέχουσες λειτουργίες ενός οργάνου ή συστήματος που εμπίπτει στην κατηγορία, αλλά εάν εξετάσουμε (ή εικάσουμε περί) την ιστορία της." (Millikan 1993e, σελ. 55).

Η έννοια της ίδιας λειτουργίας δεν είναι φυσικά χωρίς προβλήματα. Ωστόσο, η εξέταση αυτού του ζητήματος υπερβαίνει το πλαίσιο του παρόντος κειμένου. Γι' αυτό θα αναφέρουμε απλώς ενδεικτικά δύο από τα κυριότερα προβλήματα χωρίς να εξετάσουμε τις απόπειρες επίλυσής τους. (1) Κάθε "επιλεγμένη" λειτουργία πρέπει να δεχθεί ότι συμβάλλει στην αναπαραγωγή του είδους³⁷² πράγμα που δεν έχει γίνει παρά μόνο σε ελάχιστες περιπτώσεις³⁷³. (2) Η ουσιώδης εξάρτηση της ίδιας λειτουργίας από την εξελικτική ιστορία οδηγεί σε μια ριζική αποσύνδεση της ίδιας λειτουργίας από την τρέχουσα λειτουργία, δηλαδή τη συστημική λειτουργία. Για παράδειγμα, μια τεχνητή καρδιά, δεδομένου ότι δεν έχει εξελικτική ιστορία, δεν επιτρέπεται να πούμε ότι έχει ως ίδια λειτουργία την άντληση του αίματος³⁷⁴. Ή, για να πάρουμε ένα πιο ακραίο παράδειγμα³⁷⁵, αν από κάποια κοσμική σύμπτωση ένας κεραυνός που πέφτει σε έναν βάλτο δημιουργήσει ένα ανθρωποειδές πανομοιότυπο με το ανθρώπινο είδος, τότε σύμφωνα με την θεωρία περί ίδιας λειτουργίας δεν νομιμοποιούμαστε να αποδώσουμε στα όργανα του ανθρωποειδούς τις ίδιες λειτουργίες που αποδίδουμε στα ανθρώπινα όργανα. Για την ακρίβεια, μάλλον, δεν νομιμοποιούμαστε να αποδώσουμε καμία απολύτως ίδια λειτουργία γιατί το ανθρωποειδές δεν έχει εξελικτική ιστορία. Το ίδιο ισχύει, για να δώσουμε ένα ρεαλιστικό παράδειγμα, και στην περίπτωση του πρώτου ζώου στο οποίο εμφανίστηκε κάποια

³⁷² Το οποίο προφανώς δεν ταυτίζεται απλά με την *επιβίωση* μελών του είδους, διότι η επιβίωση μπορεί κάλλιστα να μην συνοδεύεται από αναπαραγωγή.

³⁷³ Οι οποίες σίγουρα δεν αφορούσαν *αναπαραστασιακές* λειτουργίες.

³⁷⁴ Ο μόνος τρόπος να της αποδώσουμε τη συγκεκριμένη ίδια λειτουργία είναι να ανατρέξουμε στις προθέσεις του κατασκευαστή της.

³⁷⁵ Πρόκειται για ένα παράδειγμα του Davidson (1994) κατάλληλα προσαρμοσμένο στην περίπτωση των ίδιων λειτουργιών.

μετάλλαξη η οποία τελικά επιλέχτηκε. Αντίστροφα, ένα έκτυπο βιολογικού οργάνου, σύμφωνα με τη θεωρία της Millikan, διαθέτει ίδια λειτουργία ακόμα και αν δεν κατορθώσει να την επιτελέσει ούτε μια φορά. Με άλλα λόγια, η ικανότητα ή μη επιτέλεσης μιας ίδιας λειτουργίας από κάποιο όργανο δεν αποτελεί αναγκαία συνθήκη για την απόδοση της ίδιας λειτουργίας σε αυτό το όργανο.

Στην επόμενη ενότητα θα δεχτούμε ότι η έννοια της ίδιας λειτουργίας δεν παρουσιάζει ανυπέρβλητα προβλήματα και θα εξετάσουμε πώς αυτή χρησιμοποιείται για την φυσικοποίηση του νοητικού περιεχομένου. Θα εστιάσουμε το ενδιαφέρον μας στη λεγόμενη "μετριοπαθή τελεολογική σημασιολογία" διότι αυτή αποτελεί την ισχυρότερη και την πιο πειστική εκδοχή της τελεολογικής σημασιολογίας. Στην συνέχεια θα ασχοληθούμε με την επάρκεια της προτεινόμενης φυσικοποίησης και τα προβλήματα που αυτή παρουσιάζει.

7.2.4.3.2 Μετριοπαθής τελεολογική σημασιολογία

Η μετριοπαθής τελεολογική σημασιολογία επιχειρεί με βάση την έννοια της ίδιας λειτουργίας να εξηγήσει την προέλευση του περιεχομένου των στοιχειωδών και έμφυτων αναπαραστασιακών καταστάσεων. Σε μια πρώτη προσέγγιση, μπορούμε να πούμε ότι το περιεχόμενο αυτών των καταστάσεων εξισώνεται με την ίδια λειτουργία τους. Ένα ερώτημα που προκύπτει άμεσα είναι εάν κάθε είδους ίδια λειτουργία είναι ταυτόχρονα και αναπαραστασιακή λειτουργία. Κάτι τέτοιο θα οδηγούσε μάλλον σε παραδοξολογίες. Για παράδειγμα, θα μπορούσε κανείς να ισχυρισθεί ότι οι καρδιές επειδή έχουν κάποια ίδια λειτουργία, δηλαδή την άντληση αίματος, αποτελούν αναπαραστάσεις που έχουν ως περιεχόμενο την άντληση αίματος. Πρέπει, λοιπόν, να τεθεί το ερώτημα: πώς διακρίνονται οι καταστάσεις που διαθέτουν αναπαραστασιακή λειτουργία από το σύνολο όλων των καταστάσεων που διαθέτουν κάποια ίδια λειτουργία; Το ερώτημα αυτό δεν τίθεται άμεσα από τους περισσότερους υποστηρικτές της τελεολογικής σημασιολογίας. Αν εξετάσει κανείς τον τρόπο που θέτουν το ζήτημα, διαπιστώνει ότι πρώτα εντοπίζουν κάποιες βιολογικές καταστάσεις που, με προθεωρητικά κριτήρια, διαθέτουν αναπαραστασιακό περιεχόμενο και μετά επιχειρούν να το φυσικοποιήσουν ανατρέχοντας στην έννοια της ίδιας λειτουργίας. Μόνο η Millikan φαίνεται να αντιμετωπίζει άμεσα το παραπάνω ερώτημα. Συγκεκριμένα, αυτό που υποστηρίζει είναι ότι απ' όλα τα πράγματα που διαθέτουν κάποια ίδια λειτουργία, μόνο ορισμένα αποτελούν αναπαραστάσεις ή, με τη δικιά της ορολογία, "αποβλεπτικές εικόνες" (intentional icons)³⁷⁶.

Οι αποβλεπτικές εικόνες πληρούν τρεις ειδικούς περιορισμούς (Millikan 1993b, σελ. 106-107). Καταρχάς, είναι ισομορφικές με τα περιβαλλοντικά χαρακτηριστικά προς τα οποία προσαρμόζονται. Αυτό

³⁷⁶ Βλ. Millikan 1984, σελ. 95-113.

σημαίνει ότι οι αποβλεπτικές εικόνες μπορούν να μεταβληθούν ως προς ορισμένες διαστάσεις (όχι κατ' ανάγκη χωρικές) κατά τέτοιο τρόπο ώστε να παραλληλίζονται με (ή να κατοπτρίζουν) αντίστοιχες μεταβολές των περιβαλλοντικών χαρακτηριστικών που χαρτογραφούν (map). Δεύτερον, υπάρχει κάποιος ειδικός μηχανισμός μέσα στον οργανισμό του οποίου η ίδια λειτουργία είναι να παράγει τις αποβλεπτικές εικόνες. Και τρίτον, υπάρχει ένας άλλος "μηχανισμός (ή μηχανισμοί) στον οργανισμό που μπορεί να χρησιμοποιήσει ή να καθοδηγηθεί από την αποβλεπτική εικόνα στην εκτέλεση ορισμένων από τις ίδιες λειτουργίες του" (Millikan 1993b, σελ. 107). Η Millikan ονομάζει αυτό το δεύτερο είδος μηχανισμών, μηχανισμούς "κατανάλωσης" αποβλεπτικών εικόνων³⁷⁷. Οι μηχανισμοί παραγωγής και κατανάλωσης μιας αποβλεπτικής εικόνας μπορούν να βρίσκονται είτε στον ίδιο είτε σε διαφορετικούς οργανισμούς (Millikan 1993b, σελ. 107-108). Η Millikan αναφέρει ως χαρακτηριστικά παραδείγματα της δεύτερης περίπτωσης το χορό των μελισσών και τα σήματα κινδύνου των ζώων.

Χωρίς την ύπαρξη μηχανισμών "κατανάλωσης" δεν μπορεί να γίνει λόγος για αποβλεπτικές εικόνες. Για παράδειγμα, παρότι υπάρχουν μηχανισμοί στο δέρμα του χαμαιλέοντα που έχουν ως λειτουργία να τροποποιούν το χρώμα του δέρματος κατά τέτοιο τρόπο ώστε να ταιριάζει με το χρώμα του περιβάλλοντος, τα διάφορα αυτά χρώματα δεν αποτελούν αποβλεπτικές εικόνες διότι δεν υπάρχουν άλλοι μηχανισμοί στον χαμαιλέοντα ή σε άλλα ζώα που να καθοδηγούνται από τα συγκεκριμένα χρώματα (Millikan 1993b, σελ. 107)³⁷⁸. Θα μπορούσε, ωστόσο, κανείς να αντιτείνει ότι ακόμα και σε αυτή την περίπτωση χρησιμοποιούνται αποβλεπτικές εικόνες μόνο που είναι εσωτερικές στον μηχανισμό αλλαγής χρώματος, δηλαδή μεσολαβούν μεταξύ του "αισθητήρα" του χρωματικού περιβάλλοντος και του "παραγωγού" του χρώματος στο δέρμα. Όμως, η αντίρρηση αυτή δεν αντικείται τελικά στη θέση της Millikan γιατί στη

³⁷⁷ Αρχικά η Millikan (1984) είχε ονομάσει αυτούς τους μηχανισμούς "ερμηνευτικούς", αλλά στη συνέχεια εγκατέλειψε αυτό τον όρο γιατί υποδήλωνε πολύ περισσότερα απ' όσα η ίδια θεωρούσε ότι έκαναν οι εν λόγω μηχανισμοί. Η Millikan με τον μηχανισμό κατανάλωσης προτείνει ουσιαστικά κάποιον υποτυπώδη χρήστη ως αναγκαία προϋπόθεση για να λειτουργήσει μια κατάσταση του οργανισμού ως αναπαράσταση. Ως αυτό το σημείο λοιπόν, φαίνεται να συμφωνεί με τον ορισμό της αναπαράστασης που δόθηκε στο τρίτο κεφάλαιο. Ωστόσο, το κρίσιμο ερώτημα είναι τι ικανότητες αποδίδει στους μηχανισμούς κατανάλωσης και πόσο αυτές υπολείπονται από τις ικανότητες των χρηστών που περιγράφονται στο τρίτο κεφάλαιο;

³⁷⁸ Η θέση αυτή της Millikan είναι αμφισβητούμενη: οι διαφορετικοί χρωματισμοί που παίρνει το δέρμα του χαμαιλέοντα ενδέχεται τελικά να αποτελούν σήματα για άλλους χαμαιλέοντες ή άλλα ζώα. Για παράδειγμα, μπορεί να ενδεικνύουν τον έλεγχο μιας περιοχής ή να εκφράζουν επιθετικότητα. Πάντως, το ζήτημα αυτό είναι καθαρά εμπειρικό και δεν επηρεάζει τη γενικότερη θέση της Millikan.

συγκεκριμένη περίπτωση, αν όντως έτσι έχουν τα πράγματα, υπάρχει μηχανισμός κατανάλωσης και, κατά συνέπεια, υπάρχουν και αποβλεπτικές εικόνες. Αντίθετα, το χρώμα του δέρματος στο χαμαιλέοντα υποτίθεται ότι δεν χρησιμοποιείται από κανένα μηχανισμό και γι' αυτό ακριβώς το λόγο δεν αποτελεί αποβλεπτική εικόνα. Όπως επισημαίνει η Millikan (1993c, σελ. 88) "μια καλή ματιά στο τμήμα του συστήματος που συνιστά τον μηχανισμό κατανάλωσης οφείλει να είναι ό,τι ακριβώς χρειάζεται για να αναπαραστασιακό περιεχόμενο". Φυσικά, αυτή η διατύπωση δεν θα πρέπει να κατανοηθεί συγχρονικά αλλά εξελικτικά. Το αναπαραστασιακό περιεχόμενο δεν προσδιορίζεται από την τρέχουσα λειτουργία του μηχανισμού κατανάλωσης, αλλά από την κανονική, την ίδια λειτουργία του, όπως αυτή καθορίζεται από την εξελικτική ιστορία. Η Millikan (1993, σελ. 93), αναφερόμενη στις αναπαραστάσεις που χρησιμοποιούν οι άνθρωποι, λέει χαρακτηριστικά: "δεν υπάρχει τρόπος να εξεταστεί μόνο κατά την παρουσία στιγμή ένας άνθρωπος, π.χ. οι τρέχουσες ομλιακές προδιαθέσεις ή οι τρέχοντες σχηματισμοί του νευρωνικού δικτύου του, ούτε ως προς να αποκαλυφθεί έστω η αποβλεπτική φύση των εκφερόμενων προτάσεων ή των εσωτερικών αναπαραστάσεών του, πολύ περισσότερο να αποκαλυφθεί τι αυτές αναπαριστούν".

Προσφιλές παράδειγμα αποβλεπτικής εικόνας αποτελεί για τους υποστηρικτές της τελεολογικής σημασιολογίας η εγκεφαλική κατάσταση ενός βάτραχου κατά τη διάρκεια ανίχνευσης και σύλληψης μυγών. Πρόκειται για μια κατάσταση που είναι καθορισμένη φυλογενετικά, υπό την έννοια ότι ο βάτραχος δεν την αποκτά μέσω μάθησης. Τέτοιες καταστάσεις, θα μπορούσαμε να πούμε, είναι εκείνες στις οποίες βρίσκονται οι φοντοριανοί αυτόνομοι υπομηχανισμοί που περιγράφηκαν στο δεύτερο κεφάλαιο. Ωστόσο, υπάρχει τουλάχιστον μία σημαντική διαφορά: οι μηχανισμοί που εξετομικεύει η μετριοπαθής τελεολογική σημασιολογία δεν μπορεί να είναι αποκλειστικά αισθητηριακοί, οφείλουν να περιλαμβάνουν και άλλους μηχανισμούς που χρησιμοποιούν τις εξόδους των αισθητηριακών μηχανισμών, διότι η ίδια λειτουργία μιας αποβλεπτικής εικόνας προσδιορίζεται από τα πλεονεκτικά αποτελέσματα που επέφερε η χρήση της στους προγόνους του οργανισμού. Όπως επισημαίνει ο Papineau (1998, σελ. 3) συνοψίζοντας τη θέση της Millikan "ο σκοπός μιας κατάστασης δεν εξαρτάται από το ποιες συνθήκες υποτίθεται ότι την προκαλούν... αλλά μάλλον από το ποιες συνθήκες θα επατρέψουν τη χρήση της κατάστασης για την εκπλήρωση κάποιων περαιτέρω βιολογικής λειτουργίας".

Ο μηχανισμός χρήσης, ή με την ορολογία της Millikan ο μηχανισμός "κατανάλωσης", πρέπει να περιλαμβάνει και κινητικά τμήματα διότι η χρήση θα πρέπει τελικά να εκδηλώνεται με κάποιο τρόπο στην συμπεριφορά του οργανισμού και στον τρόπο που αυτός

αλληλεπιδρά με το περιβάλλον³⁷⁹. Συνοψίζοντας, μπορούμε να θεωρήσουμε ότι οι αναπαραστασιακές καταστάσεις των οποίων το περιεχόμενο επιχειρεί να εξηγήσει η μετριοπαθής τελεολογική σημασιολογία είναι εκείνες οι καταστάσεις που παράγονται από αυτόνομους αισθητηριακο-κινητικούς υπομηχανισμούς. Οι συγκεκριμένες καταστάσεις προκαλούν αυτό που προθεωρητικά ονομάζουμε ενστικτώδη αντανακλαστικά.

Η ιδιαίτερη έμφαση που δίνει η Millikan στους μηχανισμούς κατανάλωσης σηματοδοτεί τη διαφοροποίησή της από τη σημασιολογία ενδεικτών. Όπως λέει η ίδια "οι αποβλεπτικές εικόνες, ως τέτοιες ή εν γένει, δεν μεταφέρουν 'φυσική πληροφορία'. Ούτε 'συμμεταβάλλονται' ('covary') με, ούτε 'παρακόλουθούν'-'ιχνηλατούν' ('track') ό,τι εικονίζουν (icon)." (Millikan 1993b, σελ. 107). Για παράδειγμα, δεν έχει ιδιαίτερη σημασία το πόσο αξιόπιστα είναι τα σήματα κινδύνου των πιθήκων. Ενδεχομένως, οι πιθήκοι παράγουν τέτοια σήματα πολύ περισσότερες φορές από όσες είναι αναγκαίο, δηλαδή ακόμα και σε περιπτώσεις που δεν υφίσταται ο κίνδυνος τον οποίο τα σήματα ενδεικνύουν. Ωστόσο, το γεγονός ότι τα σήματα κινδύνου δεν αποτελούν αξιόπιστους ενδείκτες δεν μειώνει την χρησιμότητά τους ως αποβλεπτικών εικόνων. Μπορεί η παραγωγή αυτών των σημάτων συχνά να μην ανταποκρίνεται στις κανονικές συνθήκες στις οποίες εξελικτικά οφείλουν την επιλογή τους, ωστόσο αυτή η αστοχία δεν έχει ιδιαίτερη σημασία στο βαθμό που δεν μειώνει τις πιθανότητες επιβίωσης και αναπαραγωγής του οργανισμού: "Αντί να είναι φυσική πληροφορία, το "περιεχόμενο" μιας αποβλεπτικής εικόνας περιγράφεται από το ποιο είδος δομής ή χαρακτηριστικού θα όφειλε να βρίσκεται στο περιβάλλον του οργανισμού ούτως ώστε η εικόνα να το χαρτογραφεί μέσω του δικού της κανόνα χαρτογράφησης, προκειμένου ο καταναλωτής της να την χρησιμοποιεί επιτυχώς κατά τον κανονικό τρόπο, δηλαδή τον τρόπο που ιστορικά εξηγεί τον αλληλοσυνδεδεμένο σχεδιασμό του παραγωγού, της εικόνας και του καταναλωτή." (Millikan 1993b, σελ. 107).

Μετά από αυτή τη σύντομη ανασυγκρότηση της μετριοπαθούς τελεολογικής σημασιολογίας, ας έρθουμε τώρα να εξετάσουμε το εάν καταφέρνει τελικά να αποφυγεί τα προβλήματα που αντιμετωπίζουν οι σημασιολογίες ενδεικτών. Θα ξεκινήσουμε με το πρόβλημα της λανθασμένης αναπαραστάσης. Η Millikan θεωρεί ότι αυτό το πρόβλημα επιλύεται διότι το περιεχόμενο των αποβλεπτικών εικόνων δεν καθορίζεται από την τρέχουσα κατάσταση του εξωτερικού περιβάλλοντος αλλά από την κανονική κατάσταση. Έτσι, μια αποβλεπτική εικόνα είναι δυνατόν να είναι

³⁷⁹ Βλ. Jacob (1999): "Αυτό που επιλέχθηκε μέσω της εξέλιξης δεν ήταν όντα με έναν αισθητηριακό μηχανισμό ικανό να αναπαριστά Αυτό που επιλέχθηκε μέσω της εξέλιξης ήταν όντα S με έναν συντονισμό μεταξύ του αισθητηριακού μηχανισμού M που αναπαριστά το F και ενός κινητικού συστήματος M*".

λανθασμένη, αν δεν κατορθώσει να επιτελέσει την ίδια λειτουργία της. Αν, για παράδειγμα, η ίδια λειτουργία της είναι να ενδεικνύει την παρουσία ενός ζωφίου (δηλαδή είναι μια ενδεικτική (indicative) αποβλεπτική εικόνα³⁸⁰), τότε η αποβλεπτική εικόνα είναι λανθασμένη αν δημιουργηθεί χωρίς την παρουσία κάποιου ζωφίου. Σε αυτή την περίπτωση, η αποβλεπτική εικόνα έχει ένα περιεχόμενο που δεν ανταποκρίνεται στην τρέχουσα πραγματικότητα. Αν πάλι η ίδια λειτουργία της αποβλεπτικής εικόνας είναι να προκαλεί κάποια συγκεκριμένη συμπεριφορά (δηλαδή είναι μια προστακτική (imperative) αποβλεπτική εικόνα³⁸¹), τότε είναι λανθασμένη αν η συμπεριφορά που προκαλεί δεν είναι κανονική με την τεχνική έννοια του όρου. Για παράδειγμα, αν η ίδια λειτουργία μιας αποβλεπτικής εικόνας είναι να προκαλεί την εκτίναξη της γλώσσας ενός βιτράχου για την σύλληψη κάποιου ζωφίου, τότε η αποβλεπτική εικόνα είναι λανθασμένη αν δεν οδηγεί στη σύλληψη του ζωφίου.

Η Millikan (1984, σελ. 115) θεωρεί ότι οι πιο πρατόγονες μορφές αποβλεπτικών εικόνων -τις οποίες ονομάζει "αποβλεπτικά σήματα" - είναι ταυτόχρονα προστακτικές και ενδεικτικές. Έτσι, τα δύο παραδείγματα που δώσαμε παραπάνω δεν πρέπει να θεωρηθούν ότι στην πραγματικότητα αφορούν διαφορετικές αποβλεπτικές εικόνες: η ίδια αποβλεπτική εικόνα ταυτόχρονα ενδεικνύει την παρουσία ζωφίου και "προστάζει" την εκτίναξη της γλώσσας του βιτράχου. Αν περιοριστούμε, λοιπόν, σε μια μετριοπαθή εφαρμογή της τελεοσημασιολογίας, δηλαδή επιχειρήσουμε να εξηγήσουμε την προέλευση του περιεχομένου έμφυτων και πολύ απλών νοητικών καταστάσεων, τότε το περιεχόμενο αυτών των καταστάσεων θα πρέπει να έχει τόσο προστακτική όσο και ενδεικτική λειτουργία.

Γενικότερα, η Millikan (1996) ισχυρίζεται ότι: "Η δυνατότητα λανθασμένης αναπαράστασης προέρχεται από την δυνατότητα κάποιο έκτυπο να αποτύχει στην επιτέλεση μιας λειτουργίας η οποία εξηγούσε την συνεχιζόμενη αναπαραγωγή του τύπου του". Με άλλα λόγια, η δυνατότητα λανθασμένης αναπαράστασης οφείλεται τελικά στο γεγονός ότι η ίδια λειτουργία δεν ταυτίζεται με την τρέχουσα λειτουργία: η ίδια λειτουργία καθορίζεται από τις κανονικές συνθήκες, ενώ η τρέχουσα λειτουργία μπορεί είτε να συμφωνεί είτε να μην συμφωνεί με την ίδια λειτουργία.

Αυτή είναι σε γενικές γραμμές η απάντηση της Millikan στο πρόβλημα της λανθασμένης αναπαράστασης. Ωστόσο υπάρχουν δύο τουλάχιστον προβλήματα:

1. Αν κανείς ακολουθήσει τις συνέπειες της απάντησης που δίνει η Millikan, οδηγείται στο συμπέρασμα ότι οι αποβλεπτικές εικόνες που χρησιμοποιεί ένας οργανισμός δεν χαρτογραφούν κάποιο

³⁸⁰ Βλ. Millikan 1984, σελ. 99.

³⁸¹ Βλ. Millikan 1984, σελ. 99.

χαρακτηριστικό της τρέχουσας κατάστασης κατά τη διάρκεια της οποίας παράγονται, αλλά εκείνο τον τύπο χαρακτηριστικού ο εντοπισμός και η απεικόνιση του οποίου εξελικτικά συνετέλεσε στην αναπαραγωγή των προγόνων του οργανισμού. Το πρόβλημα με αυτό το συμπέρασμα είναι ότι έτσι αποσυνδέεται πλήρως το περιεχόμενο της αποβλεπτικής εικόνας από την τρέχουσα πραγματικότητα. Αν περιοριστεί κανείς στο εννοιολογικό "οπλοστάσιο" της τελεοσημσιολογίας, τότε μόνο εκ των υστέρων μπορεί να εντοπίσει τη σχέση των αποβλεπτικών εικόνων του οργανισμού με ό,τι συμβαίνει συγχρονικά στο περιβάλλον του οργανισμού: μια αποβλεπτική εικόνα χαρτογραφεί όντως ένα χαρακτηριστικό της τρέχουσας πραγματικότητας (δηλαδή είναι ορθή) αν η αποβλεπτική εικόνα καθοδηγεί τον μηχανισμό κατανάλωσης να επιτελέσει επιτυχώς την ίδια λειτουργία του συντελώντας έτσι στην αναπαραγωγή του οργανισμού. Γενικότερα, στην απάντηση της Millikan φαίνεται να συγχέονται δύο έννοιες περιεχομένου και αλήθειας: η εξελικτική (όπου αλήθεια είναι η αντιστοιχία με τις κανονικές συνθήκες) και η προθεωρητική (όπου αλήθεια είναι η αντιστοιχία με τις τρέχουσες συνθήκες). Ακόμα και αν οι τρέχουσες συνθήκες συμπίπτουν με τις κανονικές συνθήκες δεν είναι ορθό να πούμε ότι η αποβλεπτική εικόνα χαρτογραφεί τις τρέχουσες συνθήκες, διότι εξ ορισμού αυτό που χαρτογραφεί η αποβλεπτική εικόνα δεν καθορίζεται από την τρέχουσα σχέση της με την εξωτερική πραγματικότητα αλλά από την εξελικτική της ιστορία.

2. Για να καταλάβουμε αν υπάρχει δυνατότητα λανθασμένης αναπαράστασης θα πρέπει να ξέρουμε ποια θα ήταν η ορθή αναπαράσταση, δηλαδή ποιο θα ήταν το περιεχόμενο μιας αποβλεπτικής εικόνας όταν λειτουργεί κανονικά. Έτσι, λοιπόν, βλέπουμε ότι το πρόβλημα της λανθασμένης αναπαράστασης εξαρτάται από το πρόβλημα του προσδιορισμού του αναπαραστασιακού περιεχομένου.

Φτάνουμε έτσι στο άλλο μεγάλο πρόβλημα που, όπως είδαμε στην ενότητα 7.2.4.1, οδήγησε σε αδιέξοδο τη σηματολογία ενδεικτών: το πρόβλημα της άπειρης διάζευξης. Αυτό το πρόβλημα, όπως θα δούμε, παραμένει ουσιαστικά άλυτο και στο πλαίσιο της μετριοπαθούς τελεολογικής σηματολογίας. Ας πάρουμε το παράδειγμα των αποβλεπτικών εικόνων που χρησιμοποιεί ο αισθησιοκινητικός υπομηχανισμός του βατράχου κατά την αντίχνευση και σύλληψη μυγών. Ποιο θα μπορούσε να είναι το περιεχόμενο αυτών των αποβλεπτικών εικόνων;

Ο εντοπισμός μιας μύγας προκαλεί την εκτίναξη της γλώσσας του βατράχου και τη σύλληψη της μύγας. Αυτό το ενστικτώδες αντανακλαστικό έχει προφανώς πλεονεκτικά από εξελικτική άποψη αποτελέσματα για το είδος των βατράχων, διότι επέτρεψε σε προηγούμενες γενιές βατράχων να βρίσκουν φαγητό - γεγονός που συνετέλεσε στην επιβίωση και αναπαραγωγή τους. Θα μπορούσε, λοιπόν,

το περιεχόμενο τής υπό εξέταση αναπαραστασιακής κατάστασης να είναι "αυτο-είναι-μια-μύγα"³⁸²; Η πρόταση αυτή παρουσιάζει πολλά προβλήματα. Καταρχάς, υπάρχει ένα πρόβλημα συμβολισμού: οι παύλες μεταξύ των λέξεων έχουν στόχο να δηλώσουν ότι το εν λόγω περιεχόμενο δεν έχει σύνθετη δομή. Από την άλλη μεριά, αυτές καθ' αυτές οι λέξεις που χρησιμοποιούνται προϋποθέτουν τη δομή της κατηγορήσης, δηλαδή περιεχόμενο προτασιακού τύπου³⁸³. Ωστόσο, δεν θα επιμείνουμε σε αυτό το ζήτημα γιατί αποτελεί απλά πρόβλημα επιλογής του κατάλληλου συμβολισμού. Ενδεχομένως, ένας καταλληλότερος συμβολισμός θα ήταν, παραφράζοντας τους λογικούς εμπειριστές, "εδώ-τώρα-μύγα".

Πέρα, όμως, από το πρόβλημα του συμβολισμού, υπάρχει πρόβλημα και στο ίδιο το περιεχόμενο στο οποίο αναφέρεται η αποβλεπτική εικόνα. Το "εδώ-τώρα-μύγα" ονομάζει την *απία* της αναπαραστασιακής κατάστασης του βατράχου και όχι το *αποτέλεσμα* που αυτή επιφέρει. Κατ' αυτό τον τρόπο, όμως, δεν επιχειρείται να προσδιοριστεί το περιεχόμενο της κατάστασης με βάση την ίδια λειτουργία της, αλλά την αιτία της, δηλαδή ξαναγυρνάμε στη λογική της σημασιολογίας ενδεικτών. Χαρακτηριστικό παράδειγμα είναι ο Sterelny (1990, σελ. 127) ο οποίος, παρά το ότι δηλώνει υποστηρικτής της τελεολογικής σημασιολογίας, ισχυρίζεται ότι η ίδια λειτουργία της αναπαραστασιακής κατάστασης του βατράχου είναι να ενδεικνύει την παρουσία μυγών³⁸⁴. Έτσι, αναγκάζεται να περιοριστεί στο αισθητηριακό τμήμα του υπό εξέταση εγκεφαλικού υπομηχανισμού, αλλά το σημαντικότερο και πιο προβληματικό είναι ότι προσδιορίζει την ίδια λειτουργία με βάση τις αιτίες που προκαλούν μια αναπαραστασιακή κατάσταση και όχι με βάση τα αποτελέσματα της αυτή προκαλεί. Όπως επισημαίνει και ο Papineau (1998, σελ. 6), σχολιάζοντας την αντίστοιχη με του Sterelny λογική της Neander (1995): "η ανίχνευση μικρών, μαύρων, κινούμενων αντικειμένων' δεν φαίνεται να αποτελεί με κανένα τρόπο κάποιο *αποτέλεσμα*. Τα ακόλουθα, σίγουρα, είναι αποτελέσματα του συστήματος ανίχνευσης στους βατράχους: ο βάτραχος εκτινάσσει τη γλώσσα του· συλλαμβάνει ένα μικρό μαύρο κινούμενο πράγμα· συλλαμβάνει μια μύγα· βρίσκει φαγητό κλπ.... Σε κάθε περίπτωση 'η ανίχνευση του τάδε' δεν απαντάται μεταξύ των αποτελεσμάτων". Εφόσον, λοιπόν, το "εδώ-τώρα-μύγα" ή το "εδώ-τώρα-μικρό-μαύρο-κινούμενο-

³⁸² Όπως υποστηρίζει ο Sterelny 1990, σελ. 125.

³⁸³ Η παρατήρησή μας δεν σημαίνει ότι δεν υπάρχουν ερευνητές που αποδίδουν προτασιακό περιεχόμενο στις αναπαραστασιακές καταστάσεις του βατράχου. Απλώς στην παρούσα ενότητα ενδιαφερόμαστε αποκλειστικά για τη χρήση της τελεολογικής σημασιολογίας στην εξήγηση στοιχειωδών και έμφυτων αναπαραστασιακών καταστάσεων.

³⁸⁴ Στην ίδια λογική εντάσσεται και η Neander (1995, σελ. 129-130), η οποία θεωρεί ότι η ίδια λειτουργία τής υπό εξέταση αναπαραστασιακής κατάστασης του βατράχου είναι η "ανίχνευση μικρών, μαύρων, κινούμενων πραγμάτων".

πράγμα" δεν συνιστούν αποτελέσματα δεν τίθεται καν θέμα επιλογής τους μέσω της εξελικτικής διαδικασίας³⁸⁵.

Ποιο θα μπορούσε να είναι, τότε, το περιεχόμενο της αναπαραστασιακής κατάστασης του βατράχου όταν εντοπίζει και συλλαμβάνει μία μύγα; Το πρόβλημα είναι ότι υπάρχει μια σειρά αποτελεσμάτων που προκαλούνται από την παρουσία της αναπαραστασιακής κατάστασης: εκτίναξη γλώσσας, σύλληψη μύγας, εύρεση φαγητού, ενίσχυση του οργανισμού κλπ. Ποιο από όλα αυτά αποτελεί το περιεχόμενο της κατάστασης; Μήπως το περιεχόμενο είναι η διάλυση όλων των παραπάνω αποτελεσμάτων; Σε αυτή την περίπτωση οδηγούμαστε σε μια πιο σύνθετη μορφή του προβλήματος της άπειρης διάλυσης³⁸⁶, διότι δεν έχουμε απλώς να προσδιορίσουμε αν το περιεχόμενο είναι σύλληψη μύγας ή σύλληψη κουνουπιού ή σύλληψη μικρού μαύρου κινούμενου πράγματος ή κλπ, αλλά το εάν καταρχάς αυτό έχει να κάνει με την ενέργεια της σύλληψης ή με ενέργειες που προηγούνται αυτής ή με ενέργειες που έπονται αυτής. Αν, λοιπόν, η τελεολογική σημασιολογία δεν κατορθώσει να απαντήσει σε αυτό το γενικευμένο πρόβλημα της άπειρης διάλυσης, το όλο πρόγραμμα φαίνεται να οδηγείται σε αποτυχία γιατί το αναπαραστασιακό περιεχόμενο των αποβλεπτικών εικόνων καθίσταται εντελώς απροσδιόριστο.

Κλείνοντας με το πρόγραμμα της τελεολογικής σημασιολογίας θα θέλαμε να επισημάνουμε ότι αυτό είναι εγκλωβισμένο σε μια εξαιρετικά αμφισβητήσιμη αντίληψη όσον αφορά το είδος του αναπαραστασιακού περιεχομένου που υποτίθεται ότι χρησιμοποιούν οι βιολογικοί οργανισμοί, και ειδικότερα τα λιγότερο εξελιγμένα ζώα³⁸⁷. Η αντίληψη αυτή προϋποθέτει μια αντικειμενοκρατική οντολογία, δηλαδή μια οντολογία αντικειμένων και ιδιοτήτων³⁸⁸. Ωστόσο, δεν είναι καθόλου αυτονόητο ότι ο κόσμος στον οποίο έχει πρόσβαση ένα ζώο είναι κόσμος αντικειμένων και ιδιοτήτων. Όπως επισημαίνει και η Akins (1996) εξετάζοντας το ρόλο των αισθητήρων θερμού και ψυχρού: "Αυτό για το οποίο ανησυχεί ο οργανισμός, στις καλύτερες των ναρκισιστικών παραδόσεων, είναι η ίδια η δικιά του άνεση. Το σύστημα δεν ερωτά "πώς είναι εκεί έξω;" -μια ερώτηση για τις αντικειμενικές θερμοκρασιακές καταστάσεις που έχει το

³⁸⁵ Επίσης, η ίδια η Millikan (1993, σελ. 126) επισημαίνει: "το πρόβλημα της απροσδιοριστίας [δηλαδή το πρόβλημα της άπειρης διάλυσης] που ο Dretske και ο Fodor αντιμετώπισαν επιλύεται με την προσεκτική εξέταση του ρόλου που παίζει ο καταναλωτής των αναπαραστάσεων".

³⁸⁶ Για την ακρίβεια πρόκειται για έναν συνδυασμό του προβλήματος της άπειρης διάλυσης με το πρόβλημα του βάθους.

³⁸⁷ Το ίδιο ισχύει φυσικά και για το πρόγραμμα της σημασιολογίας ενδεικτών.

³⁸⁸ Πιο συγκεκριμένα, αντικειμενοκρατική είναι η οντολογία που βασίζεται στην οντολογία της φυσικής -τουλάχιστον όπως αυτή κατανοείται από τους απλούς ανθρώπους και ορισμένους φιλοσόφους. Σε μια τέτοια οντολογία περιλαμβάνονται εκτός από αντικείμενα και ιδιότητες, συμβάντα, καταστάσεις και διαδικασίες.

δέρμα του σώματος". Η Akins, για να υποστηρίξει τη θέση της, ξεκινά από την εξέταση μιας γνωστής ψευδαισθήσης: όταν βάλουμε τα χέρια μας σε ένα δοχείο με χλιαρό νερό αφού προηγουμένως είχαμε βουτήξει το ένα σε κρύο και το άλλο σε ζεστό νερό, αισθανόμαστε διαφορετική θερμοκρασία στα δύο χέρια μας. Πρόκειται άραγε για ένα λάθος που κάνουν οι αισθητήρες μας επειδή είναι ατελείς; Η Akins θεωρεί πως όχι. Οι συγκεκριμένοι αισθητήρες θα μας φανούν λιγότερο "χαζοί" αν τους προσεγγίσουμε υπό ένα διαφορετικό πρίσμα, δηλαδή "αν κανείς σκεφτεί [τους αισθητήρες] ότι λένε πως ένα ερέθισμα επηρεάζει το δέρμα - ότι [δηλαδή] το ένα χέρι ψύχεται απότομα ενώ το άλλο θερμαίνεται απότομα. Βλάβη του δέρματος μπορεί να συμβεί από ακραίες θερμοκρασίες (από κάψιμο ή πάγωμα), αλλά επίσης και από [θερμοκρασιακές] μεταβολές που συμβαίνουν εντός του εύρους [που είναι ανεκτό από] το δέρμα"³⁸⁹. Με άλλα λόγια, όσον αφορά τους αισθητήρες θερμισμού και νυχτρού οδηγούμαστε στην λανθασμένη άποψη ότι αυτοί κάνουν λάθος στην περίπτωση που περιγράφηκε γιατί προϋποθέτουμε λανθασμένα ότι αποστολή των συγκεκριμένων αισθητήρων είναι η αντίχρευση της αντικειμενικής θερμοκρασίας. Γενικεύοντας σε αυτή τη λογική μπορούμε να συμφωνήσουμε με τον Stich (1990, σελ. 62), ότι "η φυσική επιλογή δεν ενδιαφέρεται για την αλήθεια: ενδιαφέρεται για την αναπαραγωγική επιτυχία". Κατά συνέπεια, αποτελεί αδιέξοδο δρόμο το να επιχειρεί κανείς να αναπτύξει μια θεωρία αλήθειας (και περιεχομένου) βασιζόμενος στην έννοια της αναπαραγωγικής επιτυχίας.

Με την τελεολογική σημασιολογία ολοκληρώνουμε την εξέταση των εξτερναλιστικών προσεγγίσεων για την προέλευση του νοητικού περιεχομένου. Στην επόμενη ενότητα θα εξετάσουμε τις ιντερναλιστικές προσεγγίσεις.

7.2.5 Ιντερναλιστικές προσεγγίσεις

Αφετηρία των ιντερναλιστικών προσεγγίσεων, όπως επισημάναμε στην ενότητα 7.2.3, αποτέλεσε η έννοια του στενού περιεχομένου. Στο νοητικό πείραμα του Putnam στενό περιεχόμενο είναι εκείνο το "τμήμα" του νοητικού περιεχομένου που παραμένει κοινό μεταξύ των δύο διδύμων όταν αναφέρονται σε διαφορετικές αλλά φαινομενικά μη διακρίσιμες ουσίες. Το στενό περιεχόμενο, εν προκειμένω, είναι ο κοινός τρήπος παρουσίας δύο διαφορετικών αλλά φαινομενικά μη διακρίσιμων ουσιών. Γι' αυτό το λόγο το στενό περιεχόμενο προτείνεται ως μια προσέγγιση του φρεγκεανού νοήματος.

Οι περισσότεροι από τους υποστηρικτές των ιντερναλιστικών προσεγγίσεων δεν αρνούνται ότι το στενό περιεχόμενο δεν εξαντλεί ό,τι προθεωρητικά ονομάζουμε νοητικό περιεχόμενο, αλλά εστιάζουν την προσοχή τους σε αυτό το "τμήμα" του νοητικού περιεχομένου γιατί

³⁸⁹ Akins 1996, σελ. 349.

θεωρούν ότι μόνο το στενό περιεχόμενο αφορά την ψυχολογία. Μόνο ό,τι είναι "μέσα στο κεφάλι" αποτελεί αιτία της συμπεριφοράς του υποκειμένου: Αν δύο διαφορετικές ουσίες φαίνονται ίδιες σε ένα υποκείμενο, τότε θα του προκαλέσουν το ίδιο σύνολο συμπεριφορών. Επίσης, όπως επισημαίνει ο Recanati (1994, σελ. 213) υπάρχει μια πολύ ισχυρή καρτεσιανή διαίσθηση που υποστηρίζει την έννοια του στενού περιεχομένου: "Διαφορετικές αιτίες μπορούν να παραγάγουν τα ίδια αποτελέσματα στα αισθητηριακά μας όργανα και, μέσω αυτών των αποτελεσμάτων, μπορούν να επιφέρουν τις ίδιες υποκειμενικές εμπειρίες σε μας. Από την οπτική του υποκειμένου ο τρόπος που αυτό βιώνει τον κόσμο είναι ο ίδιος όταν αντιλαμβάνεται ένα μήλο μπροστά του (και πιστεύει ότι είναι πράσινο) και όταν απλώς έχει παραισθήσεις ενός ποιοτικά ταυτόσημου μήλου ως ευρισκόμενου μπροστά του (και πιστεύει ότι είναι πράσινο): και στις δύο περιπτώσεις φαίνεται στο υποκείμενο ότι αντιλαμβάνεται ένα συγκεκριμένο μήλο"³⁹⁰.

Ένας τρόπος να διακρίνει κανείς το στενό από το ευρύ περιεχόμενο είναι να εξετάσει παραδείγματα ενδεικτικών σκέψεων (indexical thoughts). Αν ο Α και ο Β βλέποντας ένα διαφορετικό αυτοκίνητο κάνουν τη σκέψη 'Αυτό το αυτοκίνητο είναι πράσινο' [(1)], τότε η (1) ως σκέψη του Α έχει διαφορετικό ευρύ περιεχόμενο απ' ό,τι ως σκέψη του Β, γιατί το αυτοκίνητο που βλέπει ο Α είναι διαφορετικό από το αυτοκίνητο που βλέπει ο Β. Ωστόσο, υπάρχει αναμφίβολα και κάτι κοινό ανάμεσα στις δύο αυτές σκέψεις το οποίο συλλαμβάνεται με την έννοια του στενού περιεχομένου³⁹¹.

Ένας πρόσθετος λόγος για την αποδοχή της έννοιας του στενού περιεχομένου είναι ότι αν παραμείνει κανείς στην σχέση νοητικής αναπαράστασης - κόσμος δεν μπορεί να εξηγήσει τη διαφορά περιεχομένων που υπάρχει μεταξύ εννοιών που αναφέρονται στο ίδιο αντικείμενο. Πρόκειται, ουσιαστικά, για το γνωστό πρόβλημα που ανέδειξε ο Frege με το παράδειγμα του Αυγερινού και του Αποσπερίτη, όπως αυτό μεταφέρεται στην φιλοσοφία του νου. Στην προκειμένη περίπτωση η έννοια του στενού περιεχομένου προσφέρει τη δυνατότητα μιας λεπτότερης εξατομίκευσης του νοητικού περιεχομένου.

Έτσι, λοιπόν, ο στόχος που θέτουν οι ιντερναλιστές είναι να εξηγήσουν την προέλευση του στενού περιεχομένου των νοητικών καταστάσεων, με ποιο τρόπο δηλαδή μια νοητική αναπαράσταση αποκτά στενό περιεχόμενο. Στα επόμενα θα εξετάσουμε τη σημαντικότερη ιντερναλιστική θεωρία φυσικοποίησης του στενού περιεχομένου, την σημασιολογία εννοιολογικού ρόλου.

³⁹⁰ Η άποψη αυτή θα αμφισβητηθεί στο όγδοο κεφάλαιο, ενότητα 8.2.2.

³⁹¹ Βλ., επίσης, Block 1994, σελ. 84-85.

7.2.5.1 Σημασιολογία εννοιολογικού ρόλου

Πρόκειται για μια ολιστική θεωρία νόηματος η οποία αρθρώνεται γύρω από την ιδέα ότι το νόημα μιας νοητικής αναπαράστασης προκύπτει από τις σχέσεις της με τις εισόδους, τις εξόδους και τις υπόλοιπες νοητικές αναπαραστάσεις. Η σημασιολογία εννοιολογικού ρόλου (ΣΕΡ) αποτελεί μια ντερναλιστική θεωρία για την προέλευση του (στενού) περιεχομένου των νοητικών αναπαραστάσεων γιατί κάθε νοητική αναπαράσταση αποκτά το περιεχόμενό της από την λειτουργία που αυτή έχει μέσα στο γνωσιακό σύστημα του οργανισμού, δηλαδή από τις σχέσεις της με τις υπόλοιπες νοητικές αναπαραστάσεις. Με άλλα λόγια, η ΣΕΡ βασίζεται στην έννοια της συστημικής λειτουργίας (βλ. ενότητα 7.2.4.3.1).

Βασικοί υποστηρικτές της ΣΕΡ είναι ο Field (1978), ο Loar (1981), ο McGinn (1982), ο Block (1986), οι LePore & Loewer (1987), και ο Peacocke (1992). Επίσης στον χώρο της ψυχολογίας έχει αναπτυχθεί ανεξάρτητα η λεγόμενη διαδικαστική σημασιολογία (procedural semantics) (Woods 1981, Miller & Johnson-Laird 1976) η οποία βασίζεται στην ίδια αρχή με την ΣΕΡ.

Ο Block θεωρεί ότι η ΣΕΡ αποτελεί ουσιαστικά μια επέκταση της γνωστής από την φιλοσοφία της γλώσσας θεωρίας της σημασίας ως χρήσης (Wittgenstein 1953), σύμφωνα με την οποία η σημασία μιας λέξης είναι η χρήση της. Η ΣΕΡ, αντί για την εξωτερική χρήση, δηλαδή την χρήση εξωτερικών αναπαραστάσεων για την επικοινωνία και, γενικότερα, την κοινωνική αλληλεπίδραση, ανατρέπει στην *εσωτερική* χρήση αναπαραστάσεων εντός του εγκεφάλου (ή του υπολογιστή). Ωστόσο, αυτή η συσχέτιση είναι μάλλον παραπλανητική γιατί η χρήση στην ΣΕΡ ορίζεται καθαρά με συντακτικούς όρους δηλαδή, με το ποιες συντακτικές μορφές σχετίζονται με ποιες άλλες συντακτικές μορφές. Αντίθετα, στον Wittgenstein η έννοια της χρήσης σίγουρα δεν εξαντλείται στο συντακτικό επίπεδο. Αφορά και το εξωγλωσσικό (με τη στενή έννοια του όρου) πλαίσιο όπως αυτό συγκροτείται από τους θεσμούς, τις πρακτικές και τις κανόνες των χρηστών³⁹². Η ΣΕΡ θεωρείται, επίσης, ότι συγγενεύει και με τον συναγωγισμό (inferentialism) - μια επέκταση της αντίληψης του Wittgenstein με βάση την έννοια της συναγωγής (Sellars 1963) - γι' αυτό και συχνά ονομάζεται "σημασιολογία συναγωγικού ρόλου"³⁹³.

³⁹² Βλ. κεφάλαιο 3, ενότητα 3.3.1.3.2.

³⁹³ Ο Putnam (1988, σελ. 47) αναφέρει χαρακτηριστικά: "Ενδεχομένως, η σαφέστερη πηγή της έννοιας [του εννοιολογικού ρόλου] είναι το έργο του Wilfrid Sellars, ο οποίος περιγράφει τη γλώσσα ως έχουσα 'κανόνες εισόδου στη γλώσσα' (σκέψου αυτούς ως κανόνες που λένε ότι όταν ο ομιλητής έχει συγκεκριμένες εμπειρίες θα βάζει συγκεκριμένες γλωσσικές προτάσεις στο 'κουτί των πεποιθήσεων'), κανόνες 'από τη γλώσσα στη γλώσσα' (κανόνες που όταν ο ομιλητής αποδέχεται ορισμένες γλωσσικές προτάσεις τότε αποδέχεται και συγκεκριμένες άλλες) και 'κανόνες εξόδου από τη γλώσσα' (κανόνες που λένε ότι όταν ο ομιλητής έχει συγκεκριμένες γλωσσικές προτάσεις στα κουτιά των

Όμως, ακόμα πιο προφανής είναι η συγγένεια της ΣΕΡ με τον λειτουργισμό. Όπως έχουμε ήδη αναφέρει (βλ. Κεφάλαιο 1), ο λειτουργισμός συκροείται γύρω από τη θέση ότι οι νοητικές καταστάσεις εξατομικεύονται με βάση το λειτουργικό τους ρόλο, δηλαδή τις λειτουργικές τους σχέσεις με τις υπόλοιπες νοητικές καταστάσεις, τις εισόδους και τις εξόδους. Υπ' αυτή την έννοια, η ΣΕΡ δεν είναι τίποτα άλλο από μια εφαρμογή του λειτουργισμού στον χώρο των νοητικών αναπαραστάσεων. Για το λόγο αυτό, η ΣΕΡ είναι γνωστή και με το όνομα "σημασιολογία λειτουργικού ρόλου". Δεδομένου δε ότι η συγκεκριμένη θεωρία έχει ως στόχο την φυσικοποίηση του νοητικού περιεχομένου και, κατά συνέπεια, δεν μπορεί να βασίζεται σε σημασιολογικούς όρους για αυτή τη φυσικοποίηση, ο όρος "λειτουργικός ρόλος" είναι περισσότερο εύστοχος από τον όρο "εννοιολογικός ρόλος" ο οποίος έχει αναπόφευκτα σημασιολογικές συνδηλώσεις.

Ας δούμε τώρα συγκεκριμένα παραδείγματα που προτείνονται για την υποστήριξη της ΣΕΡ. Ένα πρώτο παράδειγμα αποτελούν οι λογικοί σύνδεσμοι. Η ΣΕΡ φαίνεται να εξηγεί ικανοποιητικά το πώς αποκτούν το περιεχόμενό τους οι λογικοί σύνδεσμοι. Ο Peacocke (1992, σελ. 6), συγκεκριμένα, αναφερόμενος στην έννοια της σύζευξης (Σ) προσδιορίζει την κατοχή αυτής της έννοιας ως εξής:

"ένας στοχαστής θα πρέπει να βρίσκει μεταβάσεις της μορφής

$$\frac{p}{\frac{q}{p\Sigma q}} \quad \frac{p\Sigma q}{p} \quad \frac{p\Sigma q}{q}$$

ως πρωτογενώς αναγκαστικές και ο λόγος που τις βρίσκει πρωτογενώς αναγκαστικές οφείλεται αποκλειστικά στη μορφή τους"³⁹⁴.

Ένα δεύτερο παράδειγμα που συχνά αναφέρεται προς υποστήριξη της ΣΕΡ είναι το περιεχόμενο εννοιών που αναφέρονται σε θεωρητικές οντότητες. Δεδομένου ότι τέτοιες οντότητες δεν μπορούν να δοθούν αντιληπτικά, οποιαδήποτε προσπάθεια προσδιορισμού του περιεχομένου των αντίστοιχων εννοιών με εξτερναλιστικό τρόπο είναι καταδικασμένη σε αποτυχία. Αντ' αυτού προτείνεται ότι το περιεχόμενο των εν λόγω εννοιών προσδιορίζεται από τη σχέση που έχουν μεταξύ τους μέσα στο γνωσιακό σύστημα. Στην ουσία αυτή η πρόταση, όπως επισημαίνει ο Cummins (1996, σελ. 30), αποτελεί υιοθέτηση της γραμμής που ακολούθησαν οι εμπειριστές (βλ. Hempel 1950) όσον αφορά την σημασία των

πεποιθήσεων και των επιθυμιών, τότε επιτελεί συγκεκριμένες σωματικές κινήσεις, ή λέει συγκεκριμένες λέξεις κ.λπ.)"

³⁹⁴ Ωστόσο, όπως επισημαίνει ο Rey (1997, σελ. 239) η εφαρμογή αυτής της θεωρήσης σε άλλους λογικούς συνδέσμους δημιουργεί ανυπέβλητα προβλήματα: "Σκεφτείτε απλώς το 'ή', ή τις διαμάχες γύρω από το νόμο του αποκλεισμένου τρίτου: οι αντίπαλοι σε αυτή τη διαμάχη εννοούν, άραγε, κάτι διαφορετικό με το 'ή'; Πώς μπορούν τότε στην πραγματικότητα να διαφωνούν;"

επιστημονικών θεωρητικών όρων: "Ακριβώς όπως οι εμπειριστές ήθελαν να ταυτίσουν τη σημασία ενός θεωρητικού όρου με τον εννοιολογικό ρόλο που παίζει στην οικεία θεωρία, έτσι και η ΣΕΡ ταυτίζει τη σημασία μιας νοητικής αναπαράστασης με τον εννοιολογικό ρόλο που παίζει σε ένα είδος 'αυτόματης θεωρίας', δηλαδή σε ένα γνωστικό σύστημα". Με άλλα λόγια, αυτό που συμβαίνει, σύμφωνα με τη ΣΕΡ, είναι κάτι αντίστοιχο με την περίπτωση τού προσδιορισμού εννοιών μιας επιστημονικής θεωρίας. Για παράδειγμα, στο πλαίσιο της νευτώνειας φυσικής ο όρος "δύναμη" δεν μπορεί να ορισθεί παρά μόνο σε σχέση με τους υπόλοιπους όρους της θεωρίας, όπως "μάζα", "ορμή", "ενέργεια" κ.λπ.

Ας δούμε όμως τώρα πιο συγκεκριμένα τις βασικές θέσεις και παραδοχές γύρω από τις οποίες αρθρώνεται η ΣΕΡ.

Κατ' αρχάς, η ΣΕΡ δέχεται ότι οι διάφορες νοητικές καταστάσεις και ειδικότερα οι προτασιακές στάσεις ταυτίζονται με λειτουργικές - υπολογιστικές καταστάσεις του οργανισμού. Η ταυτότητα για την οποία γίνεται εδώ λόγος είναι θεωρητική ταυτότητα. Οι προτασιακές στάσεις εκλαμβάνονται ως έννοιες μιας θεωρίας -της ψυχολογίας του κοινού νου- που μπορούν να αναχθούν σε έννοιες υπολογιστικών καταστάσεων μιας υπολογιστικής θεωρίας του νου. Αυτό, συγκεκριμένα, σημαίνει ότι οι νόμοι της ψυχολογίας του κοινού νου μπορούν να παραχθούν από τους νόμους της υπολογιστικής θεωρίας του νου. Η ΣΕΡ, λοιπόν, βασίζεται στην αποδοχή μιας ταυτότητας τύπου μεταξύ νοητικών καταστάσεων και λειτουργικών - υπολογιστικών καταστάσεων.

Το δεύτερο στοιχείο που πρέπει να επισημάνουμε σε σχέση με την ΣΕΡ αφορά τον τρόπο που ορίζονται οι όροι μιας υπολογιστικής θεωρίας. Κάθε υπολογιστική κατάσταση ορίζεται από την σχέση που έχει με όλες τις υπόλοιπες υπολογιστικές καταστάσεις, δηλαδή από την ολότητα των υπολογιστικών της καταστάσεων. Πρόκειται ουσιαστικά για κάποιο έμμεσο ορισμό. Έτσι, λοιπόν, σε μια υπολογιστική θεωρία όλες οι υπολογιστικές καταστάσεις ενός συστήματος ορίζονται ταυτόχρονα: ο ορισμός μιας κατάστασης προϋποθέτει τον ορισμό όλων των άλλων. Μία προφανής ένσταση είναι ότι δεν συμβαίνει κάτι ανάλογο και στη ψυχολογία του κοινού νου: Εκεί, προκειμένου να εξατομικεύσουμε μια νοητική κατάσταση δεν είναι αναγκαίο να γνωρίζουμε με ποιο τρόπο συνδέεται με κάθε άλλη νοητική κατάσταση. Οι υποστηρικτές της ΣΕΡ, αναγνωρίζοντας αυτή την αδυναμία, έκαναν λόγο για "ιδεατή ψυχολογική θεωρία": σε μια τέτοια θεωρία προκειμένου να ορισθεί κάποιος όρος θα πρέπει να προσδιορισθεί η σχέση του με όλους τους υπόλοιπους.

Ωστόσο, το πρόβλημα που τίθεται είναι πώς μπορεί ποτέ να κατασκευαστεί μια τέτοια θεωρία. Κανείς από τους υποστηρικτές της δεν έχει μέχρι σήμερα κατορθώσει να προχωρήσει πέρα από τη διατύπωση των βασικών αρχών που πρέπει να διέπουν την ΣΕΡ. Η ΣΕΡ, ουσιαστικά, αποτελεί μια προγραμματική διατύπωση για την ανάπτυξη μιας θεωρίας που όμως μέχρι στιγμής παραμένει σχεδόν ανύπαρκτη.

Ακόμα, όμως, και αν τελικά ξεκινήσει η κατασκευή μιας τέτοιας θεωρίας, υπάρχουν πολλά άλλα προβλήματα που θα πρέπει να επιλυθούν.

Υπάρχουν, καταρχάς, προβλήματα συμβατότητας της ΣΕΡ με την ΑΥΘΝ. Πιο συγκεκριμένα αν οι νοητικές αναπαραστάσεις της γλώσσας της σκέψης αποκτούν το περιεχόμενό τους με τον τρόπο που η ΣΕΡ περιγράφει, τότε οι νοητικές αναπαραστάσεις δεν είναι συνθετικές³⁹⁵. Αυτό συμβαίνει γιατί, σύμφωνα με τη ΣΕΡ, το περιεχόμενο μιας (συντακτικά σύνθετης) νοητικής αναπαράστασης δεν είναι συνάρτηση της συντακτικής δομής της και του περιεχομένου των λεξικών συστατικών της, αλλά συνάρτηση του τρόπου με τον οποίο η νοητική αναπαράσταση σχετίζεται με τις υπόλοιπες. Αν όμως η γλώσσα της σκέψης δεν είναι συνθετική, τότε, σύμφωνα με τον Fodor, δεν μπορεί να εξηγηθεί ούτε η παραγωγικότητα και η συστηματικότητα της σκέψης (βλ. κεφάλαιο 2, ενότητα 2.4.2).

Ένα δεύτερο πρόβλημα είναι ότι ο εγγενής ολισμός της ΣΕΡ οδηγεί σε εντελώς ανεπιθύμητα αποτελέσματα. Όπως επισημαίνουν οι Fodor & Lepore (1994, σελ. 145): "από τη στιγμή που αρχίζει κανείς να ταυτίζει το περιεχόμενο μιας πεποιθήσεως με τον συναγωγικό της ρόλο σε ένα σύστημα πεποιθήσεων ... είναι δύσκολο να αποφύγει την ταύτιση του περιεχομένου μιας πεποιθήσεως με τον *συνολικό* της συναγωγικό ρόλο στο σύστημα πεποιθήσεων. Και, έχοντας κάει τόσο μακριά, είναι δύσκολο να αποφύγει το συμπέρασμα ότι εάν [δύο] συστήματα πεποιθήσεων διαφέρουν κατά τι όσον αφορά τις προτάσεις που αποδέχονται, τότε διαφέρουν εντελώς όσον αφορά τις προτάσεις που αποδέχονται ... έχοντας φτάσει κανείς στον πάτο [αυτού του ολισθηρού δρόμου] βρίσκεται να αποδέχεται ορισμένες, εκ πρώτης όψεως, εξωφρενικές θέσεις όπως ότι δεν υπάρχουν δύο άνθρωποι που να μοιράζονται την ίδια πεποίθηση· ότι δεν υπάρχει σχέση μεταφρασής· ότι δεν υπάρχουν δύο άνθρωποι που να εννοούν το ίδιο πράγμα με αυτό που λένε [κ.λπ.]". Οι Fodor & Lepore (1992) θεωρούν ότι αυτό το συμπέρασμα μπορεί να αποφευχθεί μόνο αν κατορθώσουμε να ορίσουμε το περιεχόμενο μιας έννοιας βασιζόμενοι σε ένα υποσύνολο όλων των πεποιθήσεων στις οποίες αυτή χρησιμοποιείται. Η συγκεκριμένη κίνηση, όμως, θα ήταν δικαιολογημένη μόνο αν ίσχυε η διάκριση μεταξύ αναλυτικού και συνθετικού, δηλαδή η διάκριση μεταξύ κρίσεων που είναι αληθείς δυνάμει του νοημάτός τους και κρίσεων που είναι αληθείς δυνάμει του τρόπου που συμβαίνει να είναι ο κόσμος. Οι Fodor & Lepore, ωστόσο, αποδεχόμενοι την κριτική που έκανε ο Quine (1953) στη διάκριση αναλυτικού/ συνθετικού θεωρούν ότι κάτι τέτοιο δεν μπορεί να γίνει. Κατά συνέπεια, οι υποστηρικτές της ΣΕΡ δεν έχουν άλλο δρόμο παρά να ορίσουν το περιεχόμενο μιας έννοιας με βάση όλες τις πεποιθήσεις στις οποίες αυτή χρησιμοποιείται και έτσι να οδηγηθούν στο αδιέξοδο που περιγράφει ο Fodor.

Γενικότερα, η ΣΕΡ φαίνεται να δίνει προτεραιότητα στο περιεχόμενο των προτασιακών στάσεων έναντι του περιεχομένου των

³⁹⁵ Βλ. Fodor & Lepore 1992, σελ. 175-176.

νοητικών αναπαραστάσεων. Μια νοητική αναπαράσταση αποκτά το περιεχόμενό της από το σύνολο των προτασιακών στάσεων στις οποίες χρησιμοποιείται καθώς και από το σύνολο των συναγωγών στις οποίες συμμετέχουν οι προτασιακές στάσεις που περιέχουν την αναπαράσταση. Αν ακολουθήσουμε, μάλιστα, με συνέπεια τη λογική της ΣΕΡ ίσως δεν έχει καν νόημα να μιλάμε για νοητικές αναπαραστάσεις που συνιστούν μια προτασιακή στάση. Με άλλα λόγια, μικρότερη μονάδα ανάλυσης του νοητικού περιεχομένου φαίνεται να καθίσταται η προτασιακή στάση στο σύνολό της. Μια απόδειξη αυτού είναι ότι η πεποίθηση ότι p και η επιθυμία ότι p , σύμφωνα με τη ΣΕΡ, πρέπει να έχουν διαφορετικό περιεχόμενο, διότι μια πεποίθηση συνδέεται με άλλο σύνολο συναγωγών απ' ό,τι μια επιθυμία³⁹⁶. Το ότι, δηλαδή, και στις δύο περιπτώσεις πρόκειται για το ίδιο προτασιακό περιεχόμενο p , αυτό δεν είναι κάτι που μπορεί να εξασφαλισθεί με το οπλοστάσιο της ΣΕΡ.

Ένα άλλο πρόβλημα της ΣΕΡ προκύπτει από προβληματισμούς που έχουν αναπτυχθεί στο πλαίσιο της μαθηματικής λογικής. Η ΣΕΡ μπορεί ουσιαστικά να θεωρηθεί ως μια τυπική θεωρία που ορίζει τη σχέση μεταξύ των νοητικών αναπαραστάσεων με μορφικούς όρους. Όμως, σύμφωνα με το θεώρημα των Loewenheim-Skolem, είναι πάντα δυνατό να βρούμε ασύμβατες ερμηνείες μιας τυπικής θεωρίας στις οποίες όλα τα θεωρήματά της παραμένουν αληθή. Έτσι, μπορούμε να φανταστούμε δύο υπολογιστές οι οποίοι χρησιμοποιούν το ίδιο ακριβώς πρόγραμμα για να επιτελέσουν δύο εντελώς διαφορετικές εργασίες. Με άλλα λόγια, αν περιοριστεί κανείς στο "οπλοστάσιο" της ΣΕΡ υπάρχει ένας υποκαθορισμός του περιεχομένου των νοητικών αναπαραστάσεων. Ουσιαστικά, δηλαδή, η ΣΕΡ οδηγείται και αυτή σε μια μορφή του προβλήματος της άπειρης διάζευξης.

Τέλος, το σημαντικότερο πρόβλημα της ΣΕΡ είναι ότι δεν μπορεί να εξηγήσει την σύνδεση των νοητικών αναπαραστάσεων με τον κόσμο και έτσι, ουσιαστικά, αποτελεί μια σολιψιστική θεωρία. Εφόσον, το περιεχόμενο μιας νοητικής αναπαράστασης οφείλεται στον τρόπο που σχετίζεται με τις υπόλοιπες αναπαραστάσεις, δεν υπάρχει τρόπος να εξηγηθεί η αναφορική διάσταση του περιεχομένου της νοητικής αναπαράστασης. Να υπενθυμίσουμε σε αυτό το σημείο ότι η σχέση των αναπαραστάσεων αναμεταξύ τους ορίζεται λειτουργικά - συντακτικά. Όπως, ανέδειξε και το επιχείρημα του κινέζικου δωματίου (βλ. ενότητα 7.2) ο απλός χειρισμός συμβόλων με βάση κάποιους κανόνες που υπαγορεύουν ποια σύμβολα συνδέονται με ποια άλλα σύμβολα δεν μπορεί να εξηγήσει την αποβλεπτικότητα³⁹⁷.

³⁹⁶ Βλ. Cummins 1996, σελ. 32.

³⁹⁷ Γι' αυτό και στο τρίτο κεφάλαιο (ενότητα 3.3.1.4) υποστηρίξαμε ότι η μετάβαση από μια εγγραφή σε μια άλλη με βάση αποκλειστικά τη μορφή των εγγραφών δεν αποτελεί μετάφραση αλλά μετεγγραφή. Με άλλα λόγια, πρόκειται

Η ΣΕΡ, στην καλύτερη περίπτωση, μπορεί να αποτελέσει μια θεωρία των ορθών τρόπων συλλογισμού, δηλαδή μια θεωρία που, δεδομένων κάποιων αληθών αναπαραστάσεων, καθορίζει ποια συμπεράσματα είναι δικαιολογημένα. Αν η ΣΕΡ καθόριζε και την αλήθεια των νοητικών αναπαραστάσεων, τότε αυτό δεν θα μπορούσε να γίνει παρά μονάχα με αναφορά στο σύνολο των συναγωγών στις οποίες η αναπαρασταση θα επιτρεπόταν να συμμετέχει. Κατ' αυτόν τον τρόπο όμως η αλήθεια θα ταυτιζόταν με τη δικαιολόγηση, δηλαδή την ορθή συναγωγή.

7.2.6 Θεωρία διπλής όψης

Αρκετοί υποστηρικτές της ΣΕΡ αναγνωρίζοντας την αδυναμία της να εξηγήσει την αναφορική διάσταση του περιεχομένου των αναπαραστάσεων έσπευσαν να υιοθετήσουν συμπληρωματικά την αιτιακή θεωρία του νοητικού περιεχομένου. Το θεωρητικό υβρίδιο που προέκυψε ονομάστηκε θεωρία διπλής όψης (double aspect theory) και αποτελεί ουσιαστικά ένα πλαίσιο που έχει ως στόχο να συνδυάσει τα πλεονεκτήματα των ιντερναλιστικών και των εξτερναλιστικών θεωριών. Το πλαίσιο αυτό έχει ήδη τεθεί από τον Putnam (1975) με τη διάκριση του νοητικού περιεχομένου σε ευρύ και στενό: Από τη στιγμή που κανείς δέχεται ως βάσιμη τη συγκεκριμένη διάκριση δεν μπορεί παρά να δεχθεί ότι πρόκειμένου να φυσικοποιηθεί το νοητικό περιεχόμενο χρειάζονται δύο ειδών θεωρίες -μια για κάθε διάστασή του.

Ωστόσο, ο συνδυασμός δύο αποτυχημένων θεωριών δεν οδηγεί σε ένα επιτυχές υβρίδιο. Όπως είδαμε, οι εξτερναλιστικές θεωρίες παρότι αναπτύχθηκαν για να εξηγήσουν την αναφορική διάσταση των νοητικών αναπαραστάσεων, απέτυχαν σε αυτό το στόχο. Κατά συνέπεια, ο συνδυασμός τους με την ΣΕΡ δεν παρέχει τίποτε παραπάνω για την επίτευξη του εν λόγω στόχου. Αλλά και αντίστροφα, η ΣΕΡ η οποία αναπτύχθηκε για την εξήγηση της νοηματικής διάστασης του νοητικού περιεχομένου δεν κατόρθωσε να φωτίσει το συγκεκριμένο πρόβλημα. Οπότε ο συνδυασμός της ΣΕΡ με κάποια εξτερναλιστική θεωρία δεν συμβάλλει ούτε ως προς αυτή την κατεύθυνση.

7.3 Μια κοινά αποδεκτή διατύπωση του ιντερναλισμού και του εξτερναλισμού

Μέχρις εδώ εξετάσαμε τις σημαντικότερες εξτερναλιστικές και ιντερναλιστικές θεωρίες για την προέλευση του νοητικού περιεχομένου. Διαπιστώσαμε ότι και τα δύο είδη θεωριών αντιμετωπίζουν ανυπέρβλητα προβλήματα. Καμία από αυτές τις θεωρίες δεν επιτυγχάνει να εξηγήσει το πώς οι νοητικές αναπαραστάσεις αποκτούν το περιεχόμενό τους. Ωστόσο, χωρίς μια τέτοια εξήγηση, όλη η υπόθεση ότι τα νοητικά φαινόμενα

για μια μη αναπαραστασιακή χρήση των εγγραφών, δηλαδή μια χρήση που δεν τις καθιστά αναπαραστάσεις. Θα επανέλθουμε σε αυτό το ζήτημα στον Επύλογο.

οφείλονται στην ύπαρξη νοητικών αναπαραστάσεων καθώς και οι επακόλουθες διαμάχες για το ποια ακριβώς οφείλει να είναι η μορφή αυτών των αναπαραστάσεων, καθίστανται κυριολεκτικά μετέωρες. Έχοντας φτάσει σε αυτό το σημείο δεν μας απομένουν παρά δύο δρόμοι, είτε, πρώτον, να επιμείνουμε στην αναζήτηση κάποιας λύσης ιντερναλιστικού ή εξτερναλιστικού τύπου, είτε, δεύτερον, να απορρίψουμε την υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων και να θεωρήσουμε ως ψευδοπρόβλημα το επακόλουθο πρόβλημα της προέλευσης του περιεχομένου των νοητικών αναπαραστάσεων.

Οι περισσότεροι ερευνητές στον χώρο της σύγχρονης φιλοσοφίας του νου ακολουθούν τον πρώτο δρόμο. Εμείς, αντίθετα, θεωρούμε ότι τα προβλήματα που δημιουργεί η υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων είναι τόσο πολλά και μεγάλα ώστε αξίζει τον κόπο να διερευνηθούν άλλες προσεγγίσεις που απορρίπτουν αυτή την υπόθεση. Πώς θα μπορούσαν, όμως, να κατανοηθούν τα νοητικά φαινόμενα χωρίς αναφορά σε κάποιες εσωτερικές στον οργανισμό αναπαραστάσεις;

Προκειμένου να προχωρήσουμε προς αυτή την κατεύθυνση θα πρέπει να επανεξετάσουμε το ζήτημα των νοητικών φαινομένων. Μια νέα προσέγγιση του ζητήματος μπορεί να γίνει αν εντοπίσουμε το τι διαφεύγει από τον τρόπο που οι ιντερναλιστές και οι εξτερναλιστές εκλαμβάνουν τα νοητικά φαινόμενα. Θα αναζητήσουμε, λοιπόν, κοινές παραδοχές μεταξύ εξτερναλισμού και ιντερναλισμού οι οποίες μπορούν να τεθούν υπό αμφισβήτηση.

Ωστόσο, προκειμένου να ξεκινήσουμε μια τέτοια έρευνα πρέπει προηγουμένως να έχουμε καταλήξει σε μια διατύπωση του ιντερναλισμού και του εξτερναλισμού η οποία να είναι κοινά αποδεκτή από τους υποστηρικτές των διαφόρων θεωριών φυσικοποίησης του περιεχομένου των νοητικών αναπαραστάσεων. Στην παρούσα ενότητα θα αναζητήσουμε αυτή την κοινά αποδεκτή διατύπωση του ιντερναλισμού και του εξτερναλισμού, ενώ στο επόμενο κεφάλαιο θα εξετάσουμε κοινές παραδοχές των παραπάνω θεωριών η αμφισβήτηση των οποίων οδηγεί σε μια μη αναπαραστασιακή κατανόηση των νοητικών φαινομένων.

Μπορούμε να διακρίνουμε δύο κυρίως ερμηνείες του εξτερναλισμού και του ιντερναλισμού. Θα ονομάσουμε την πρώτη ερμηνεία "οντολογική" και τη δεύτερη "γνωσιολογική". Οι ερμηνείες αυτές, όπως θα δούμε, ορισμένες φορές στην πράξη συγχέονται.

Σύμφωνα με την οντολογική ερμηνεία, ο μιν εξτερναλισμός προσβέυει ότι οι νοητικές καταστάσεις βρίσκονται (τουλάχιστον, εν μέρει) έξω από το κεφάλι, ο δε ιντερναλισμός αποκλειστικά μέσα στο κεφάλι. Μια άλλη διατύπωση, που καθιστά πιο κατανοητή αυτή την ερμηνεία, είναι με όρους αιτίας και αποτελέσματος: αν θεωρήσουμε ότι οι νοητικές καταστάσεις αποτελούν τις αιτίες της συμπεριφοράς, τότε ο εξτερναλισμός προσβέυει ότι αυτές οι αιτίες βρίσκονται (τουλάχιστον, εν μέρει) έξω από το κεφάλι, ενώ ο ιντερναλισμός ότι βρίσκονται αποκλειστικά μέσα στο κεφάλι.

Ακολουθώντας τον Kim (1996, σελ. 200), μπορούμε να θεωρήσουμε ότι η οντολογική ερμηνεία έχει δύο τουλάχιστον εκδοχές: μία ισχυρή και μια ασθενή³⁹⁸. Σύμφωνα με την ισχυρή εκδοχή του οντολογικού εξτερναλισμού, οι νοητικές καταστάσεις *συγκροτούνται οντολογικά* εν μέρει από την αναφορά τους: "Θα μπορούσε κανείς να πει ότι η πεποίθηση ότι το νερό και το λάδι δεν αναμιγνύονται συνίσταται εν μέρει από νερό και λάδι, ότι η πεποίθηση η ίδια, κατά κάποιον τρόπο, ενέχει τα πραγματικά υλικά, το νερό και το λάδι, πέραν από το πρόσωπο που έχει την πεποίθηση (ή το "κεφάλι" του)... Η γενική ιδέα είναι ότι όλοι οι παράγοντες που παίζουν κάποιο ρόλο στον καθορισμό του περιεχομένου μιας πεποίθησης συγκροτούν οντολογικά αυτή την πεποίθηση, ότι [δηλαδή] η πεποίθηση είναι μια κατάσταση που περιλαμβάνει αυτά τα πράγματα εντός της". Αντίστοιχα, η ισχυρή εκδοχή του οντολογικού ιντερναλισμού είναι ότι οι νοητικές καταστάσεις δεν συγκροτούνται οντολογικά από οτιδήποτε βρίσκεται εκτός του κεφαλιού (ή, ακριβέστερα, εκτός του οργανισμού).

Να επισημάνουμε σε αυτό το σημείο ότι η έκφραση "έξω από το κεφάλι" δεν πρέπει να συγχέεται με την έκφραση "έξω από το νου". Οι νοητικές καταστάσεις, σύμφωνα με την οντολογική ερμηνεία του εξτερναλισμού, όντας νοητικές, είναι *μέσα στο νου*. Με άλλα λόγια, ο νους δεν ταυτίζεται με ό,τι βρίσκεται μέσα στο κεφάλι, δηλαδή με τον εγκέφαλο, αλλά περιλαμβάνει και πράγματα εκτός των ορίων του κεφαλιού, για την ακρίβεια, εκτός των ορίων του σώματος.

Η ισχυρή εκδοχή της οντολογικής ερμηνείας βασίζεται ουσιαστικά στην υποστασιοποίηση του νοητικού. Οι υποστάσεις, όπως επισημαίνει ο McGinn (1989, σελ. 21), έχουν μια χωρική διακρίσιμτητα που τους παρέχει αυτονομία και αποκλειστικότητα: "οι υποστάσεις είναι χωρικά όντα των οποίων η διαμόρφωση είναι ανεξάρτητη από οτιδήποτε κείται πέραν αυτών· δεν ενσωματώνουν συστατικά από αλλού. Σε αυτό έγκειται, ουσιαστικά, το αυθυτόστατό τους". Οι νοητικές καταστάσεις εκλαμβάνονται ως υποστάσεις που συνίστανται, στην περίπτωση του ιντερναλισμού, αποκλειστικά από κάτι που βρίσκεται μέσα στο κεφάλι, ενώ στην περίπτωση του εξτερναλισμού, από κάτι που βρίσκεται μέσα στο κεφάλι και από τμήματα του κόσμου.

Ωστόσο, όσον αφορά τον εξτερναλισμό, θεωρούμε ότι η ισχυρή εκδοχή δεν αντέχει σε κριτική διότι, δυνάμει, καθιστά τμήμα του νου οτιδήποτε υφίσταται εκτός του κεφαλιού. Κάθε νοητική κατάσταση κατανοείται ως μια διαφορετική υπόσταση που περιλαμβάνει το εγκεφαλικό υπόστρωμα μαζί με τα εκάστοτε όντα στα οποία αναφέρεται η νοητική κατάσταση. Έτσι, το νοητικό επεκτείνεται χωρίς όριο και μπορεί να συμπεριλάβει οποιοδήποτε φυσικό αντικείμενο. Αυτό σημαίνει, ταυτόχρονα, ότι το νοητικό καθίσταται διαρκώς μεταβαλλόμενο και έτσι,

³⁹⁸ Οι όροι "ισχυρή οντολογική ερμηνεία" και "ασθενής οντολογική ερμηνεία" είναι δικό μου και όχι του Kim.

ουσιαστικά, χάνει τη χωρική διακριτότητά του. Επιπλέον, εφόσον τα νοητικά φαινόμενα δεν περιορίζονται σε ό,τι βρίσκεται μέσα στο κεφάλι, δεν είναι δυνατόν ο ισχυρός οντολογικός εξτερναλισμός να συμβιβαστεί με την υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων.

Δεν ισχύει, όμως, το ίδιο και για την ισχυρή εκδοχή του οντολογικού ιντερναλισμού. Τόσο οι θεωρίες ταυτότητας όσο και η θεωρία επιγένεσης που κυριαρχούν στην φιλοσοφία του νου (βλέπε κεφάλαιο 1 (ενότητες 1.4.2 και 1.4.3) και κεφάλαιο 7 (ενότητα 7.2.1)) όσον αφορά τη σχέση των νοητικών καταστάσεων με τις εγκεφαλικές καταστάσεις, ουσιαστικά προϋποθέτουν, τις περισσότερες φορές, μια υποστασιακή αντίληψη για το νοητικό. Εξάλλου, ο ίδιος ο όρος "νοητική κατάσταση" με τον οποίο προσδιορίζονται τα νοητικά φαινόμενα παραπέμπει σε κάποιο είδος υπόστασης: Ο ισχυρός οντολογικός ιντερναλισμός μπορεί να θεωρηθεί ως μια μορφή φυσικοποιημένου καρτεσιανισμού, διότι διατηρεί την ιδέα του ριζικού διαχωρισμού νου - κόσμου την οποία εντοπίζει όχι στην διαφορά δύο ειδών υπόστασης (*res cogitans* και *res extensa*), αλλά στον *τοπολογικό χωρισμό* δύο υποστάσεων του ίδιου είδους: ο νους όντας μέσα στο κεφάλι είναι διακριτός από τον εξωτερικό κόσμο.

Ας δούμε, όμως τώρα, την δεύτερη εκδοχή της οντολογικής ερμηνείας, την οποία θα ονομάσουμε "ασθενή οντολογική ερμηνεία". Σύμφωνα με την ασθενή οντολογική ερμηνεία του εξτερναλισμού οι νοητικές καταστάσεις είναι *σχεσιακές ιδιότητες* που υφίστανται μεταξύ του οργανισμού και πραγμάτων του κόσμου. Ο Kim (1996, σελ. 200) επισημαίνει χαρακτηριστικά: "Η πεποίθηση ότι το νερό και το λάδι δεν αναμιγνύονται θα μπορούσε να θεωρηθεί ως μια συγκεκριμένη σχέση που ισχύει μεταξύ του υποκειμένου από τη μια μεριά και του νερού και του λαδιού από την άλλη. Ή, εναλλακτικά, η πεποίθηση μπορεί να θεωρηθεί ως μια *σχεσιακή ιδιότητα* του υποκειμένου που εμπλέκει το νερό και το λάδι". Κατ' αντιστοιχία, η ασθενής οντολογική ερμηνεία του ιντερναλισμού θα είναι ότι οι νοητικές καταστάσεις αποτελούν *εγγενείς (intrinsic) ιδιότητες του οργανισμού*³⁹⁹.

³⁹⁹ Βασικός υποστηρικτής αυτής της θέσης είναι ο Searle (βλ. Searle 1983 και Searle 1992), ωστόσο αυτός δεν υποστηρίζει, ταυτόχρονα, την υπόθεση της ΑΥΘΝ περί ύπαρξης νοητικών αναπαραστάσεων. Είναι, επίσης, αξιοσημείωτο ότι η σημασιολογία εννοιολογικού ρόλου, αν και ιντερναλιστική προσέγγιση, δεν ανατρέπει σε κάποιες *εγγενείς ιδιότητες των νοητικών συμβόλων*, αλλά στις μεταξύ τους λειτουργικές σχέσεις, δηλαδή σε *σχεσιακές ιδιότητες των νοητικών συμβόλων*. Ωστόσο, η προσέγγιση αυτή παραμένει ιντερναλιστική γιατί οι εν λόγω λειτουργικές ιδιότητες ορίζονται αποκλειστικά από τις σχέσεις οντοτήτων που βρίσκονται μέσα στο κεφάλι. Μπορούμε, με άλλα λόγια, να ισχυρισθούμε ότι οι λειτουργικές ιδιότητες είναι εγγενείς ιδιότητες του *οργανισμού* με την έννοια ότι ορίζονται αποκλειστικά με ό,τι συμβαίνει εντός του κεφαλιού. Εξάλλου, σύμφωνα με την ΑΥΘΝ, οι σημασιολογικές σχέσεις των νοητικών συμβόλων

Η ασθενής οντολογική ερμηνεία του εξτερναλισμού είναι, ίσως, η περισσότερο αποδεκτή απ' όσους συντάσσονται με τον εξτερναλισμό. Όπως επιστημαίνει και η Baker (1995, σελ. 63): "Υπάρχει μια αυξανόμενη συναίνεση, με την οποία συμφωνώ, ότι η ιδιότητα τού να έχει κανείς μια πεποιθήση ότι το νερό είναι υγρό είναι ... μια σχεσιακή ιδιότητα". Επίσης, ο McGinn, βασικός υποστηρικτής του εξτερναλισμού ως προς ορισμένα είδη νοητικών φαινομένων, δηλώνει: "Κατά τον εξτερναλισμό, ο νους είναι ουσιαστικά ένα σύστημα σχέσεων μεταξύ του υποκειμένου και των εγκόσμιων αντικειμένων των νοητικών του καταστάσεων" (McGinn 1989, σελ. 29, υποσημ. 39).

Τι ακριβώς, όμως, είναι μια σχεσιακή ιδιότητα; Ας ξεκινήσουμε με ένα παράδειγμα. Αν πούμε ότι μια πέτρα είναι από πυρίτιο, τότε αναφερόμαστε σε μια εγγενή ιδιότητα της πέτρας. Αν, όμως, πούμε ότι η πέτρα είναι πλανήτης, τότε σίγουρα αναφερόμαστε σε μια σχεσιακή ιδιότητα της πέτρας: ένα σώμα Α είναι πλανήτης μόνο αν περιστρέφεται γύρω από κάποιο άλλο σώμα Β. Με άλλα λόγια, η ύπαρξη αυτής της ιδιότητας για το σώμα Α προϋποθέτει την ύπαρξη του σώματος Β⁴⁰⁰. Το παράδειγμα αυτό δίνει μια αίσθηση τού τι είναι σχεσιακή ιδιότητα, ωστόσο αν προσπαθήσει κανείς να προσδιορίσει θεωρητικά την έννοια δημιουργούνται πολλά προβλήματα⁴⁰¹. Χωρίς να εμπλακούμε στην λογική του εντοπισμού ενός επαρκούς ορισμού των σχεσιακών ιδιοτήτων - ένα έργο που βαρύνει τους υποστηρικτές του εξτερναλισμού⁴⁰² - θα προβούμε απλώς σε ορισμένες παρατηρήσεις.

Καταρχάς, όπως είδαμε στις προηγούμενες ενότητες, οι σχεσιακές ιδιότητες που εξετάζονται από τους εξτερναλιστές είναι είτε αιτιακές - πληροφοριακές, είτε ιστορικές - τελεολογικές. Και στις δύο περιπτώσεις, οι σχέσεις αυτές ισχύουν μεταξύ των νοητικών καταστάσεων του οργανισμού και αντικειμένων του κόσμου, δηλαδή μεταξύ δύο υποστάσεων. Με άλλα λόγια, ο ασθενής οντολογικός εξτερναλισμός κατανοεί τα νοητικά φαινόμενα ως μία σχεσιακή ιδιότητα κάποιων υποστάσεων που βρίσκονται εντός του κεφαλιού, δηλαδή των

κατοπιρίζονται στις συντακτικές τους σχέσεις, οι οποίες, αν νοηθούν ως μορφικές σχέσεις, αποτελούν εγγενείς ιδιότητες των συμβόλων.

⁴⁰⁰ Ένα άλλο παράδειγμα σχεσιακής ιδιότητας είναι η αξία των μετοχών. Το πόσο αποτιμάται μια μετοχή δεν είναι κάτι που εξαρτάται από τις εγγενείς της ιδιότητες αλλά από τη "θέση" της στην αγορά.

⁴⁰¹ Βλ. Lewis 1983, Dunn 1990, Sider (υπό έκδοση).

⁴⁰² Ένας πρώτος ορισμός είναι αυτός που δίνει η Baker (1995, σελ. 63): "Λέμε ότι η R είναι μια σχεσιακή ιδιότητα αν και μόνο αν: το να έχει το x την R συνεπάγεται ότι υπάρχει κάποιο y διακριτό από το x". Ωστόσο αυτός ο ορισμός είναι μάλλον ελλειπής διότι δεν αναφέρεται στο εάν και πώς το x σχετίζεται με το y. Εξάλλου, η Baker υποστηρίζει ένα διαφορετικό είδος εξτερναλισμού που δεν θα μας απασχολήσει γιατί δεν αποδέχεται την υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων. Βλ. επίσης, Brown 1992, σελ. 345-349.

αναπαραστασιακών φορέων των νοητικών αναπαραστάσεων. Οι αναπαραστασιακοί φορείς, όπως κάθε υπόσταση, εκτός από τις εγγενείς ιδιότητες (όπως το σχήμα και το υλικό) μπορούν να διαθέτουν και κάποιες σχεσιακές ιδιότητες. Το αναπαραστασιακό περιεχόμενο είναι, σύμφωνα με τον ασθενή οντολογικό εξτερναλισμό, μια τέτοια σχεσιακή ιδιότητα των εσωτερικών αναπαραστασιακών φορέων. Οι εξτερναλιστικές προσεγγίσεις που εξετάσαμε στην ενότητα 7.2.4 μπορούν να θεωρηθούν ως προσπάθειες προσδιορισμού των συνθηκών που πρέπει να πληρούν σχεσιακές ιδιότητες προκειμένου να συνιστούν αναπαραστασιακές-νοητικές ιδιότητες.

Ωστόσο, αν κανείς, τελικά, δεχθεί την ασθενή εκδοχή του οντολογικού εξτερναλισμού τίθεται ένα γενικότερο ζήτημα που αφορά το οντολογικό καθεστώς των σχεσιακών ιδιοτήτων: είναι, άραγε, οι σχεσιακές ιδιότητες πραγματικές ιδιότητες των υποστάσεων ή απλώς αποτελούν τρόπους περιγραφής των υποστάσεων; Αν κανείς επιλέξει το δεύτερο δρόμο, οδηγείται σε μια αντιρεαλιστική θέση ως προς τα νοητικά φαινόμενα: οι σχεσιακές ιδιότητες που χαρακτηρίζουν τις εσωτερικές καταστάσεις του οργανισμού δεν αποτελούν πραγματικές ιδιότητές τους, αλλά προσδιορισμούς που ο ερευνητής-ερμηνευτής αποδίδει σε αυτές προκειμένου να τις εξατομικεύσει⁴⁰³. Οι προσδιορισμοί αυτοί αποτελούν απλώς σχεσιακές περιγραφές των εσωτερικών καταστάσεων του οργανισμού με βάση τις συνθήκες που επικρατούν εκτός του οργανισμού (είτε συγχρονικά είτε ιστορικά). Ανεξάρτητα, όμως, από το εάν κανείς, τελικά, υιοθετεί μια ρεαλιστική ή μια αντιρεαλιστική στάση ως προς τις σχεσιακές ιδιότητες, αυτό που έχει σημασία είναι ότι ένας ερευνητής προκειμένου να εντοπίσει (αν είναι ρεαλιστής) ή να αποδώσει (αν είναι αντιρεαλιστής) μια σχεσιακή ιδιότητα σε μια εσωτερική κατάσταση του οργανισμού πρέπει να γνωρίζει τις συνθήκες που ισχύουν εκτός του οργανισμού. Με άλλα λόγια, πέρα από την οντολογική διάσταση υπάρχει και μια γνωσιολογική διάσταση που αφορά τον τρόπο που μπορούμε να γνωρίσουμε τα νοητικά φαινόμενα⁴⁰⁴.

Φτάνουμε, έτσι, στην ερμηνεία του εξτερναλισμού που ονομάσαμε "γνωσιολογική ερμηνεία". Σύμφωνα με αυτή την ερμηνεία, ο εξτερναλισμός είναι η θέση ότι, προκειμένου ο ερευνητής να εξατομικεύσει και να περιγράψει τις νοητικές καταστάσεις ενός υποκειμένου, χρειάζεται να γνωρίζει τις συνθήκες μέσα στις οποίες αυτό είναι ενταγμένο. Οι υπό εξέταση συνθήκες μπορεί να είναι φυσικές, κοινωνικές, ιστορικές, αναπτυξιακές ή εξελικτικές. Αντίστοιχα, μπορεί να

⁴⁰³ Βασικός εκφραστής αυτής της θέσης είναι ο Dennett 1969 ο οποίος προτείνει την αποβλεπτική στάση (intentional stance) ως την ερευνητική στάση που προσδιάζει στην ψυχολογία.

⁴⁰⁴ Φυσικά, όπως θα δούμε λίγο παρακάτω, στην περίπτωση της αντιρεαλιστικής στάσης αυτές οι δύο διαστάσεις συμψύρονται.

διατυπωθεί και η γνωσιολογική ερμηνεία του ιντερναλισμού: ο ερευνητής, προκειμένου να εξετασθεί και να περιγράψει τις νοητικές καταστάσεις ενός υποκειμένου, δεν χρειάζεται να γνωρίζει στιδήποτε συμβαίνει εκτός του υποκειμένου.

Η γνωσιολογική ερμηνεία αποτελεί μια διακριτή ερμηνεία μόνο στην περίπτωση που συνδυάζεται με την αντιρεαλιστική θέση ως προς τα νοητικά φαινόμενα. Σε αυτή την περίπτωση, τα νοητικά φαινόμενα δεν θεωρούνται ότι υπάρχουν πραγματικά, απλά ο ερευνητής - ερμηνευτής τα αποδίδει στον υπό εξέταση οργανισμό προκειμένου να εξηγήσει την συμπεριφορά του. Η ίδια η πράξη της ερμηνείας στην οποία προβαίνει ο ερευνητής-ερμηνευτής, εκτός από γνωσιολογικό, παίζει και οντολογικό ρόλο: θεμελιώνει την ύπαρξη των αποδιδόμενων νοητικών φαινομένων (βλ. Kim 1996, σελ. 189-190). Με άλλα λόγια, τα νοητικά φαινόμενα υπάρχουν μόνο στο βαθμό που υπάρχει κάποιος ερμηνευτής που τα αποδίδει σε ένα υποκείμενο. Κατ' αυτό τον τρόπο όμως, τα νοητικά φαινόμενα καθίστανται μια απλή προβολή. Ο ερευνητής - ερμηνευτής υποθετώντας μια *αποβλεπτική* στάση συμπεριφέρεται ως εάν να υπήρχαν τα νοητικά φαινόμενα στα οποία αναφέρεται (Dennett 1969, 1987).

Ωστόσο, αυτή η προσέγγιση δεν μπορεί να συμπεριληφθεί στις θεωρίες φυσικοποίησης του νοητικού περιεχομένου διότι επιχειρεί να εξηγήσει τα νοητικά φαινόμενα ανατρέχοντας στην αποβλεπτική στάση του ερμηνευτή. Αλλά η αποβλεπτική στάση του ερμηνευτή αποτελεί, με τη σειρά της, και αυτή ένα νοητικό φαινόμενο, οπότε οδηγούμαστε σε φαύλο κύκλο. Όπως επισημαίνει και ο Putnam (1998, σελ.19-20): "...στις τελευταίες δεκαετίες του εικοστού αιώνα ... έχουμε ευφυνείς φιλοσόφους που ισχυρίζονται ότι η ίδια η αποβλεπτικότητα⁴⁰⁵ είναι κάτι που προβάλλουμε παίρνοντας μια 'στάση' απέναντι σε ορισμένα τμήματα του κόσμου (λες και το ίδιο το γεγονός τού να 'παίρνεις μια στάση' δεν είναι μια αποβλεπτική έννοια!)". Γι' αυτούς τους λόγους δεν μας απασχολήσει περαιτέρω η αντιρεαλιστική γνωσιολογική ερμηνεία.

Αν η γνωσιολογική ερμηνεία συνδυαστεί με την ρεαλιστική θέση, τότε είναι συμβατή και με τις δύο εκδοχές της οντολογικής ερμηνείας. Σε αυτή την περίπτωση, αποτελεί απλώς μια κατευθυντήρια οδηγία για το πώς μπορεί να εντοπισθεί από τον ερευνητή ένα *υπάρχον* νοητικό φαινόμενο, είτε κατανοείται ως υπόσταση είτε ως ιδιότητα. Υπ' αυτή την έννοια, η ρεαλιστική γνωσιολογική ερμηνεία προϋποθέτει την ισχύ μιας εκ των δύο εκδοχών της οντολογικής ερμηνείας.

Αν εξαιρέσουμε την ισχυρή ερμηνεία του οντολογικού εξτερναλισμού και την αντιρεαλιστική γνωσιολογική ερμηνεία, οι οποίες, όπως είδαμε, αποτελούν προβληματικές θέσεις, όλες οι υπόλοιπες

⁴⁰⁵ Για λόγους ομοιογένειας με το υπόλοιπο κείμενο αποδίδω τον όρο intentionality ως αποβλεπτικότητα παρότι στην ελληνική μετάφραση του βιβλίου του Putnam αποδίδεται ως προθετικότητα.

ερμηνείες του ιντερναλισμού και του εξτερναλισμού φαίνονται, καταρχήν, να μην είναι άμεσα απορρίψιμες ενώ, ταυτόχρονα, είναι συμβατές με την υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων. Δεδομένου δε ότι η ρεαλιστική γνωσιολογική ερμηνεία προϋποθέτει την ισχύ μιας εκ των δύο εκδοχών της οντολογικής ερμηνείας, μπορούμε να διατυπώσουμε με ένα συνοψιστικό και ενοποιητικό τρόπο τις αντίστοιχες θέσεις του ιντερναλισμού και του εξτερναλισμού που προϋποτίθενται των διαφόρων προγραμμάτων φυσικοποίησης του νοητικού περιεχομένου:

Ιντερναλισμός: ο ερευνητής, προκειμένου να εξατομικεύσει και να περιγράψει τα νοητικά φαινόμενα ενός υποκειμένου (είτε αυτά νοούνται ως εγγενείς ιδιότητες των εσωτερικών του καταστάσεων είτε ως υποστάσεις εσωτερικές στο κεφάλι), δεν χρειάζεται να γνωρίζει τίποτε απ' ό,τι συμβαίνει εκτός του υποκειμένου.

Εξτερναλισμός: ο ερευνητής, προκειμένου να εξατομικεύσει και να περιγράψει τα νοητικά φαινόμενα ενός υποκειμένου (τα οποία νοούνται ως σχεσιακές ιδιότητες των εσωτερικών του καταστάσεων), χρειάζεται να γνωρίζει τις φυσικές, κοινωνικές, ιστορικές, αναπτυξιακές ή εξελικτικές συνθήκες μέσα στις οποίες αυτό είναι ενταγμένο.

7.4 *Επίλογος*

Σε αυτό το κεφάλαιο εξετάσαμε κριτικά τα σημαντικότερα προγράμματα φυσικοποίησης του περιεχομένου των νοητικών αναπαραστάσεων. Τόσο τα ιντερναλιστικά όσο και τα εξτερναλιστικά προγράμματα αποδείχθηκαν ανεπαρκή στο να εξηγήσουν την αποβλεπτικότητα των νοητικών αναπαραστάσεων. Χωρίς μια τέτοια εξήγηση, ωστόσο, η ίδια η έννοια της νοητικής αναπαράστασης καθίσταται μετέωρη. Ολόκληρη η κλασική γνωσιακή επιστήμη βρίσκεται να υιοθετεί και να χρησιμοποιεί μια θεωρητική οντότητα για την οποία δεν μπορεί να εξηγηθεί η δυνατότητα ύπαρξής της: τόσο οι ιντερναλιστικές όσο και οι εξτερναλιστικές θεωρίες αποτυγχάνουν να εξηγήσουν πώς μια εσωτερική κατάσταση του εγκεφάλου μπορεί να αποκτήσει αποβλεπτικότητα. Μπροστά σε αυτό το αδιέξοδο θεωρήσαμε ότι είναι σκόπιμο να εγκαταλείψουμε την υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων και να αναζητήσουμε μια μη αναπαραστασιακή προσέγγιση των νοητικών φαινομένων. Αφού, τόσο η ιντερναλιστική όσο και η εξτερναλιστική προσέγγιση αποδείχθηκαν ανεπαρκείς, ένας τρόπος να προχωρήσουμε είναι να αναζητήσουμε κοινές παραδοχές των δύο προσεγγίσεων που μπορούν να τεθούν υπό αμφισβήτηση. Στόχος αυτής της μεθοδολογικής γραμμής είναι να φτάσουμε σε μια πιο πιστή περιγραφή των νοητικών φαινομένων. Στο επόμενο κεφάλαιο, έχοντας αφαιρέσει την κοινή αποδεκτή διατύπωση του ιντερναλισμού και του εξτερναλισμού στην οποία καταλήξαμε, θα προβούμε λοιπόν σε μια *εξωτερική* κριτική. Μέσα από την αμφισβήτηση των κοινών παραδοχών θα αναδειχθεί μια περιγραφή των νοητικών φαινομένων όπως αυτά μας

δίδονται πρωταρχικά και η οποία δεν θα στηρίζεται στην έννοια της νοητικής αναπαράστασης.

ΚΕΦΑΛΑΙΟ ΟΓΔΩΟ

Κοινές παραδοχές του ιντερναλισμού και του εξτερναλισμού και ο δρόμος για μια μη αναπαραστασιακή θεώρηση των νοητικών φαινομένων

8.1 Εισαγωγή

Στο τέλος του προηγούμενου κεφαλαίου καταλήξαμε σε μια διατύπωση του ιντερναλισμού και του εξτερναλισμού η οποία είναι κοινά αποδεκτή από τους υποστηρικτές των διαφόρων θεωριών φυσικοποίησης του περιεχομένου των νοητικών αναπαραστάσεων. Οι αντίστοιχες διατυπώσεις έχουν ως εξής:

Ιντερναλισμός: ο ερευνητής, προκειμένου να εξατομικεύσει και να περιγράψει τα νοητικά φαινόμενα ενός υποκειμένου (είτε αυτά νοούνται ως εγγενείς ιδιότητες των εσωτερικών του καταστάσεων είτε ως υποστάσεις εσωτερικές στο κεφάλι), δεν χρειάζεται να γνωρίζει τίποτε απ' ό,τι συμβαίνει εκτός του υποκειμένου.

Εξτερναλισμός: ο ερευνητής, προκειμένου να εξατομικεύσει και να περιγράψει τα νοητικά φαινόμενα ενός υποκειμένου (τα οποία νοούνται ως σχεσιακές ιδιότητες των εσωτερικών του καταστάσεων), χρειάζεται να γνωρίζει τις φυσικές, κοινωνικές, ιστορικές, αναπτυξιακές ή εξελικτικές συνθήκες μέσα στις οποίες αυτό είναι ενταγμένο.

Προκειμένου να προχωρήσουμε προς την κατεύθυνση μιας αντι-αναπαραστασιακής προσέγγισης των νοητικών φαινομένων πρέπει να εντοπίσουμε τι διαφεύγει από τον τρόπο που οι ιντερναλιστές και οι εξτερναλιστές κατανοούν τα νοητικά φαινόμενα. Θα αναζητήσουμε, λοιπόν, κοινές παραδοχές μεταξύ εξτερναλισμού και ιντερναλισμού οι οποίες μπορούν να αμφισβητηθούν. Συγκεκριμένα, θα αναφερθούμε σε δύο τέτοιες παραδοχές οι οποίες θεωρούμε ότι ευθύνονται σε σημαντικό βαθμό για τα προβλήματα που αντιμετωπίζουμε στην κατανόηση των νοητικών φαινομένων. Παράλληλα, η αμφισβήτηση αυτών των παραδοχών θα αποτελέσει και μια έσχατη κριτική στην προσπάθεια κατανόησης των νοητικών φαινομένων μέσω της έννοιας της νοητικής αναπαράστασης.

8.2 Πρώτη κοινή παραδοχή: Η οπτική του τρίτου προσώπου

Μια πρώτη κοινή παραδοχή μεταξύ της ιντερναλιστικής και της εξτερναλιστικής εκδοχής της Αναπαραστασιακής Υπολογιστικής Θεωρίας του Νου (ΑΥΘΝ) είναι η υιοθέτηση της οπτικής του τρίτου προσώπου. Βέβαια, όταν στην καθημερινή μας ζωή αποδίδουμε από την οπτική του τρίτου προσώπου σε ένα υποκείμενο νοητικά φαινόμενα δεν ανατρέχουμε σε κάποιες εσωτερικές αναπαραστάσεις του υποκειμένου τις οποίες εξατομικεύουμε εξετάζοντας αποκλειστικά τον εγκέφαλό του είτε και το πλαίσιο- με τη γενικότερη έννοια του όρου (τρέχουσες συνθήκες, ιστορία κ.λπ.) - εντός του οποίου βρίσκεται το εν λόγω υποκείμενο. Στην καθημερινή μας ζωή, για να αποδώσουμε από την οπτική του τρίτου προσώπου ένα νοητικό φαινόμενο σε ένα υποκείμενο, ανατρέχουμε στη συμπεριφορά του - στο τι λέει και το τι κάνει - βασιζόμενοι πάντα στο ευρύτερο πλαίσιο μέσα στο οποίο εντάσσεται. Με άλλα λόγια, στην καθημερινή μας ζωή, όταν προσεγγίζουμε από την οπτική του τρίτου

προσώπου τα νοητικά φαινόμενα, υιοθετούμε την εξτερναλιστική στάση, με τη διαφορά ότι σταματάμε στα όρια του οργανισμού και δεν προχωρούμε στο τι συμβαίνει στο εσωτερικό του υποκειμένου και, ειδικότερα, στον εγκέφαλό του. Αν, όμως, δεχόμαστε την υπόθεση της ΑΥΘΝ ότι τα νοητικά φαινόμενα δεν είναι τίποτα άλλο από την κατοχή και το χειρισμό εσωτερικών αναπαραστάσεων στον εγκέφαλο, τότε θα ήμασταν υποχρεωμένοι να δεχτούμε ότι στην καθημερινή μας ζωή δεν έχουμε άμεση πρόσβαση στα ίδια τα νοητικά φαινόμενα αλλά σε ορισμένα αποτελέσματά τους.

Άμεση πρόσβαση σημαίνει, σε πρώτη προσέγγιση, *αντιληπτική πρόσβαση*. Τα αισθητήρια όργανα που διαθέτουμε μας επιτρέπουν να αντιλαμβανόμαστε το τι συμβαίνει μέχρι τα εξωτερικά όρια των άλλων. Ωστόσο, αυτός ο περιορισμός θα μπορούσε κανείς να πει ότι είναι καθαρά *εμπειρικός*: αν η υπόθεση της ΑΥΘΝ αποδεικνυόταν σωστή, τότε τίποτε δεν θα απαγόρευε το ενδεχόμενο να κατόρθωνε η επιστήμη κάποια στιγμή στο μέλλον να επέκτεινε τις αντιληπτικές μας ικανότητες και να μας έδινε τη δυνατότητα με εξελεγμένες συσκευές -ας τις ονομάσουμε "εγκεφαλοσκόπια" - να βλέπουμε τι συμβαίνει στο εσωτερικό του υποκειμένου και ειδικότερα στον εγκέφαλό του⁴⁰⁶. Η *ιντερναλιστική εκδοχή* ενός τέτοιου εγκεφαλοσκόπου θα συνέλεγε στοιχεία αποκλειστικά από το εσωτερικό του υποκειμένου, ενώ η *εξτερναλιστική* θα συνέλεγε στοιχεία και από το εξωτερικό περιβάλλον. Πάντως και στις δύο αυτές υποθετικές περιπτώσεις, θα μπορούσαμε να παρατηρούμε από την οπτική του τρίτου προσώπου τις νοητικές αναπαραστάσεις, δηλαδή, σύμφωνα με την ΑΥΘΝ, τα ίδια τα νοητικά φαινόμενα.

Γιατί, όμως, είναι αμφισβητήσιμη η κοινή παραδοχή της *ιντερναλιστικής* και της *εξτερναλιστικής* εκδοχής της ΑΥΘΝ ότι τα νοητικά φαινόμενα προσεγγίζονται από την οπτική του τρίτου προσώπου; Μία πρώτη απάντηση είναι ότι η υιοθέτηση αυτής της οπτικής ως *αποκλειστικού* τρόπου προσέγγισης των νοητικών φαινομένων οδηγεί σε

⁴⁰⁶ Ορισμένοι μάλιστα, όχι μόνο δημοσιογράφοι, θεωρούν ότι αυτό το ενδεχόμενο είναι πλέον πολύ κοντά στο να γίνει πραγματικότητα. Όλο και συχνότερα διαβάζουμε "ειδήσεις" όπως αυτή της *Ελευθεροτυπίας* 2/9/2000 με τίτλο "Συσκευές θα καταγράφουν μνήμη και μυρωδιές": "Τη μνήμη και τις εικόνες από το μυαλό του ανθρώπου θα μπορούν να καταγράφουν οι επιστήμονες με τη χρήση μιας εξαιρετικά εξελεγμένης ηλεκτρονικής συσκευής, σύμφωνα με την έκθεση επιτροπής για την καταπολέμηση του εγκλήματος που εξέτασε το θέμα υπό την αιγίδα του βρετανικού υπουργείου Εμπορίου και Βιομηχανίας. Σύμφωνα με την έκθεση 'Just Around the Corner', η 'νευροχημική τεχνολογία', όπως αναπτύσσεται, θα επιτρέπει την καταγραφή, για παράδειγμα, της τελευταίας εικόνας στο μυαλό ενός θύματος, όπως άλλωστε και ευχάριστες μνήμες ανθρώπων. 'Με τη χρήση της κυβερνητικής, θα δημιουργηθεί συσκευή που θα προσαρμόζεται στον ανθρώπινο οργανισμό και θα καταγράφει τις μνήμες του. Στην περίπτωση εγκλήματος θα καταγράφεται το πρόσωπο του εγκληματία' αναφέρεται στην έκθεση".

μια "διαστρεβλωμένη" εικόνα των νοητικών φαινομένων: έτσι αποσιωπάται και δεν εξηγείται η γνωσιολογική ασυμμετρία μεταξύ της οπτικής του τρίτου προσώπου και εκείνης του πρώτου. Πιο συγκεκριμένα, τα νοητικά φαινόμενα, όταν είναι συνειδητά, προσεγγίζονται και από την οπτική του πρώτου προσώπου, οπότε, σε αυτή την περίπτωση, παρουσιάζουν ορισμένα χαρακτηριστικά που δεν συμβιβάζονται με ό,τι εμφανίζεται από την οπτική του τρίτου προσώπου. Όταν ένα υποκείμενο Α προσεγγίζει τα ίδια τα δικά του νοητικά φαινόμενα, δηλαδή προσεγγίζει τα νοητικά φαινόμενα από την οπτική του πρώτου προσώπου, τότε αυτά εμφανίζονται με διαφορετικό τρόπο από τον τρόπο που εμφανίζονται όταν προσεγγίζονται από την οπτική του τρίτου προσώπου, δηλαδή από ένα άλλο υποκείμενο που παρατηρεί είτε (α) το υποκείμενο Α εξωτερικά (όπως συμβαίνει στην καθημερινή μας ζωή), είτε (β) τον εγκέφαλο του υποκειμένου Α μέσω κάποιου εγκεφαλοσκόπιου (όπως θα συνέβαινε αν δεχόμασταν την ΑΥΘΝ). Θα ονομάσουμε την περίπτωση (α) "καθημερινή οπτική του τρίτου προσώπου" και την περίπτωση (β) "επιστημονική"⁴⁰⁷ οπτική του τρίτου προσώπου". Οι δύο αυτές οπτικές διαφέρουν όχι μόνο από την οπτική του πρώτου προσώπου αλλά και μεταξύ τους. Τόσο η καθημερινή οπτική του τρίτου προσώπου όσο και η οπτική του πρώτου προσώπου αφορούν αυτό που θα λέγαμε "προσωπικό" επίπεδο. Με άλλα λόγια, οι δύο αυτές οπτικές έχουν ως έσχατη μονάδα ανάλυσης το υποκείμενο στο σύνολό του. Αντίθετα, η επιστημονική οπτική του τρίτου προσώπου αφορά το υποπροσωπικό επίπεδο, είτε αυτό νοείται ως λειτουργικό επίπεδο είτε ως νευροφυσιολογικό. Στα επόμενα, θα εστιάσουμε την προσοχή μας στις διαφορές της οπτικής του πρώτου προσώπου από την επιστημονική οπτική του τρίτου προσώπου (δηλαδή, τον τρόπο που η ΑΥΘΝ επιτρέπει ότι θα μπορούσαν να προσεγγιστούν τα νοητικά φαινόμενα).

Στο μεθοδολογικό επίπεδο θα δείξουμε ότι ακόμα και αν δεχόμασταν ως ορθή την υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων και, κατά συνέπεια, δεχόμασταν την ορθότητα της επιστημονικής οπτικής του τρίτου προσώπου, πάλι θα υπήρχαν συγκροτητικά χαρακτηριστικά των νοητικών φαινομένων -όπως αυτά αποκαλύπτονται από την οπτική του πρώτου προσώπου- τα οποία δεν θα μπορούσαν να δοθούν από την επιστημονική οπτική του τρίτου προσώπου. Για το σκοπό αυτό θα προβούμε σε μια συστηματική σύγκριση της επιστημονικής οπτικής του τρίτου προσώπου με την οπτική του πρώτου

⁴⁰⁷ Το ότι αποκαλούμε αυτή την οπτική "επιστημονική" δεν σημαίνει ότι θεωρούμε ότι κάθε επιστημονική προσέγγιση γίνεται από την επιστημονική οπτική του τρίτου προσώπου. Για παράδειγμα, η ψυχαναλυτική προσέγγιση δεν αποτελεί προσέγγιση από την επιστημονική οπτική του τρίτου προσώπου. Στο παρόν κεφάλαιο, ο όρος "επιστημονική οπτική του τρίτου προσώπου" αφορά αποκλειστικά την οπτική της γνωσιακής επιστήμης που βασίζεται στην ΑΥΘΝ.

προσώπου. Εφόσον δε η επιστημονική οπτική του τρίτου προσώπου, όπως την ορίσαμε, προϋποθέτει την ορθότητα της ΑΥΘΝ, μπορούμε να δεχθούμε και την δυνατότητα κατασκευής της συσκευής που ονομάσαμε "εγκεφαλοσκόπιο". Έτσι, προκειμένου να συγκρίνουμε τις δύο προαναφερθείσες οπτικές θα συγκρίνουμε πώς το ίδιο νοητικό φαινόμενο θα δίδεται μέσω του εγκεφαλοσκοπίου και της οπτικής του πρώτου προσώπου.

Πριν ξεκινήσουμε, ωστόσο, αυτή τη σύγκριση οφείλουμε να απαντήσουμε σε μια εύλογη ένσταση: γιατί θα πρέπει να δώσουμε κάποια προτεραιότητα στην οπτική του πρώτου προσώπου έναντι της επιστημονικής οπτικής του τρίτου προσώπου; Γιατί, με άλλα λόγια, θεωρούμε εκ των προτέρων ότι τα χαρακτηριστικά των νοητικών φαινομένων που αποκαλύπτονται από την οπτική του πρώτου προσώπου και, ταυτόχρονα, διαφεύγουν από την επιστημονική οπτική του τρίτου προσώπου αποτελούν συγκροτητικά χαρακτηριστικά των νοητικών φαινομένων;

Καταρχάς, πρέπει να τονίσουμε ότι η επιστημονική οπτική του τρίτου προσώπου αποτελεί απλά μια υπόθεση εργασίας: αυτή η οπτική ενδέχεται να υπάρξει κάποτε μόνο αν η ΑΥΘΝ αποδειχθεί σωστή, δηλαδή μόνο αν υπάρχουν όντως νοητικές αναπαραστάσεις. Μάλιστα, όπως δείξαμε σε προηγούμενα κεφάλαια, η υπόθεση περί ύπαρξης νοητικών αναπαραστάσεων αντιμετωπίζει ανυπερβλήτα προβλήματα που την καθιστούν εντελώς αμφίβολη, αν όχι ανυπόστατη. Από την άλλη μεριά, η οπτική του πρώτου προσώπου είναι μια πραγματικότητα: όλοι μας ζούμε το μεγαλύτερο μέρος της ζωής μας σε καθεστώς συνείδησης, δηλαδή από την οπτική του πρώτου προσώπου. Με άλλα λόγια, η οπτική του πρώτου προσώπου αποτελεί τον πρωταρχικό και πιο άμεσο τρόπο με τον οποίο ερχόμαστε σε επαφή με τα νοητικά φαινόμενα. Τα προς εξήγηση φαινόμενα, δηλαδή, είναι αυτά που μας δίδονται από την οπτική του πρώτου προσώπου και κάθε προσέγγιση με βάση μια διαφορετική οπτική οφείλει να εξηγήσει αυτά ακριβώς τα φαινόμενα.

Φυσικά, το γεγονός ότι επικαλούμαστε το πρωτείο της οπτικής του πρώτου προσώπου δεν σημαίνει ότι διαθέτουμε και μια καθαρή εικόνα για το πώς ακριβώς "λειτουργεί" η ίδια η οπτική του πρώτου προσώπου. Εκείνο που είναι φανερό σε μας είναι, απλά, η φαινομενολογία των νοητικών φαινομένων, δηλαδή το τι μας δίδεται μέσω της οπτικής του πρώτου προσώπου. Στρηγίζόμενοι σε αυτήν τη φαινομενολογία θα επιχειρήσουμε ταυτόχρονα και μια κριτική της κυρίαρχης -καρτεσιανής προέλευσης- αντίληψης περί της οπτικής του πρώτου προσώπου, σύμφωνα με την οποία η συγκεκριμένη οπτική κατανοείται ως ενδοσκόπηση.

8.2.1 Υποκειμενικότητα

Ας ξεκινήσουμε, λοιπόν, τη σύγκριση των δύο οπτικών. Θα εστιάσουμε το ενδιαφέρον μας στις *συνειδητές* προτασιακές στάσεις και αρχικά στην περίπτωση της αντίληψης. Η πρώτη βασική διαφορά που

διαπιστώνουμε είναι ότι, ενώ από την οπτική του πρώτου προσώπου τα νοητικά φαινόμενα εμφανίζονται ως ουσιαστές ανήκοντα σε κάποιον, από την επιστημονική οπτική του τρίτου προσώπου δεν παρουσιάζουν αυτό το χαρακτηριστικό. Όταν βλέπω ένα δένδρο τότε γνωρίζω άμεσα ότι εγώ είμαι αυτός που το βλέπει, δηλαδή γνωρίζω άμεσα σε ποιον ανήκει η συγκεκριμένη αντίληψη. Αντίθετα, ο νευροφυσιολόγος που παρατηρεί με το εγκεφαλοσκόπιο τη νοητική μου κατάσταση δεν μπορεί να μάθει σε ποιον ανήκει η παρατηρούμενη νοητική αναπαράσταση βασίζόμενος αποκλειστικά σε ό,τι βλέπει μέσω του εγκεφαλοσκόπιου. Θα ονομάσουμε προσωρινά αυτό το χαρακτηριστικό των νοητικών φαινομένων που δίδεται από την οπτική του πρώτου προσώπου και αποκρύπτεται από την επιστημονική οπτική του τρίτου προσώπου "άμεση γνώση της πατρότητας"⁴⁰⁸ των νοητικών φαινομένων".

8.2.1.1 Μπορεί η γνώση της πατρότητας να εξηγηθεί από την επιστημονική οπτική του τρίτου προσώπου;

8.2.1.1.1 Αποτελεί η γνώση της πατρότητας μια αναπαραστασιακή μορφή γνώσης;

Τι ακριβώς μου δίδεται ως δικό μου από την οπτική του πρώτου προσώπου; Σίγουρα, όχι το αναπαραστασιακό περιεχόμενο. Το δένδρο δεν μου εμφανίζεται ως δικό μου, αλλά ως ένα αντικείμενο του κόσμου. Αυτό που μου εμφανίζεται ως δικό μου είναι η προτασιακή στάση με την οποία προσεγγίζω το αντικείμενο, το ότι, δηλαδή, εγώ είμαι αυτός που βλέπει το δένδρο.

Η συγκεκριμένη παρατήρηση δίνει και μια προφανή απάντηση στο γιατί ο νευροφυσιολόγος, όταν παρατηρεί μέσα από τη συσκευή τη νοητική μου αναπαράσταση, δεν καταλαβαίνει σε ποιον ανήκει. Αυτό συμβαίνει, απλά, γιατί δεν παρατηρεί την προτασιακή στάση στο σύνολό της, αλλά μόνο το νοητικό περιεχόμενο της στάσης όπως δίδεται από την νοητική αναπαράσταση.

Βέβαια, θα μπορούσε κανείς να αντιτείνει ότι στην ΑΥΘΝ υπάρχει θέση και για τις προτασιακές στάσεις, μόνο που αυτές δεν "τοποθετούνται μέσα" στις νοητικές αναπαραστάσεις. Το εάν ένας οργανισμός πιστεύει ότι p ή επιθυμεί ότι p καθορίζεται από το είδος της υπολογιστικής σχέσης που υπάρχει, στο λειτουργικό επίπεδο, με τη νοητική αναπαράσταση που σημαίνει ότι p . Με άλλα λόγια, το είδος της προτασιακής στάσης δεν καθορίζεται μέσα στο πλαίσιο της νοητικής αναπαράστασης. Κατά συνέπεια, είναι άδικο το να καταλογίζουμε στο νευροφυσιολόγο ότι με την συσκευή του δεν μπορεί να παρατηρήσει

⁴⁰⁸ Η, εναλλακτικά, της κυριότητας. Βλέπε, υποσημείωση 423.

προτασιακές στάσεις. Θα μπορούσαμε, μάλιστα, να φανταστούμε μια βελτιωμένη εκδοχή της συσκευής η οποία ανιχνεύει, μαζί με την εκάστοτε νοητική αναπαράσταση, και το είδος της υπολογιστικής σχέσης στην οποία η αναπαράσταση χρησιμοποιείται. Σε αυτή την περίπτωση, θα μπορούσε κανείς να ισχυριστεί ότι ο νευροφυσιολόγος, κυριολεκτικά, παρατηρεί προτασιακές στάσεις.

Παρακάμπτοντας, προς στιγμήν, αντιρρήσεις περί της δυνατότητας κατασκευής μιας τέτοιας συσκευής (δεδομένης και της θέσης του Fodor περί ολιστικής λειτουργίας των κεντρικών συστημάτων⁴⁰⁹), πρέπει να παρατηρήσουμε ότι το μόνο που θα αποδείκνυε η κατασκευή της δεν λόγω συσκευής θα ήταν ότι ο νευροφυσιολόγος θα είχε, απλά, την επιπρόσθετη ικανότητα να καθορίζει το είδος της προτασιακής στάσης που έχει το εξεταζόμενο υποκείμενο: το αν δηλαδή θυμάται, φαντάζεται, βλέπει, πιστεύει κ.λπ. την εκάστοτε νοητική αναπαράσταση. Τίποτε, ωστόσο, απ' όσα θα παρατηρούσε δεν θα του επέτρεπε να προσδιορίσει σε ποιον ανήκει η παρατηρούμενη προτασιακή στάση. Αντίθετα, από την οπτική του πρώτου προσώπου, όταν έχω μια προτασιακή στάση, γνωρίζω άμεσα ότι αυτή είναι δικιά μου. Άρα, λοιπόν, αυτή η γνώση της πατρότητας των νοητικών φαινομένων είναι κάτι που διαφεύγει από την επιστημονική οπτική του τρίτου προσώπου, η οποία, σε πρώτη προσέγγιση, μπορεί να θεωρηθεί η "οπτική" του εγκεφάλου. Κατά συνέπεια, αν δεχθεί κανείς την υπόθεση ότι τα νοητικά φαινόμενα δεν είναι τίποτε άλλο από νοητικές αναπαραστάσεις και υπολογιστικές σχέσεις μεταξύ αυτών, τότε δεν μπορεί να εξηγήσει τη γνώση της πατρότητας των νοητικών φαινομένων όπως αυτή δίδεται από την οπτική του πρώτου προσώπου.

Όμως, ένα τέτοιο συμπέρασμα μοιάζει πρόωρο: θα μπορούσαμε να φανταστούμε ένα ακόμα πιο βελτιωμένο μοντέλο του εγκεφαλοσκόπου το οποίο, εκτός από το περιεχόμενο και το είδος της προτασιακής στάσης, θα αναγράφει και το όνομα του φορέα του⁴¹⁰. Σε αυτή την περίπτωση, δεν θα ήταν σωστό να πούμε ότι ο νευροφυσιολόγος γνωρίζει την πατρότητα της παρατηρούμενης προτασιακής στάσης;

Η συγκεκριμένη αντίρρηση δεν αναιρεί το συμπέρασμα στο οποίο καταλήξαμε, αλλά, απλά, μας υποχρεώνει να το διασαφηνίσουμε. Ο νευροφυσιολόγος δεν θα γνώριζε την πατρότητα της προτασιακής στάσης κατά τον τρόπο που τη γνωρίζει το ίδιο το υποκείμενο που έχει τη προτασιακή στάση. Η γνώση της πατρότητας των προτασιακών στάσεων – όπως αυτή δίδεται από την οπτική του πρώτου προσώπου – έχει μια αμεσότητα που δεν απαντάται στον τρόπο που θα γνωρίζαμε την πατρότητα των προτασιακών στάσεων από την οπτική του εγκεφαλοσκόπου. Για παράδειγμα, αν ο μελλοντικός νευροφυσιολόγος

⁴⁰⁹ Βλέπε Κεφάλαιο 2, ενότητα 2.6.1.

⁴¹⁰ Για παράδειγμα, μέσω ανάλυσης του DNA.

έστρεψε το εγκεφαλοσκόπιο στον ίδιο του τον εγκέφαλο, τότε ο τρόπος που θα γνώριζε την πατρότητα των παρατηρούμενων προτασιακών στάσεων μέσω του εγκεφαλοσκόπιου θα ήταν διαφορετικός από τον τρόπο που θα την γνώριζε από την οπτική του πρώτου προσώπου. Η διαφορά γίνεται φανερή αν σκεφτούμε την περίπτωση ενός αμνησιακού νευροφυσιολόγου -ας τον ονομάσουμε X. Ας υποθέσουμε ότι ενόσω ο X κοιμάται τον δένουμε μπροστά στην οθόνη του εγκεφαλοσκόπιου και παράλληλα συνδέουμε τον εγκέφαλό του με τη συσκευή. Όταν ο X ξυπνήσει θα διαπιστώσει ότι βλέπει μέσω του εγκεφαλοσκόπιου τις προτασιακές στάσεις κάποιου ονόματι X, αλλά σίγουρα δεν θα καταλάβει άμεσα ότι οι παρατηρούμενες προτασιακές στάσεις είναι δικές του⁴¹¹. Αλλά, ακόμα και αν στην θέση του ονόματος του X εμφανιζόταν οποιαδήποτε καθοριστική περιγραφή (definite description) του X, πάλι ο X δεν θα καταλάβαινε άμεσα ότι πρόκειται για δικιά του προτασιακή στάση⁴¹². Αντίθετα, από την οπτική του πρώτου προσώπου ο X, όταν κοιτάει την οθόνη του εγκεφαλοσκόπιου, γνωρίζει άμεσα ότι αυτός είναι που την κοιτάει -ανεξάρτητα από το αν δεν θυμάται τίποτα για τον εαυτό του και την ιστορία του.

Ωστόσο, ο νευροφυσιολόγος θα μπορούσε, ενδεχομένως, να αποκτήσει γνώση της πατρότητας με έναν διαφορετικό τρόπο: θα μπορούσε να τροποποιήσει το εγκεφαλοσκόπιο ούτως ώστε αντί του ονόματος ή κάποιας καθοριστικής περιγραφής να παρουσιάζεται η εξωτερική εμφάνιση του παρατηρούμενου υποκειμένου. Δεν θα ήταν αυτός ο καλύτερος τρόπος προκειμένου ο νευροφυσιολόγος να γνωρίζει άμεσα την πατρότητα των παρατηρούμενων προτασιακών στάσεων; Το πρόβλημα και σε αυτήν την περίπτωση είναι ότι ο νευροφυσιολόγος δεν θα μπορούσε να συλλάβει τη γνώση της πατρότητας κατά τον τρόπο που το ίδιο το παρατηρούμενο υποκείμενο γνωρίζει ότι μια νοητική κατάσταση είναι δικιά του. Αν, για παράδειγμα, ο νευροφυσιολόγος έπασχε από μια ακραία μορφή προσωποαγνωσίας και κοιτούσε σε ένα εγκεφαλοσκόπιο χωρίς να γνωρίζει ότι αυτό είναι συνδεδεμένο με τον εγκέφαλό του, σίγουρα δεν θα καταλάβαινε άμεσα ότι οι παρατηρούμενες νοητικές καταστάσεις είναι δικές του⁴¹³. Από την άλλη μεριά, ο

⁴¹¹ Θα χρειαστεί ενδεχομένως κάποια λεπτά ή δευτερόλεπτα για να συναγάγει από το περιεχόμενο των παρατηρούμενων νοητικών καταστάσεων ότι αυτές είναι δικές του.

⁴¹² Πρβλ. Shoemaker (1992, σελ. 468): "κανείς μπορεί να αναφερθεί στον ίδιο με το 'εγώ' χωρίς να γνωρίζει καμιά περιγραφή που θα μπορούσε να χρησιμοποιηθεί για την παγίωση της αναφοράς του".

⁴¹³ Πρβλ., επίσης, Wittgenstein 1984, σελ. 105: "Είναι δυνατόν σ' ένα δυστύχημα, ας πούμε, να νιώσω πόνο στο χέρι μου, να δω ένα σπασμένο χέρι δίπλα μου και να το πάρω για δικό μου, ενώ στην πραγματικότητα είναι του διπλανού μου. Και θα μπορούσα, κοιτώντας σ' έναν καθρέφτη, να πάρω ένα καρούμπαλο στο κούτελό του για δικό μου".

προσωποαγνωσιακός που, μη γνωρίζοντας ότι κοιτάει σε έναν καθρέπτη, ρωτά τον σερβιτόρο ποιος είναι αυτός απέναντι που τον κοιτά επίμονα⁴¹⁴, σίγουρα δεν έχει καμιά αμφιβολία για το ποιος βλέπει τον άγνωστο άνδρα.

Τα παραπάνω παραδείγματα και, κυρίως, το παράδειγμα του αμνησιακού δείχνουν ότι το πρόβλημα της γνώσης της πατρότητας δεν πρέπει να συγχέεται με το πρόβλημα της προσωπικής ταυτότητας, δηλαδή της ενότητας του εαυτού μέσα στο χρόνο. Το πρόβλημα της γνώσης της πατρότητας μιας νοητικής κατάστασης είναι εντοπισμένο χρονικά γιατί, όταν κανείς βρίσκεται σε μια νοητική κατάσταση, κάθε χρονική στιγμή γνωρίζει άμεσα ότι η συγκεκριμένη νοητική κατάσταση είναι δική του. Ωστόσο, το πρόβλημα της γνώσης της πατρότητας δεν πρέπει να εξισωθεί ούτε με το πρόβλημα της ενότητας του εαυτού σε μια χρονική στιγμή⁴¹⁵, αλλά πρέπει να θεωρηθεί ως ένα υποπρόβλημα του τελευταίου. Το πρόβλημα της ενότητας του εαυτού σε μια χρονική στιγμή παραμένει ανοικτό τόσο για τη φιλοσοφία όσο και για τη ψυχολογία και επίσης θεωρείται στενά συνδεδεμένο με το πρόβλημα της ενότητας του εαυτού μέσα στον χρόνο. Για το λόγο αυτό δεν θα προβούμε παρά σε μια σχεδόν ταυτολογική διαπίστωση: Όταν κάποιος χάνει εντελώς⁴¹⁶ την ενότητα του εαυτού, παύει να υφίσταται ο υποκειμενικός πόλος στον οποίον μπορούν να ανήκουν τα νοητικά φαινόμενα. Ενδεχομένως μια τέτοια διανοητική κατάσταση απαντάται κατά τη διαδικασία της αποπροσωποποίησης στη σχιζοφρένια και πιο συγκεκριμένα κατά τη φάση της εκμηδένισης⁴¹⁷. Αν

⁴¹⁴ Βλ. Sacks 1990, σελ. 39.

⁴¹⁵ Το πρόβλημα αυτό ονομάζεται συχνά "πρόβλημα της ενότητας της συνείδησης" (βλ., για παράδειγμα, Kolak & Martin 1991).

⁴¹⁶ Λέμε "εντελώς" γιατί δεν θέλουμε να συμπεριλάβουμε περιπτώσεις πολλαπλών προσωπικοτήτων.

⁴¹⁷ "Αποπροσωποποίηση: Συναίσθημα απώλειας του εαυτού που δοκιμάζουν ορισμένα άτομα, [είτε] ως προς τη σωματική και οργανική τους ακεραιότητα, [είτε] ως προς το σύνολο αυτών των συστατικών στοιχείων της προσωπικότητας." *Αλφαριθμητικό Εγχειρίδιο Ψυχιατρικής* Α. Porot (αναφέρεται στο Καγγελάρης 1985, σελ. 11). Χαρακτηριστική είναι η περιγραφή μιας ασθενούς για το πώς ένοιωθε κατά τη φάση της εκμηδένισης που αποτελεί ένα στάδιο της διαδικασίας της αποπροσωποποίησης: "Είναι τρομερό, ένοιωθα να χάνομαι μέσα σ' ένα λευκό καπνό, οι ήχοι δεν ακούγονταν πια, υπήρχαν μόνο ουρλιαχτά χωρίς ήχο. Τα χρώματα υψώνονταν από τα πράγματα σαν πολύχρωμα σύννεφα, απομακρυνόμενα, δεν είχα δύναμη να φωνάξω, χανόμενα, εξαφανιζόμενα, γινόμουν ένα με τα χρώματα και τους ήχους, χανόμενα στην ομίχλη, όλα γίνονταν γκριζα, κι εσύ ήσουνα ένα μέρος της ομίχλης. Έβλεπα την κεραία της τηλεόρασης, την κρατούσα μπροστά μου ενώ ο καπνός ανέβαινε. Όταν όλα είχαν εξαφανιστεί, χαμένη μέσα στην ομίχλη έβλεπα ακόμα την άκρη της κεραίας, ώσπου ξαφνικά χάθηκε κι αυτή. Ήμουν εκεί, πού όμως; Δεν είχα πεθάνει, ένοιωθα σαν ένα μικρό συννεφάκι καπνού, όλα γύρω μου ήταν μικρά σύννεφα καπνού. Δεν ήταν ακριβώς σύννεφα, δεν ξέρω πως να το πω. Καταλαβαίνετε; Μέχρι τη στιγμή που τα σύννεφα χάθηκαν το ένα μέσα σ' άλλο, δεν υπήρχε παρά καπνός. Μάντευα

δεχθούμε ότι υπάρχει δυνατότητα μιας τέτοιας καθολικής απώλειας της ενότητας του εαυτού, τότε είναι εύλογο να υποθέσουμε ότι σε αυτή την περίπτωση το "υποκειμένο" είναι αδύνατο να γνωρίζει άμεσα ότι οι (όποιες) προτασιακές στάσεις ανήκουν στο ίδιο.

Ισχύει άραγε και το αντίστροφο; Συνεπάγεται η απώλεια της άμεσης γνώσης της πατρότητας την καθολική απώλεια του εαυτού; Κάτι τέτοιο ισχύει μόνο αν πρόκειται για την απώλεια της άμεσης γνώσης της πατρότητας κάθε νοητικού φαινομένου. Διαφορετικά, μπορούμε να φανταστούμε την περίπτωση ενός υποκειμένου που ενώ διατηρεί την αίσθηση του εαυτού, δηλαδή μια αίσθηση ενός ενιαίου εγώ, του εμφανίζονται κάποιες μεμονωμένες σκέψεις, πεποιθήσεις, επιθυμίες κ.λπ. που δεν τις αντιλαμβάνεται άμεσα ως δικές του. Σε αυτή την περίπτωση το υποκειμένο, ενώ έχει αυτό το ίδιο τις συγκεκριμένες προτασιακές στάσεις και έχει συνείδηση τόσο του περιεχομένου όσο και του είδους τους, δεν τις αντιλαμβάνεται ως δικές του. Εμπειρικά μια τέτοια διανοητική κατάσταση είναι, ενδεχομένως, ορισμένες παραισθήσεις των σχιζοφρενών όπως η παρεμβολή σκέψεων (thought insertion): μια φωνή που προστάζει το σχιζοφρενή να κάνει κάτι είναι, ίσως, μια πρόθεση του ίδιου του σχιζοφρενή από την οποία έχει "απολεσθεί" η άμεση γνώση της πατρότητάς της⁴¹⁸.

Κατά συνέπεια, η ενότητα του εαυτού σε μια χρονική στιγμή αποτελεί προϋπόθεση για την άμεση γνώση της πατρότητας των προτασιακών στάσεων. Μέχρι στιγμής, λοιπόν, έχουμε δει ότι η γνώση της πατρότητας μιας νοητικής κατάστασης από την οπτική του πρώτου προσώπου δεν συνίσταται στη γνώση του ονόματος, της εξωτερικής εμφάνισης ή οποιασδήποτε καθοριστικής περιγραφής του υποκειμένου που φέρει τη νοητική κατάσταση. Ο νευροφυσιολόγος, ενώ βλέπει μέσα από το γκεφαλοσκόπιο το περιεχόμενο και το είδος της προτασιακής στάσης μαζί με το όνομα ή μια καθοριστική περιγραφή ή, τέλος, μια εικόνα του εαυτού του, είναι δυνατόν να μην καταλάβει άμεσα ότι η προτασιακή στάση που παρατηρεί είναι δικιά του -πράγμα που δεν ισχύει όταν "προσεγγίζει" το ίδιο νοητικό φαινόμενο από την οπτική του πρώτου

που βρίσκονταν τα πράγματα πριν. Μετά τίποτα... πιστεύω πως πα δεν φοβόμουν, δεν ξέρω πώς να το πω, όλα είχαν τελειώσει" (παρατίθεται από τον Καγγελάρη 1985, σελ. 239-240). Είναι ενδιαφέρον να επισημάνουμε ότι η αφήγηση της ασθενούς γίνεται από την οπτική του πρώτου προσώπου. Παρόλα αυτά, η αυξανόμενη αδυναμία της στο να βρει επαρκείς εκφράσεις φαίνεται να είναι εκείνη το στοιχείο που δηλώνει την προϊούσα εξαφάνιση του εγώ (η οποία συμβαίνει ταυτόχρονα με την προϊούσα εξαφάνιση του κόσμου).

⁴¹⁸ Πρβλ. Snyder 1974, σελ. 121: "Οι φωνές είναι ακριβώς οι ίδιες οι σκέψεις του ασθενούς τις οποίες επέλεξε, κατά πάσα πιθανότητα, χωρίς συνειδητή επίγνωση, να προβάλει στον εξωτερικό κόσμο" (αναφέρεται στο Graham & Stephens 1994a, σελ. 95). Βλ., επίσης, Campbell 1999. Θα επανέλθουμε σε αυτό το ζήτημα αναλυτικά σε επόμενες ενότητες του παρόντος κεφαλαίου.

προσώπου. Συναφές εδώ είναι και το παράδειγμα των εκ γενετής τυφλών: το γεγονός ότι δεν γνωρίζουν παρά ελάχιστα για την προσωπική τους εμφάνιση δεν επηρεάζει στο παραμικρό την ικανότητά τους να γνωρίζουν άμεσα την πατρότητα των νοητικών τους καταστάσεων. Βέβαια, η επισήμανση αυτή δεν σημαίνει ότι και από γενετικής απόψεως ο εκ γενετής τυφλός φτάνει με την ίδια ευκολία στην απόκτηση αυτής της ικανότητας (βλ. Hobson 1993). Εδώ πρέπει να διακρίνουμε το οντογενετικό ερώτημα σχετικά με το πώς φτάνουμε να αποκτήσουμε την ικανότητα να έχουμε μια άμεση γνώση της πατρότητας των νοητικών καταστάσεων από το ερώτημα που μας ενδιαφέρει στην παρούσα ενότητα, δηλαδή το ερώτημα σχετικά με το ποιο είναι το περιεχόμενο αυτής της γνώσης. Μια απάντηση στο πρώτο ερώτημα δεν συνιστά απάντηση και του δεύτερου ερωτήματος⁴¹⁹.

Συμπερασματικά, η γνώση της πατρότητας που έχει ένα υποκείμενο Α για μια νοητική κατάσταση στην οποία βρίσκεται χαρακτηρίζεται από μια αμεσότητα η οποία δεν μπορεί να συλληφθεί μέσω κάποιας εικονιστικής ή συμβολικής αναπαράστασης του υποκειμένου Α. Μπορούμε, λοιπόν, εύλογα να συνάγουμε ότι η άμεση γνώση της πατρότητας μιας προτασιακής στάσης δεν είναι μια αναπαραστασιακή μορφή γνώσης.

8.2.1.1.2 Αποτελεί η γνώση της πατρότητας μια μορφή άμεσης αντίληψης;

Εξάλλου, η ίδια η έννοια της άμεσης γνώσης δεν συμβιβάζεται με την έννοια της αναπαραστασιακής γνώσης, διότι "άμεση", εν προκειμένω, φαίνεται να σημαίνει "αδιαμεσολάβητη" και "μη συναγωγική" (non-inferential). Έτσι, η άμεση γνώση ενός πράγματος μπορεί να κατανοηθεί ως μια άμεση, αδιαμεσολάβητη, μη συναγωγική πρόσβαση στο πράγμα, δηλαδή ως μια πρόσβαση που δεν βασίζεται σε αναπαραστάσεις του πράγματος. Μια τέτοια γνώση είναι και η άμεση, αδιαμεσολάβητη αντίληψη⁴²⁰.

Θα μπορούσε, λοιπόν, κανείς να ισχυρισθεί ότι αυτό που, από την οπτική του πρώτου προσώπου, ονομάζουμε "άμεση γνώση της πατρότητας μιας προτασιακής στάσης", από την επιστημονική οπτική του τρίτου προσώπου δεν είναι τίποτε άλλο παρά ο εντοπισμός κάποιου είδους αντικειμένου (και όχι κάποιας αναπαράστασης) στον εγκέφαλο. Πρόκειται, με άλλα λόγια, για μια μορφή άμεσης αντίληψης, εφόσον αυτό που εντοπίζεται είναι κάποιο αντικείμενο -ή, έστω, κατάσταση του

⁴¹⁹ Πρβλ. Lacan 1949/1984.

⁴²⁰ Εδώ υποθέτουμε ότι υπάρχει μια άμεση και αδιαμεσολάβητη αντίληψη, δεχόμεστε, με άλλα λόγια, ότι η αντίληψη δεν είναι αναπαραστασιακή. Αυτή η παραδοχή θα υποστηριχθεί στην ενότητα 8.2.2.2.

εγκεφάλου⁴²¹ - και όχι μια αναπαράσταση αντικείμενου. Ένα τέτοιο αντικείμενο θα μπορούσε, δύναμει, να εντοπισθεί με το εγκεφαλοσκόπιο. Έτσι, η άμεση γνώση της πατρότητας μιας προτασιακής στάσης από την οπτική του πρώτου προσώπου θα ισοδυναμούσε με την παρουσία του συγκεκριμένου αντικείμενου στο νευροφυσιολογικό επίπεδο⁴²². Σε μια τέτοια περίπτωση, η άμεση γνώση της πατρότητας των προτασιακών στάσεων δεν θα αποτελούσε κάποιο χαρακτηριστικό των νοητικών φαινομένων που, ενώ μας δίδεται από την οπτική του πρώτου προσώπου, τελικά διαφεύγει από την επιστημονική οπτική του τρίτου προσώπου.

Η ένσταση αυτή είναι πολλαπλώς προβληματική. Καταρχάς, ακόμα και αν υπήρχε ένα τέτοιο εσωτερικό αντικείμενο στον οργανισμό που η παρουσία του δημιουργούσε στο φαινομενολογικό επίπεδο αυτό που ονομάσαμε "άμεση γνώση της πατρότητας μιας προτασιακής στάσης", ο εντοπισμός του συγκεκριμένου αντικείμενου από τον νευροφυσιολόγο μέσω του εγκεφαλοσκόπου σίγουρα δεν θα συνοδευόταν από μια παρόμοια άμεση γνώση της πατρότητας του αντικείμενου. Απεναντίας, είναι μάλλον προφανές ότι ο νευροφυσιολόγος, ακόμα και αν παρατηρούσε το δικό του εγκέφαλο, και εντόπιζε αυτό το αντικείμενο ή την κατάσταση που ευθύνεται για την άμεση γνώση της πατρότητας των νοητικών του ενεργημάτων στο φαινομενολογικό επίπεδο, πάλι δεν θα είχε μια άμεση γνώση της πατρότητας του παρατηρούμενου αντικείμενου.

Μια μικρή ένσταση εδώ είναι ότι ο νευροφυσιολόγος όταν παρατηρεί το συγκεκριμένο αντικείμενο δεν το παρατηρεί άμεσα, αλλά μέσω των ενδείξεων του εγκεφαλοσκόπου, άρα, λοιπόν, δεν πρόκειται πραγματικά για κάποια περίπτωση άμεσης αντίληψης. Η ένσταση αυτή, παρότι σωστή, δεν αναιρεί τη θέση ότι η άμεση γνώση της πατρότητας μιας προτασιακής στάσης δεν μπορεί να εξηγηθεί με βάση την έννοια της άμεσης αντίληψης ενός αντικείμενου. Διότι ακόμα και αν, *per impossibile*, ο νευροφυσιολόγος κατόρθωνε να παρατηρήσει άμεσα -χωρίς το εγκεφαλοσκόπιο - το συγκεκριμένο εγκεφαλικό αντικείμενο, πάλι δεν θα είχε μια άμεση γνώση της πατρότητας αυτού. Ένα αντικείμενο, εξάλλου, μπορούμε να το αντιλαμβανόμαστε άμεσα χωρίς να το αντιλαμβανόμαστε ως δικό μας ή, έστω, ως ανήκον σε κάποιον άλλον. Ακόμα, όμως, και αν το αντιλαμβανόμαστε ως δικό μας, αυτή η έννοια πατρότητας-κυριότητας⁴²³ σίγουρα είναι πολύ διαφορετική από την πατρότητα-

⁴²¹ Στα επόμενα, χάριν συντομίας, δεν θα επαναλαμβάνουμε αυτή την εναλλακτική περίπτωση και θα κάνουμε λόγο μόνο για κάποιο αντικείμενο.

⁴²² Ορισμένοι θα την ονόμαζαν "νευρωνικό σύστοιχο της συνείδησης" (*neural correlate of consciousness*). Υπάρχει μια διαρκώς ογκούμενη βιβλιογραφία γύρω από αυτό το ζήτημα κυρίως μετά τη δεκαετία του '90. Ειδικότερα, στο περιοδικό *Journal of Consciousness Studies* μπορεί κανείς να βρει πολλές διαφορετικές προτάσεις για το εν λόγω ζήτημα.

⁴²³ Σε αυτό το πλαίσιο είναι καταλληλότερο να μιλήσει κανείς για κυριότητα (*ownership*) και όχι για πατρότητα. Η πατρότητα ενός αντικείμενου αφορά το

κυριότητα μιας προτασιακής στάσης. Όταν αντιλαμβανόμαστε ένα αντικείμενο, ακόμα και αν είναι δικό μου, το νοώ πάντα ως κάτι *έτερο*, ως κάτι με το οποίο έχω μια σχέση εξωτερικότητας. Γι' αυτό, *άλλωστε*, και ένα αντικείμενο μπορεί να πάψει να είναι δικό μου. Αντίθετα, μια νοητική κατάσταση, εφόσον μου ανήκει, είναι αδύνατον να "μεταβιβαστεί", διότι μου ανήκει κατά αναγκαίο τρόπο. Έχουμε λοιπόν εδώ να κάνουμε με δύο εντελώς διαφορετικές έννοιες πατρότητας - κυριότητας: την πατρότητα - κυριότητα του *έτερου* και την πατρότητα - κυριότητα του *ιδίου*. Όταν αναγνωρίζω κάτι *έτερο* προς εμένα ως δικό μου, τότε μπορώ κάλλιστα να κάνω λάθος: μπορώ, για παράδειγμα, να θεωρήσω ότι το ποδήλατο που βλέπω απέναντί μου είναι το δικό μου ποδήλατο και λίγο αργότερα να διαπιστώσω ότι έκανα λάθος. Αντίθετα, όταν βρίσκομαι σε μια νοητική κατάσταση, παρότι μπορώ να αμφιβάλω για το ακριβές της περιεχόμενο, είναι αδύνατον να αμφιβάλω για το ποιος βρίσκεται σε αυτή την κατάσταση. Γι' αυτό και η έκφραση "κάποιος πιστεύει ότι θα βρέξει αλλά δεν είμαι σίγουρος αν αυτός είμαι εγώ" δεν έχει νόημα⁴²⁴ παρά μόνο στο στόμα ανθρώπων με πρόβλημα πολλαπλών προσωπικοτήτων. Από την οπτική του πρώτου προσώπου, η γνώση της πατρότητας μιας νοητικής κατάστασης είναι μια αλάνθαστη γνώση: στους φυσιολογικούς ανθρώπους δεν υπάρχει περιθώριο λανθασμένης αναγνώρισης του ποιος έχει τη νοητική κατάσταση⁴²⁵. Να επιστημόνουμε σε αυτό το σημείο ότι η θέση στην οποία αναφερόμαστε δεν πρέπει να συγχέεται με τις καρτεσιανές θέσεις της μη-διορθωσιμότητας (incorrigibility) και της προνομιακής πρόσβασης (privileged access) που υποτίθεται ότι χαρακτηρίζουν την

ποιος το παρήγαγε, ενώ η κυριότητα το ποιος το κατέχει. Ούτε η πατρότητα συνεπάγεται κατ' ανάγκη την κυριότητα, αλλά ούτε και το αντίστροφο. Στην περίπτωση των νοητικών φαινομένων, αντίθετα, αν κανείς "παράγει" μια προτασιακή στάση, τότε κατ' ανάγκη είναι κύριος της προτασιακής στάσης και το αντίστροφο. Για περισσότερα επ' αυτής της διάκρισης βλέπε την ενότητα 8.2.1.4 όπου αναφέρεται μια παθολογική περίπτωση προτασιακής στάσης στην οποία η γνώση της πατρότητας μπορεί να αποσυνδεθεί από τη γνώση της κυριότητας.

⁴²⁴ Βλ. Wittgenstein 1984, σελ. 105: "...Ενώ δεν υπάρχει θέμα αναγνώρισης προσώπου, όταν λέω πως έχω πονόδοντο. Και θάταν ανοησία να ρωτήσω 'είσαι σίγουρος πως σού πονάς'. Το ότι σ' αυτή την περίπτωση δεν υπάρχει δυνατότητα σφάλματος οφείλεται στο ότι η κίνηση που ίσως τείνουμε να θεωρούμε εσφαλμένη, μια 'κακή κίνηση', δεν είναι διόλου κίνηση αυτού του παιχνιδιού". Βλ., επίσης, Wittgenstein 1967, σελ. 121.

⁴²⁵ Στο γλωσσικό επίπεδο, το ίδιο φαινόμενο παρατηρείται με την προσωπική αντωνυμία "εγώ" όταν χρησιμοποιείται σε νοητικά κατηγορήματα. Όταν πω "Εγώ έχω πονόδοντο" δεν έχει νόημα να με ρωτήσει κάποιος "Είσαι σίγουρος ότι εσύ είσαι που έχεις πονόδοντο;" (βλ. Wittgenstein 1958, σελ. 67). Ο Shoemaker (1963/1994, σελ. 81) ονομάζει αυτό το χαρακτηριστικό του "εγώ" όταν χρησιμοποιείται σε τέτοιου είδους προτάσεις "ασυλία από λάθος που οφείλεται σε λανθασμένη ταυτοποίηση σε σχέση με την αντωνυμία πρώτου προσώπου" (immunity to error through misidentification relative to the first-person pronouns).

οπτική του πρώτου προσώπου, διότι οι τελευταίες αφορούν το περιεχόμενο των προτασιακών στάσεων.

Συνεπώς, δείξαμε ότι από την επιστημονική οπτική του τρίτου προσώπου δεν είναι δυνατόν να γνωρίσει κανείς άμεσα την πατρότητα των παρατηρούμενων προτασιακών στάσεων, δηλαδή να την γνωρίσει κατά τον τρόπο που τη γνωρίζει το υποκείμενο που τις έχει.

Όμως, ίσως, το συμπέρασμά μας να είναι ακόμα πρόωρο. Θα μπορούσε κανείς να αντιτείνει ότι από το γεγονός ότι δεν μπορούμε να γνωρίσουμε άμεσα την πατρότητα των προτασιακών στάσεων από την επιστημονική οπτική του τρίτου προσώπου δεν συνάγεται ότι οι παρατηρούμενες προτασιακές στάσεις δεν διαθέτουν όντως το συγκεκριμένο χαρακτηριστικό. Με άλλα λόγια, η ένσταση θα μπορούσε να είναι ότι περνάμε άκριτα από το γνωσιολογικό στο οντολογικό επίπεδο: ό,τι δεν μπορούμε να γνωρίσουμε δεν σημαίνει ότι δεν υπάρχει και δεν λειτουργεί από μόνο του.

Ωστόσο, τι θα μπορούσε πραγματικά να σημαίνει μια τέτοια ένσταση; Το "χαρακτηριστικό" που εξετάζουμε προϋποθέτει εγγενώς κάποιον γνωρίζον υποκείμενο, διότι πρόκειται για την άμεση γνώση της πατρότητας μιας προτασιακής στάσης. Με άλλα λόγια, το ίδιο το προς εξήγηση φαινόμενο αποτελεί μια γνωσιακή κατάσταση από την οποία προφανώς δεν μπορεί να εξαλειφθεί το γνωρίζον υποκείμενο. Σε αυτή τη λογική, το μόνο που θα μπορούσε κανείς να ισχυρισθεί είναι ότι η άμεση γνώση της πατρότητας των προτασιακών στάσεων είναι μια μορφή γνώσης που δεν μπορεί να δοθεί σε εξωτερικούς παρατηρητές, αλλά δίδεται μόνο στο ίδιο το υποκείμενο που έχει τις προτασιακές στάσεις. Δηλαδή, φτάνουμε πάλι στην αρχική θέση από την οποία ξεκινήσαμε.

Ίσως, όμως, υπάρχει ένας άλλος τρόπος με τον οποίον μπορεί να κατανοηθεί η παραπάνω ένσταση: το αντικείμενο που παρατηρεί ο εγκεφαλοσκόπος αποτελεί απλώς την αιτία της εμφάνισης του "χαρακτηριστικού" που αποκαλύπτεται από την οπτική του πρώτου προσώπου. Με άλλα λόγια, το αναζητούμενο "χαρακτηριστικό" δεν αποτελεί ιδιότητα του παρατηρούμενου εγκεφαλικού αντικειμένου, αλλά αιτιακά αποτέλεσμα του.

Μια τέτοια ένσταση, ωστόσο, κάθε άλλο παρά υπονομεύει την αρχική μας θέση: η άμεση γνώση της πατρότητας των προτασιακών στάσεων δεν αποτελεί μια γνώση που μπορεί να αποκτήσει κανείς από την επιστημονική οπτική του τρίτου προσώπου. Το αντικείμενο (ή η κατάσταση) στον εγκέφαλο, που υποτίθεται ότι προκαλεί την άμεση γνώση της πατρότητας των προτασιακών στάσεων στο φαινομενολογικό επίπεδο, δεν πρέπει να συγχέεται με τα ίδια τα νοητικά φαινόμενα ή κάποιο χαρακτηριστικό τους.

Τέλος, μια έσχατη ένσταση θα μπορούσε να είναι ότι ο εγκεφαλοσκόπος, σε αντίθεση με το ίδιο το υποκείμενο, δεν συνδέεται με τον κατάλληλο τρόπο με το συγκεκριμένο εγκεφαλικό αντικείμενο, γι' αυτό και δεν έχει μια άμεση γνώση της πατρότητας των προτασιακών στάσεων. Με άλλα λόγια, το εν λόγω χαρακτηριστικό των προτασιακών στάσεων

μπορεί να δοθεί μόνο στο ίδιο το υποκείμενο που διαθέτει το υποτιθέμενο εγκεφαλικό αντικείμενο.

Όμως, και πάλι, τι ακριβώς αμφισβητεί αυτή η ένσταση; Σίγουρα όχι το ότι από την επιστημονική οπτική του τρίτου προσώπου δεν μπορούμε να έχουμε πρόσβαση στο συγκεκριμένο χαρακτηριστικό. Αντίθετα, αυτό που αναδεικνύει είναι ότι το εν λόγω χαρακτηριστικό δεν μπορεί να δοθεί στην αντίληψη, δηλαδή στις αισθήσεις με τις οποίες προσεγγίζουμε τα αντικείμενα και τις ιδιότητες του εξωτερικού κόσμου.

Θεωρούμε, λοιπόν, ότι η πρώτη μεγάλη διαφορά που διαπιστώσαμε μεταξύ της οπτικής του πρώτου προσώπου και της επιστημονικής οπτικής του τρίτου προσώπου αποτελεί πράγματι μια μη εξαιρεϊσμηκή διαφορά που αναδεικνύει μια όψη της γνωσιολογικής ασυμμετρίας των δύο οπτικών. Προτού περάσουμε στην άλλη μεγάλη διαφορά, θα εξετάσουμε δύο αλληλένδετα ερωτήματα, η απάντηση των οποίων θα συνεισφέρει στο γενικότερο ερώτημα για το πώς πρέπει να κατανοούνται τα νοητικά φαινόμενα. Το πρώτο ερώτημα αφορά το πώς μπορεί να εξηγηθεί η άμεση γνώση της πατρότητας των προτασιακών στάσεων όπως αυτή δίδεται από την οπτική του πρώτου προσώπου, ενώ το δεύτερο ερώτημα έχει να κάνει με το ποια είναι τελικά η χρησιμότητα του συγκεκριμένου είδους γνώσης.

8.2.1.2 Πώς μπορεί να εξηγηθεί από την οπτική του πρώτου προσώπου η άμεση γνώση της πατρότητας;

Θα ξεκινήσουμε με το πρώτο ερώτημα. Πώς μπορεί να εξηγηθεί, λοιπόν, το ότι από την οπτική του πρώτου προσώπου έχουμε άμεση γνώση της πατρότητας των προτασιακών στάσεων; Η απάντηση σε αυτό το ερώτημα προϋποθέτει μια θεωρία για το τι είναι η οπτική του πρώτου προσώπου και πώς "λειτουργεί".

8.2.1.2.1 Η θεωρία της ενδοσκόπησης

Η κυριότερη θεωρία που έχει προταθεί είναι η θεωρία της εσωτερικής αντίληψης. Πιο συγκεκριμένα, η οπτική του πρώτου προσώπου θεωρείται ότι οφείλεται στην ύπαρξη μιας ιδιαίτερης μορφής αντίληψης η οποία, σε αντίθεση με την συνήθη αντίληψη, δεν αφορά τα αντικείμενα του εξωτερικού κόσμου, αλλά τις νοητικές καταστάσεις του υποκειμένου. Γι' αυτό το λόγο, η εσωτερική αντίληψη ονομάζεται και "ενδοσκόπηση"⁴²⁶. Η ενδοσκόπηση, λοιπόν, θεωρείται ένα είδος αντίληψης που διαφοροποιείται από την συνήθη αντίληψη με βάση το είδος των αντικειμένων στα οποία αυτή έχει πρόσβαση: ενώ η συνήθης αντίληψη (η "εξωτερική" αντίληψη) έχει πρόσβαση σε εξωτερικά αντικείμενα, η ενδοσκόπηση έχει πρόσβαση σε εσωτερικά αντικείμενα. Η

⁴²⁶ Λατινικά: introspection → spicere (= κοιτώ) intra (= εντός).

ενδοσκόπηση ως μέθοδος για την μελέτη των νοητικών φαινομένων, όπως είδαμε στο πρώτο κεφάλαιο, έφτασε στο απόγειό της με την ενδοσκοπική ψυχολογία.

Ακόμα κι αν δεχτούμε ως ορθή τη θεωρία της ενδοσκόπησης⁴²⁷, το ερώτημα που τίθεται είναι αν μπορούμε με βάση αυτήν τη θεωρία να εξηγήσουμε την άμεση γνώση της πατρότητας των προτασιακών στάσεων που χαρακτηρίζει την οπτική του πρώτου προσώπου.

Η απάντηση είναι προφανώς όχι, διότι στο βαθμό που η πρόσβαση στα νοητικά φαινόμενα από την οπτική του πρώτου προσώπου κατανοείται ως ένα είδος *αντίληψης*, ισχύουν τα ίδια επιχειρήματα που αναπτύχθηκαν κατά της δυνατότητας άμεσης γνώσης της πατρότητας των προτασιακών στάσεων από την επιστημονική οπτική του τρίτου προσώπου. Εφόσον, με άλλα λόγια, δείξαμε ότι η αντίληψη νοητικών αναπαραστάσεων ή αντικειμένων στο νευροφυσιολογικό / λειτουργικό επίπεδο δεν επαρκεί για να αποκτήσουμε μια άμεση γνώση της πατρότητας των προτασιακών στάσεων, το ίδιο ισχύει και στην περίπτωση της "εσωτερικής" αντίληψης νοητικών αναπαραστάσεων ή αντικειμένων που θα εμφανίζονταν στο φαινομενολογικό επίπεδο.

Να επισημάνουμε για άλλη μια φορά ότι, σε αυτό το επίπεδο ανάλυσης, δεν επερωτούμε προς το παρόν την ορθότητα της θέσης ότι από την οπτική του πρώτου προσώπου έχουμε τη δυνατότητα αντίληψης εσωτερικών αντικειμένων ή εσωτερικών αναπαραστάσεων. Αυτό που υποστηρίζουμε απλά είναι ότι *ακόμα και αν είχαμε* μια τέτοια δυνατότητα ενδοσκόπησης στο φαινομενολογικό επίπεδο, αυτή δεν θα επαρκούσε για να εξηγήσει την άμεση γνώση της πατρότητας των προτασιακών στάσεων για τους ίδιους λόγους που αποδείχτηκε ότι δεν ήταν επαρκής η "εξωτερική" αντίληψη νοητικών αναπαραστάσεων ή αντικειμένων στο νευροφυσιολογικό ή το λειτουργικό επίπεδο.

8.2.1.2.2 Η θεωρία του αναστοχασμού

Μια δεύτερη θεωρία για το τι είναι η οπτική του πρώτου προσώπου και πώς "λειτουργεί" είναι η θεωρία του αναστοχασμού (reflection). Η έννοια του αναστοχασμού αρχικά χρησιμοποιήθηκε ως σχεδόν συνώνυμη της έννοιας της ενδοσκόπησης. Όπως επισημαίνει ο Lyons στο βιβλίο του *The Disappearance of Introspection*, "παρά τη μεγάλη χρήση που έγινε [κατά τον 17^ο αιώνα]... στους όρους 'ενδοσκόπηση', 'εσωτερική αίσθηση' και 'αναστοχασμός', δεν κατέστη σαφές εάν αυτοί οι όροι έπρεπε να θεωρηθούν συνώνυμοι" (Lyons 1988, σελ. 3). Η έννοια του αναστοχασμού αρχικά χρησιμοποιήθηκε για να δηλώσει μια εμπειρική λειτουργία, ένα είδος αντίληψης που στρέφεται σε

⁴²⁷ Η ορθότητα της θεωρίας της ενδοσκόπησης θα τεθεί υπό αμφισβήτηση όταν εξετάσουμε τη δεύτερη μεγάλη διαφορά που παρουσιάζει η οπτική του πρώτου προσώπου από την επιστημονική οπτική του τρίτου προσώπου.

μια εσωτερική πραγματικότητα⁴²⁸. Ενώ, η αίσθηση στρέφεται στην εξωτερική πραγματικότητα, ο αναστοχασμός στρέφεται στην εσωτερική πραγματικότητα, δηλαδή στις νοητικές μας καταστάσεις⁴²⁹. Ο αναστοχασμός, κατανοημένος με βάση το μοντέλο της αντίληψης αντικειμένων, είναι ουσιαστικά ταυτόσημος με την ενδοσκόπηση.

Η έννοια του αναστοχασμού, ωστόσο, έχει κατανοηθεί και με ορισμένους άλλους τρόπους στην ιστορία της νεώτερης φιλοσοφίας. Καταρχάς, στον Leibniz ο αναστοχασμός, σύμφωνα με τον Gasché (1986) χάνει τον εμπειρικό χαρακτήρα που είχε στον Locke και κατανοείται ως λογικός αναστοχασμός⁴³⁰. Αργότερα, ο Kant, εισάγει την έννοια του υπερβατολογικού αναστοχασμού διακρίνοντάς την τόσο από την έννοια της εσωτερικής αίσθησης (ό,τι ονομάσαμε "ενδοσκόπηση", ή, στον Locke, "(εμπειρικό) αναστοχασμό"), όσο και από τον (λογικό) αναστοχασμό του Leibniz.⁴³¹

⁴²⁸ "Η άλλη πηγή από την οποία η εμπειρία προμηθεύει τη διάνοια με ιδέες είναι η αντίληψη των λειτουργιών του ίδιου του δικού μας νου εντός μας όπως όπως εφαρμόζεται στις ιδέες που έχει αποκτήσει ... Αυτή την πηγή ιδεών κάθε άνθρωπος την έχει εξ' ολοκλήρου εντός του και παρότι δεν είναι αίσθηση, καθώς δεν σχετίζεται καθόλου με εξωτερικά αντικείμενα, είναι πολύ όμοια με αυτήν και θα έπρεπε κανονικά να ονομάζεται *εσωτερική αίσθηση*. Αλλά όπως ονομάζω την άλλη Αίσθηση, έτσι αποκαλώ αυτήν ΑΝΑΣΤΟΧΑΣΜΟ, οι ιδέες που παρέχει όντας τέτοιες ώστε ο νους τις βρίσκει αναστοχαζόμενος τις ίδιες τις εντός αυτού λειτουργίες" (Locke 1984, Βιβλίο II, Κεφάλαιο 1, §4).

⁴²⁹ Όπως επισημαίνει και ο Husserl (1900/1970, σελ. 853): "ο Locke, χρησιμοποιώντας τους όρους 'αναστοχασμός' και 'αίσθηση', εισήγαγε τις δύο αντίστοιχες τάξεις αντίληψεων στη νεώτερη φιλοσοφία. Αυτός ο χωρισμός παραμένει σε ισχύ μέχρι σήμερα. Η εξωτερική αντίληψη θεωρήθηκε, ακολουθώντας τον Locke, αντίληψη σωμάτων, ενώ η εσωτερική αντίληψη ήταν η αντίληψη που τα 'πνεύματα' ή οι 'ψυχές' μας έχουν για τις ίδιες τις δικές τους δραστηριότητες (τις cogitationes τους με την καρτεσιανή έννοια)".

⁴³⁰ Χαρακτηριστικά ο Gasché (1986, σελ. 18) επισημαίνει: "Για τον Leibniz, επίσης, ο αναστοχασμός είναι μια προσοχή που κατευθύνεται στον εαυτό, αλλά είναι μια διαδικασία που λαμβάνει χώρα με την στροφή από οτιδήποτε το εμπειρικό προς μια διανοητική ενεργοποίηση των έμφυτων ιδεών. Αντίθετα προς τον εμπειρικό αναστοχασμό του Locke, ο αναστοχασμός του Leibniz πρέπει να ονομαστεί λογικός". Βλέπε, επίσης, Kant 1979, σελ. 278, A262/B318.

⁴³¹ Ο υπερβατολογικός αναστοχασμός είναι η έρευνα για τις συνθήκες δυνατότητας της γνώσης: "Ο αναστοχασμός (reflexio) δεν ασχολείται με τα αντικείμενα τα ίδια με σκοπό να ποριστεί έννοιες απευθείας από αυτά αλλά είναι η κατάσταση εκείνη του πνεύματος, όπου αρχίζουμε να προετοιμαζόμαστε για να ανακαλύψουμε τους υποκειμενικούς όρους που μόνο υπ' αυτούς είναι δυνατό να φτάσουμε σε έννοιες" (Kant 1979, σελ. 313, A260/B316). Όμως, όπως επισημαίνει ο Gasché (1986, σελ. 18): "...ο υπερβατολογικός αναστοχασμός είναι επίσης η μελέτη της ενότητας της λειτουργίας του ίδιου του αναστοχασμού". Πρόκειται για την καθαρή συνθετική ενότητα του "εγώ νώ" που, σύμφωνα με τον Kant, πρέπει να συνοδεύει όλες μου τις παραστάσεις. Το "εγώ νώ" είναι η συνείδηση του

Υπάρχει, τέλος, και μια τέταρτη χρήση του όρου "αναστοχασμός" η οποία απαντά στη σύγχρονη φιλοσοφία του νου. Σύμφωνα με αυτήν, ο αναστοχασμός είναι ένα νοητικό ενέργημα δεύτερης τάξης, δηλαδή ένα είδος σκέψης πάνω στη σκέψη (μια "νόησις νοήσεως"). Με άλλα λόγια, ο αναστοχασμός είναι ένα νοητικό ενέργημα που έχει ως περιεχόμενό του ένα νοητικό ενέργημα πρώτης τάξης⁴³². Αυτή η έννοια αναστοχασμού χρησιμοποιείται σε μια σύγχρονη θεωρία για την εξήγηση της αυτοσυνειδησίας. Η συγκεκριμένη θεωρία ονομάζεται "Σκέψη υψηλότερης Τάξης" (Higher Order Thought, συντμ.: HOT)⁴³³. Σύμφωνα λοιπόν με τη Σκέψη υψηλότερης Τάξης, αποκτούμε αυτοσυνειδησία όταν θεματοποιούμε τις ίδιες μας τις σκέψεις. Ο όρος "αυτοσυνειδησία" στη συγκεκριμένη θεωρία χρησιμοποιείται αρκετά γενικά ώστε να σημαίνει τόσο την προσωπική ταυτότητα συγχρονικά και διαχρονικά, όσο και την άμεση γνώση της πατρότητας των προτασιακών στάσεων.

Οι τρεις τελευταίες χρήσεις της έννοιας του αναστοχασμού, παρότι δεν στηρίζονται στο μοντέλο της αντίληψης, διατηρούν τη βασική μεταφορά της "στροφής" πάνω στα ίδια τα νοητικά ενεργήματα. Ωστόσο, δεν θα εξετάσουμε περαιτέρω τι ακριβώς μπορεί να σημαίνει αυτή η μεταφορά της "στροφής" στο έργο του Leibniz και του Kant διότι μια

εαυτού, ή όπως λέγει ο Kant, η κατάληψη (apperception): "Η συνείδηση του εαυτού (κατάληψη) είναι η απλή παράσταση του εγώ" (Kant 1976, σελ. 143, B68). Για την ακρίβεια πρόκειται για την υπερβατολογική κατάληψη, την οποία ο Kant ρητά διακρίνει από την εμπειρική κατάληψη ή εσωτερική αίσθηση (ό,τι, δηλαδή, ονομάσαμε εμπειρικό αναστοχασμό ή ενδοσκοπήση): "Η συνείδηση του εαυτού, όπως καθορίζεται από την κατάσταση μας, στην εσωτερική αίσθηση είναι απλώς εμπειρική, μεταβαλλόμενη κάθε φορά: μέσα στη ροή αυτή των εσωτερικών φαινομένων δεν μπορεί να υπάρχει ένα σταθερό και μόνιμο Εγώ· αυτό ονομάζεται συνήθως εσωτερική αίσθηση ή εμπειρική κατάληψη. ... Τώρα, γνώσεις δεν μπορούν να γεννηθούν μέσα μας [ούτε] καμιά σύνδεση και ενότητα των ιδίων μεταξύ τους, χωρίς εκείνη την ενότητα της συνειδήσεως, η οποία προηγείται απ' όλα τα δεδομένα των εποπτειών και που μόνο σε αναφορά προς αυτή είναι δυνατή κάθε παράσταση αντικειμένων. Αυτή την καθαρή πρωταρχική, μη υποκείμενη σε μεταβολή συνείδηση θέλω τώρα να την ονομάσω υπερβατολογική κατάληψη" (Kant 1979, σελ. 98-99, A107). Η έννοια του υπερβατολογικού αναστοχασμού χρησιμοποιήθηκε από τον Kant και, κυρίως, από τον γερμανικό ιδεαλισμό για την ανάλυση της έννοιας της αυτοσυνειδησίας. Συγκεκριμένα, θεωρήθηκε ότι η δομή του αναστοχασμού συγκροτεί την αυτοσυνειδησία. Σύμφωνα με αυτό το μοντέλο, θα μπορούσε κανείς να ισχυρισθεί ότι η άμεση γνώση της πατρότητας των νοητικών ενεργημάτων βασίζεται στον αναστοχασμό του εγώ από το ίδιο το εγώ. Έτσι, το υποκείμενο γνωρίζει τον εαυτό του στρεφόμενο αναστοχαστικά στον ίδιο του τον εαυτό. Αυτή η λειτουργία μπορεί να ονομαστεί για συντομία "αυτό-αναστοχασμός".

⁴³² Τα νοητικά φαινόμενα εδώ παύουν να κατανοούνται ως εσωτερικά αντικείμενα και εκλαμβάνονται ως ενεργήματα (πρβλ. υποσημείωση 101)

⁴³³ Βλέπε, Rosenthal 1986.

τέτοια έρευνα υπερβαίνει κατά πολύ το πλαίσιο του κεφαλαίου. Θα περιοριστούμε στην εξέταση της τέταρτης χρήσης της έννοιας του αναστοχασμού η οποία αποτελεί ουσιαστικά τον τρόπο που χρησιμοποιείται η έννοια στη σύγχρονη φιλοσοφία του νου. Μεθοδολογικά, θα ακολουθήσουμε τον ίδιο δρόμο που ακολουθήσαμε και κατά την εξέταση του μοντέλου της ενδοσκόπησης: χωρίς να αμφισβητήσουμε συνολικά το μοντέλο του αναστοχασμού, θα εξετάσουμε αν με βάση αυτό μπορεί να εξηγηθεί η άμεση γνώση της πατρότητας των προτασιακών στάσεων.

Ο αναστοχασμός αποτελεί ένα νοητικό ενέργημα δεύτερης τάξης, δηλαδή ένα ενέργημα Α που έχει ως περιεχόμενο του ένα ενέργημα Β. Το ενέργημα Β μπορεί με τη σειρά του να είναι ένα αναστοχαστικό ή μη-αναστοχαστικό ενέργημα. Θα εξετάσουμε την περίπτωση που ισχύει το δεύτερο. Το ενέργημα Β πρέπει να βοηθεί ως τρέχον ενέργημα, διαφορετικά ο αναστοχασμός δεν θα διακρινόταν από την *ανάμνηση* προτασιακών στάσεων. Ο αναστοχασμός, λοιπόν, προϋποθέτει τη δυνατότητα ταυτόχρονης επιτέλεσης δύο ενεργημάτων όπου το ένα θεματοποιεί το άλλο.

Είναι δυνατή η εξήγηση της άμεσης γνώσης της πατρότητας με τη βοήθεια αυτού του μοντέλου για το πώς "λειτουργεί" η οπτική του πρώτου προσώπου; Θεωρούμε πως όχι. Ας ξεκινήσουμε με ένα παράδειγμα. Ένα αναστοχαστικό ενέργημα θα μπορούσε να εκφραστεί γλωσσικά ως εξής: "(εγώ) πιστεύω ότι (εγώ) βλέπω ότι ρ". Αν δεχτούμε ότι η λέξη "εγώ" εκφράζει αυτό που ονομάσαμε "άμεση γνώση της πατρότητας ενός ενεργήματος", τότε, ακόμα και αν θεωρήσουμε ότι το αναστοχαστικό ενέργημα εξηγεί την γνώση της πατρότητας του ενεργήματος 1^{ης} τάξης, σίγουρα δεν μπορεί να εξηγήσει τη γνώση της πατρότητας του ενεργήματος 2^{ης} τάξης. Η επιτέλεση ενός πρόσθετου νοητικού ενεργήματος για τη θεματοποίηση του νοητικού ενεργήματος 2^{ης} τάξης, προφανώς, δεν επιλύει το πρόβλημα, αλλά οδηγεί σε μια επ' άπειρον αναδρομή. Γι' αυτό, όπως επισημαίνει και ο Dieter Heinrich (1971, σελ. 11): "Εκείνο το οποίο πρόκειται να γίνει ρητά συνειδητό μέσω του αναστοχασμού πρέπει να είναι παρόν, τουλάχιστον άρρητα, έτσι ώστε να προκαλέσει το ενέργημα του αναστοχασμού που κατευθύνεται προς αυτό. Στον αναστοχασμό προϋποτίθεται μια συνείδηση του υποκειμένου... η θεωρία του αναστοχασμού μπορεί το πολύ να εξηγήσει τη ρητή εμπειρία του εαυτού, αλλά όχι την ίδια την αυτοσυνειδησία". Με άλλα λόγια, η θεωρία του αναστοχασμού, η οποία υποτίθεται ότι εξηγεί την προέλευση της αυτοσυνειδησίας (ή, στην προκειμένη περίπτωση, της άμεσης γνώσης της πατρότητας), προϋποθέτει την ίδια την αυτοσυνειδησία που επιχειρεί να εξηγήσει.⁴³⁴

⁴³⁴ Για μια συστηματική ανασυγκρότηση της κριτικής του Heinrich βλέπε Βηροπαϊδης 1995α.

Ο Henrich προτείνει, επίσης, και ένα άλλο επιχείρημα που καταδεικνύει την κυκλικότητα της θεωρίας του αναστοχασμού ως μοντέλου για την εξήγηση της αυτοσυνειδησίας. Το εν λόγω επιχείρημα βασίζεται στην γενικότερη ιδέα ότι οποιοδήποτε ενέργημα σύλληψης κάποιου πράγματος (είτε αυτό το ενέργημα κατανοείται ως μια μορφή αντίληψης είτε, με οποιοδήποτε άλλο τρόπο) προϋποθέτει μια προγενέστερη δυνατότητα αναγνώρισης αυτού που συλλαμβάνεται. Κατά συνέπεια, το ενέργημα του αναστοχασμού δεν είναι δυνατόν να συγκροτεί την αυτοσυνειδησία (ή, εν προκειμένω, την άμεση γνώση της πατρότητας) γιατί ήδη πάντα την προϋποθέτει⁴³⁵. Ένα παρόμοιο επιχείρημα έχει αναπτυχθεί ανεξάρτητα και από τον Shoemaker (1968/1994, σελ. 88) για την περίπτωση της ενδοσκοπησης: "για να αναγνωρίσω έναν εαυτό ως εμένα τον ίδιο μέσω της κατοχής αυτής της ιδιότητας, θα έπρεπε να γνωρίζω ότι εγώ τον παρατηρώ μέσω της εσωτερικής αίσθησης, και αυτή η αυτογνωσία, όντας το θεμέλιο της αναγνώρισης του εαυτού μου ως εμένα τον ίδιο, δεν θα μπορούσε η ίδια να θεμελιωθεί σε αυτή την αναγνώριση".

Από τα παραπάνω επιχειρήματα γίνεται φανερό ότι η αυτοσυνειδησία, αλλά και, ειδικότερα, η άμεση γνώση τη πατρότητας ενός ενεργήματος, δεν μπορούν να "αναχθούν" -για να χρησιμοποιήσουμε μια σύγχρονη ορολογία- στον αναστοχασμό. Με άλλα λόγια, μπορούμε να συμπεράνουμε ότι το μοντέλο του αναστοχασμού δεν επαρκεί για να εξηγήσει το πώς είναι δυνατόν από την οπτική του πρώτου προσώπου να έχουμε μια άμεση γνώση της πατρότητας των επιτελούμενων νοητικών ενεργημάτων, διότι η συγκεκριμένη γνώση προϋποτίθεται του αναστοχασμού. Αν υπάρχει κάτι που ενδεχομένως μπορεί να εξηγήσει ο αναστοχασμός, αυτό είναι η *θεματική* αυτοσυνειδησία, δηλαδή η θεωρητική συνείδηση του εαυτού. Ωστόσο, η άμεση γνώση της πατρότητας ενός ενεργήματος αποτελεί πρωταρχικά μια *μη θεματική, προαναστοχαστική* μορφή αυτοσυνειδησίας.

8.2.1.2.3 Η άμεση γνώση της πατρότητας δεν οφείλεται σε ένα πρόσθετο ενέργημα

⁴³⁵ "Σε κάθε περίπτωση, ανεξάρτητα από το πώς το 'εγώ' φτάνει να σχετιστεί με τον εαυτό του, είτε μέσω ενός ενεργήματος αναστοχασμού είτε με κάποιο άλλο τρόπο, το 'εγώ' πρέπει να συλλάβει τον εαυτό του στην αυτοσυνειδησία. Εφόσον η σύλληψη του εαυτού πρέπει συγκεκριμένα να είναι μια συνειδητή σύλληψη, το 'εγώ' πρέπει να έχει κάποια ιδέα ότι εκείνο με το οποίο εξοικειώνεται στην αυτοσυνειδησία είναι αυτό το ίδιο. Για αυτό το σκοπό, δεν είναι αναγκαίο να έχει κάποιο είδος εννοιολογικής γνώσης του εαυτού ή να είναι ικανό να δώσει μια περιγραφή του εαυτού. Αλλά σε κάθε περίπτωση, θα πρέπει να είναι ικανό να βεβαιώσει ότι είναι αυτό το ίδιο εκείνο με το οποίο εξοικειώνεται στην αυτοσυνειδησία, ανεξάρτητα από το εάν αυτή η αυτό-εξοικείωση απορρέει από τον αναστοχασμό ή με κάποιον άλλον τρόπο" (Henrich 1971, σελ. 11).

Κατά συνέπεια, αυτό που ονομάσαμε "άμεση γνώση της πατρότητας ενός ενεργήματος" δεν μπορεί να εξηγηθεί ούτε με το μοντέλο της ενδοσκοπησης ούτε με εκείνο του αναστοχασμού. Αν επιμείνει κανείς στην ύπαρξη της εν λόγω άμεσης γνώσης βασίζόμενος στη φαινομενολογία των νοητικών ενεργημάτων, γίνεται φανερό ότι τα συγκεκριμένα μοντέλα οδηγούν σε ανυπέβλητα λογικά αδιέξοδα. Η άμεση γνώση της πατρότητας ενός ενεργήματος κατά τη στιγμή της επιτέλεσής του δεν πρέπει να θεωρηθεί ότι οφείλεται σε ένα πρόσθετο ενέργημα: *Η ίδια η επιτέλεση ενός ενεργήματος επιφέρει και την άμεση γνώση της πατρότητάς του.* Το ενέργημα και αυτός που το επιτελεί δεν είναι δύο διακριτά φαινόμενα που συνδέονται εξωτερικά, αλλά αποτελούν ένα ενιαίο φαινόμενο – μια οργανική ενότητα. Αυτός είναι και ο λόγος που τα μοντέλα της (εσωτερικής) αντίληψης και του αναστοχασμού δεν μπορούν να χρησιμοποιηθούν για την εξήγηση της αδιάσπαστης ενότητας που υπάρχει ανάμεσα στο ενέργημα και αυτόν που το επιτελεί. Τα εν λόγω μοντέλα προσποθούν την ύπαρξη δύο ξεχωριστών πόλων που μόνο εκ των υστέρων έρχονται σε επαφή μέσω κάποιας αντιληπτικής ή αναστοχαστικής διαδικασίας η οποία πηγάζει από τον έναν πόλο και κατευθύνεται στον άλλο. Γι' αυτό από εδώ και στο εξής κάθε φορά που χρησιμοποιούμε τους όρους "αντικειμενικός πόλος" και "υποκειμενικός πόλος" θα θέτουμε τη λέξη "πόλο" μέσα σε εισαγωγικά προκειμένου να δηλώσουμε ότι η συγκεκριμένη διάκριση δεν πρέπει να κατανοηθεί οντολογικά.

Παρότι ο "χώρος" που εξετάζουμε είναι, εκ συστάσεως, ιδιαίτερα ασαφής και, κατά συνέπεια, υπάρχει πάντα περιθώριο "ελλιγμών", θα θεωρήσουμε ότι τα μοντέλα της ενδοσκοπησης και του αναστοχασμού δεν επαρκούν για την εξήγηση της άμεσης γνώσης της πατρότητας των νοητικών ενεργημάτων. Σε ένα τόσο ασαφή "χώρο" είναι σκόπιμο να ακολουθούμε πρωτίστως την *φαινομενολογική προφάνεια* και όχι τις οποιεσδήποτε θεωρητικές κατασκευές. Θα κρατήσουμε, λοιπόν, από τα προηγούμενα ότι υπάρχει μια ιδιαίτερη φαινομενολογική ποιότητα που συνοδεύει κάθε ενέργημά μας, και ότι αυτή δεν μπορεί να εξηγηθεί με το μοντέλο της ενδοσκοπησης ή του αναστοχασμού. Επίσης, το γεγονός ότι η εν λόγω φαινομενολογική ποιότητα δεν μπορεί να προκύψει μέσα από την ταυτόχρονη αντιληπτική ή ενδοσκοπική πρόσβαση στο ενέργημα και τον "φορέα" του δείχνει ότι δεν πρόκειται για κάποια προτασιακή γνώση, δηλαδή, γνώση, εκ μέρους του υποκειμένου, ότι κάποιο ενέργημα είναι δικό του. Στο ίδιο συμπέρασμα, εξάλλου, οδηγηθήκαμε και από την ιστορία του αμνησιακού όπου φάνηκε καθαρά ότι η άμεση γνώση της πατρότητας ενός ενεργήματος υπάρχει ακόμα και αν ο ίδιος δεν έχει καμιά προτασιακή ή, γενικότερα, αναπαραστασιακή, γνώση περί του εαυτού του. Η φαινομενολογική ποιότητα που εντοπίζουμε έχει να κάνει με το *πώς είναι να είναι* δικό μου το επιτελούμενο ενέργημα. Κατά συνέπεια, είναι

ίσως παραπλανητικό το να ονομάζουμε αυτή τη φαινομενολογική ποιότητα "άμεση γνώση της πατρότητας ενός ενεργήματος", διότι, όπως προέκυψε, δεν πρόκειται για γνώση⁴³⁶, τουλάχιστον αναπαραστασιακή.

Στη σύγχρονη φιλοσοφία του νου τέτοιου τύπου φαινομενολογικές ποιότητες ονομάζονται συλλήβδην "qualia" - έναν όρο που θα μπορούσαμε να αποδώσουμε ως "υποκειμενικές" ή "φαινομενικές ποιότητες". Δύο τέτοιες "διάσημες" υποκειμενικές ποιότητες είναι το πώς είναι να είναι κανείς νυχτερίδα⁴³⁷ και το πώς είναι να βλέπει κανείς κόκκινο⁴³⁸. Τα qualia θεωρείται ότι χαρακτηρίζουν αυτό που στην σύγχρονη φιλοσοφία του νου αποκαλείται "φαινομενική συνείδηση" (phenomenal consciousness). Ωστόσο, θα πρέπει να επισημάνουμε ότι τα qualia- οι υποκειμενικές ποιότητες- έχουν κυρίως συνδεθεί με την αντίληψη αισθημάτων⁴³⁹ και όχι με την επιτέλεση ενεργημάτων. Με άλλα λόγια, έχει θεωρηθεί ότι συνοδεύουν μη αποβλεπτικές καταστάσεις. Η γραμμή σκέψης εδώ είναι ότι, εφόσον τα qualia αφορούν μη αποβλεπτικές καταστάσεις, δεν σχετίζονται με γνωσιακά φαινόμενα, δηλαδή τη σκέψη και τη γνώση, αλλά μόνο την εμπειρία⁴⁴⁰. Δεδομένου, λοιπόν, ότι τα qualia δεν αποτελούν χαρακτηριστικά των γνωσιακών φαινομένων - για τα οποία, κατά κύριο λόγο, ενδιαφέρεται η γνωσιακή επιστήμη- δεν πρέπει να θεωρείται ότι η ύπαρξη των qualia δημιουργεί κάποιο πρόσθετο πρόβλημα. Αντίθετα, η φαινομενική ποιότητα που αναδεικνύουμε στο παρόν κεφάλαιο είναι συναφόμενη και με αποβλεπτικές καταστάσεις, δηλαδή τις εκάστοτε προτασιακές στάσεις. Υπ' αυτή την έννοια, το πρόβλημα που τίθεται από την ύπαρξη της εν λόγω φαινομενικής ποιότητας δεν μπορεί να παρακαμφθεί με τον παραπάνω τρόπο.

Μια άλλη έννοια με την οποία συσχετίζονται τα qualia είναι εκείνη της υποκειμενικότητας: τα νοητικά ενεργήματα που επιτελώ είναι κατά έναν ουσιαστικό τρόπο *δικά μου*. Ωστόσο, το πρόβλημα της υποκειμενικότητας δεν θα πρέπει να ταυτισθεί, κατ' ανάγκη, με το

⁴³⁶ Σε αυτό το σημείο θα μπορούσε κανείς να διαμαρτυρηθεί ότι καταλήξαμε απλώς στην πασίγνωστη βιτγκεσταϊνική θέση ότι για να μιλήσει κανείς για γνώση χρειάζεται να υπάρχει περιθώριο αμφιβολίας (πρβλ. υποσημείωση 424) και κατά συνέπεια αν την λαμβάναμε εξ αρχής υπόψη δεν θα χρειαζόταν να κάνουμε όλη αυτή τη διαδρομή. Ωστόσο, η ίδια η διατύπωση της διαμαρτυρίας δίνει και την απάντηση για τη σκοπιμότητα του εγχειρήματος. Αν παραμέναμε στο βιτγκεσταϊνικό 'δόγμα' θα ήμασταν απλώς οπαδοί του Wittgenstein και δεν θα ακολουθούσαμε μια ανεξάρτητη διαδρομή που διαφαίνεται πιστεύουμε από διαφορετική σκοπιά τα αδιέξοδα μιας σύγχρονης μυθολογία που διαπνέει τη γνωσιακή επιστήμη.

⁴³⁷ Βλ. Nagel 1974.

⁴³⁸ Βλ. Jackson 1982.

⁴³⁹ Π.χ. το αίσθημα του πόνου ή το αίσθημα του κόκκινου.

⁴⁴⁰ Βλέπε, για παράδειγμα, τα επιχειρήματα στο δεύτερο κεφάλαιο του Jackson 1977.

πρόβλημα των φαινομενικών ή υποκειμενικών ποιότητων. Θα μπορούσε κανείς να φαντασθεί όντα που έχουν πρόσβαση σε υποκειμενικές ποιότητες χωρίς οι τελευταίες να τους δίδονται ως δικές τους⁴⁴¹. Η φαινομενική ποιότητα που εξετάσαμε στην παρούσα ενότητα εντάσσεται αποκλειστικά σε ό,τι στη φιλοσοφία του νου ονομάζεται "υποκειμενικότητα". Για τον λόγο αυτό δεν θα χρησιμοποιούμε τον όρο *qualia*, ο οποίος, όντας γενικότερος, αναφέρεται σε πολλά διαφορετικά φαινόμενα. Ούτε όμως και ο όρος που χρησιμοποιήσαμε από την αρχή, ήτοι "άμεση γνώση της πατρότητας μιας προτασιακής στάσης", είναι κατάλληλος, διότι η εν λόγω υποκειμενική ποιότητα δεν αποτελεί κάποιου είδους γνώση, τουλάχιστον αναπαραστασιακή. Πρόκειται μάλλον για μια μη-θεματική, προαναστοχαστική επίγνωση (*awareness*) ή, συντομότερα, μια συναίσθηση⁴⁴² της πατρότητας των προτασιακών στάσεων.

8.2.1.3 Ποια είναι η χρησιμότητα της συναίσθησης της πατρότητας;

Μπορούμε τώρα να περάσουμε στο δεύτερο ερώτημα που προαναγγειλάμε λίγες σελίδες παραπάνω, δηλαδή ποια είναι, τελικά, η χρησιμότητα αυτού του πρώτου "χαρακτηριστικού" που διαφεύγει από την επιστημονική οπτική του τρίτου προσώπου. Παιζει όντως κάποιο γνωστικό ρόλο η μη θεματική επίγνωση της πατρότητας των νοητικών ενεργημάτων;

Αν ακολουθήσουμε την κυρίαρχη τάση στην σύγχρονη φιλοσοφία του νου όσον αφορά τα *qualia* και την υποκειμενικότητα, θα πρέπει να δεχθούμε ότι πρόκειται για κάποιο *επιφανόμενο* που ως τέτοιο δεν έχει καμία αιτιακή επίδραση στα πραγματικά φαινόμενα. Συγκεκριμένα, θα μπορούσε κανείς να ισχυρισθεί ότι ένα νοητικό ενέργημα καθορίζεται

⁴⁴¹ Ενδεχομένως κάτι τέτοιο μπορούμε να υποθέσουμε ότι συμβαίνει στα ζώα και τα νήπια. Επίσης, σε αυτή την κατάσταση πλησιάζουν σχιζοφρενείς που βρίσκονται στη φάση της αποπροσωποποίησης (Βλ. Καγγελάρης 1985).

⁴⁴² Τόσο η συναίσθηση όσο και το συναίσθημα (*affect, feeling*)-νοημένο ως αυτό που "δίδεται" στη συναίσθηση- δεν ανήκουν, μιλώντας με καντιανούς όρους, στο χώρο της διάνοιας, διότι δεν έχουν αποβλεπτικότητα, δεν κατευθύνονται προς κάποιο αντικείμενο ή κατάσταση πραγμάτων. Ούτε, όμως, θα πρέπει να κατανοηθούν ως αίσθηση και αίσθημα, αντίστοιχα, διότι δεν αφορούν τα αντικείμενα (δηλαδή κάτι έτερο) αλλά το ίδιο το υποκείμενο. Χαρακτηριστική, προς αυτή την κατεύθυνση, είναι η επισήμανση του Kant στην *Κριτική της Κριτικής Δύναμης* για τα συναισθήματα της ηδονής και λύπης: "Εάν ένας προσδιορισμός του συναισθήματος της ηδονής ή λύπης αποκαλείται αίσθημα, τότε η έκφραση αυτή σημαίνει κάτι τελείως διαφορετικό από αυτό που σημαίνει όταν τη χρησιμοποιώ για να ονομάσω την παράσταση ενός πράγματος (μέσω των αισθήσεων, δηλαδή μια δεκτικότητα που ανήκει στη γνωστική ικανότητα). Γιατί στην τελευταία περίπτωση η παράσταση αναφέρεται στο αντικείμενο, ενώ στην πρώτη [αυτή αναφέρεται] απλώς στο υποκείμενο και δεν χρησιμεύει σε καμιά γνώση, ούτε καν σε εκείνη μέσω της οποίας το υποκείμενο γνωρίζει τον ίδιο του τον εαυτό" (§ 3, 206).

πλήρως αν είναι γνωστό το περιεχόμενο και το είδος του καθώς και το ποιος το επιτελεί, δηλαδή στοιχεία που μπορούν να δοθούν από την επιστημονική οπτική του τρίτου προσώπου. Η μη θεματική επίγνωση της πατρότητας του ενεργήματος, η οποία, όπως δείξαμε, δίνεται μόνο από την οπτική του πρώτου προσώπου, δεν συμβάλλει περαιτέρω στη συγκρότησή του ίδιου του ενεργήματος. Με άλλα λόγια, πρόκειται απλά για κάποια ποιότητα που απλώς συνοδεύει τα συνειδητά νοητικά φαινόμενα, χωρίς όμως να συμβάλλει το παραμικρό στο γνωσιακό τους ρόλο.

Αν ισχύει κάτι τέτοιο, τότε, σίγουρα, κάναμε πολύ φασαρία για το τίποτε. Ωστόσο, δεν θεωρούμε ότι αυτό είναι αληθές. Αντίθετα πιστεύουμε ότι η προαναστοχαστική επίγνωση για την οποία μιλήσαμε συμβάλλει καθοριστικά στη συγκρότηση της εκάστοτε προτασιακής στάσης και αυτό θα επιχειρήσουμε να δείξουμε στη συνέχεια.

Θα ξεκινήσουμε με ένα ερώτημα. Έστω ένα νοητικό φαινόμενο A η παρουσία του οποίου σε ένα υποκείμενο B προκαλεί σε δεδομένες συνθήκες μια συγκεκριμένη συμπεριφορά και συγκεκριμένες προδιαθέσεις για συμπεριφορά. Το ερώτημα είναι: αν, υπό τις ίδιες συνθήκες, εξέλει από το νοητικό φαινόμενο A η συναίσθηση της πατρότητας, θα επέφερε αυτή η απώλεια κάποια αλλαγή στην συμπεριφορά και στις προδιαθέσεις για συμπεριφορά του υποκειμένου B; Με άλλα λόγια, στην προτεινόμενη υποθετική περίπτωση παραμένουν σταθερά τόσο το περιεχόμενο και το είδος του νοητικού φαινομένου όσο και οι εξωτερικές συνθήκες, δηλαδή όλα τα στοιχεία που οι εξτερναλιστικές και οι ιντερναλιστικές θεωρίες λαμβάνουν υπόψη. Το μόνο που μεταβάλλεται είναι η συναίσθηση της πατρότητας. Θα προκαλέσει αυτή η διαφορά κάποια αλλαγή στη συμπεριφορά και στις προδιαθέσεις για συμπεριφορά του υποκειμένου;

Θεωρούμε ότι η διαφορά που θα επέλθει θα είναι δραματική. Έστω ότι το νοητικό φαινόμενο είναι η πρόθεσή μου να πάω στη κουζίνα να πω νερό. Αν από αυτή την πρόθεση χαθεί η συναίσθηση της πατρότητας τότε απομένω με μια σκέψη που δεν είναι δικιά μου και, παρόλα αυτά βρίσκεται "μέσα" στο νου μου. Πρόκειται για ένα εντελώς ανοίκειο συναίσθημα⁴⁴³. Είναι σαν κάποιος άλλος να βρίσκεται "μέσα" στο νου μου και να με διατάζει να πάω στην κουζίνα να πω νερό. Η παρουσία

⁴⁴³ Αρκετοί ερευνητές θεωρούν ότι ορισμένα βασικά συμπτώματα της σχιζοφρένειας οφείλονται στην απώλεια του συναίσθηματος οικειότητας με τα πράγματα και τον εαυτό. Χαρακτηριστικό παράδειγμα αποτελεί η πλάνη (delusion) Capgras. Οι ασθενείς που πάσχουν από την πλάνη Capgras πιστεύουν ότι "άλλα πρόσωπα (συνήθως, της οικογένειας) και ορισμένες φορές οικεία μέρη και αντικείμενα έχουν αντικατασταθεί από αντίγραφα ή ψώστες" (Gerrans 1999, σελ. 600). Σύμφωνα με την εξηγητική υπόθεση, οι ασθενείς φτάνουν σε αυτή την πλάνη όχι επειδή έχει μεταβληθεί κάτι στο περιεχόμενο των αντιληπτικών τους ενεργημάτων (εξάλλου, όλα τους εμφανίζονται όπως και πριν), αλλά επειδή έχει μεταβληθεί το συναίσθημα της οικειότητας που συνοδεύει την κανονική αντίληψη (βλ. Gerrans 1999).

ενός τέτοιου νοητικού φαινομένου είναι πιθανότερο να μου προκαλέσει ένα συναίσθημα βαθύτατου τρόμου και μια τάση φυγής παρά να με κινητοποιήσει να πάω στην κουζίνα να πω νερό. Ακόμα, όμως, και αν συμβεί το τελευταίο, η πράξη μου αυτή δεν θα έχει υπαγορευθεί από μια δικιά μου επιθυμία αλλά από μια διαταγή ενός "τρίτου". Δεν θα πρόκειται, με άλλα λόγια, για μια εκούσια πράξη αλλά για μια εξαναγκασμένη πράξη.

Όταν, λοιπόν, από ένα νοητικό φαινόμενο χαθεί η συναίσθηση της πατρότητας, τότε αυτό αποκόπτεται από το ψυχικό βίο του υποκειμένου - που, σε μια πρώτη προσέγγιση, μπορούμε να τον νοήσουμε ως ένα πλέγμα πεποιθήσεων και επιθυμιών. Το νοητικό φαινόμενο δεν συμπλέκεται με τα περιεχόμενα των πεποιθήσεων και των επιθυμιών μου, ούτε μπορεί να ενταχθεί σε αυτά. Έτσι, εμφανίζεται ως μια μεμονωμένη σκέψη που για να εκλογικευτεί πρέπει να αποδοθεί σε έναν τρίτο. Τέτοιες είναι οι φωνές που ακούν οι σχιζοφρενείς. Οι διαφορές στη συμπεριφορά και τις προδιαθέσεις για συμπεριφορά που προκαλούνται από ένα νοητικό φαινόμενο στο οποίο υπάρχει η συναίσθηση της πατρότητας και από ένα άλλο στο οποίο δεν υπάρχει είναι τόσο ριζικές όσο ριζική είναι η διαφορά μεταξύ της συμπεριφοράς (και των προδιαθέσεων) ενός φυσιολογικού ανθρώπου και ενός σχιζοφρενή.

Κατά συνέπεια, το πρώτο βασικό "χαρακτηριστικό" που διαπιστώσαμε ότι διαφεύγει από την επιστημονική οπτική του τρίτου προσώπου δεν αποτελεί κάποιο απλό επιφανόμενο αλλά έχει συγκροτητικό ρόλο στη γνωσιακή συμπεριφορά του υποκειμένου.

8.2.1.4 Η συναίσθηση της πατρότητας δεν είναι πρωταρχικά μια συναίσθηση της εσωτερικότητας αλλά μια συναίσθηση της ικανότητας του πράττειν

Προτού προχωρήσουμε στην εξέταση της δεύτερης μεγάλης διαφοράς μεταξύ της οπτικής του πρώτου προσώπου και της επιστημονικής οπτικής του τρίτου προσώπου θα αναφερθούμε σε μια διάκριση που προτείνεται από τους Graham & Stephens (1994a) η οποία μπορεί να αποτελέσει την αφετηρία για μια θετική προσέγγιση αυτού που ονομάσαμε "συναίσθηση της πατρότητας". Οι σχιζοφρενείς δεν αντιλαμβάνονται μόνο φωνές (ή, ακριβέστερα, ομιλιακά ενεργήματα). Ορισμένες φορές αντιλαμβάνονται και σκέψεις, σκέψεις όμως που δεν τις κατανοούν ως δικές τους. Το ενδιαφέρον με την δεύτερη περίπτωση είναι ότι αυτό που αντιλαμβάνονται οι σχιζοφρενείς δεν κατανοείται ως κάτι που βρίσκεται εκτός τους και, κατά συνέπεια, δίδεται από τις αισθήσεις, αλλά ως κάτι που βρίσκεται "εντός" τους και μάλιστα "εντός" του νου. Γι' αυτό άλλωστε, στην προκειμένη περίπτωση, αυτό που τους εμφανίζεται κατανοείται ως σκέψη και όχι ως φωνή.

Το σημαντικό στοιχείο που αναδεικνύει η παραπάνω διάκριση είναι ότι η συναίσθηση της πατρότητας δεν μπορεί να εξηγηθεί με βάση τη διάκριση εντός του νου / εκτός του νου: ακόμα και καταστάσεις που εμφανίζονται ως "εντός" του νου, δηλαδή ως σκέψεις, μπορούν να μην συνοδεύονται από τη συναίσθηση της πατρότητας. Με άλλα λόγια, η

συναίσθηση της πατρότητας δεν μπορεί να ταυτιστεί με τη συναίσθηση ότι μια κατάσταση είναι "εντός" του νου. Πώς, λοιπόν, θα μπορούσε να αναλυθεί η συναίσθηση της πατρότητας; Μια λύση που προτείνεται από τους Graham & Stephens (1994a) είναι ότι πρέπει να διακριθούν δύο είδη αυτοσυνειδησίας: η συνείδηση της υποκειμενικότητας ή των ορίων του εγώ και η συνείδηση της ικανότητας του πράττειν (agency). Αν δεχτούμε αυτή τη διάκριση, η συναίσθηση της πατρότητας θα πρέπει να αναλυθεί με βάση την έννοια της συναίσθησης της ικανότητας του πράττειν: ένα νοητικό φαινόμενο συνοδεύεται από τη συναίσθηση ότι είναι δικό μου αν κατανοείται ως μια πράξη και όχι ως ένα αντανakλαστικό. Ο σχιζοφρενής μπορεί να "κατέχει" ένα νοητικό φαινόμενο χωρίς όμως να αισθάνεται ότι το έχει παράγει. Βλέπουμε, λοιπόν, τελικά ότι η συναίσθηση της πατρότητας πρέπει να διακριθεί από τη συναίσθηση της κυριότητας (ownership): Η συναίσθηση της κυριότητας δεν συνεπάγεται κατ' ανάγκη τη συναίσθηση της πατρότητας⁴⁴. Θεωρούμε ότι η υπόδειξη των Graham & Stephens αποτελεί, ουσιαστικά, ένα πρώτο βήμα για την απόρριψη της καρτεσιανής προέλευσης διάκρισης μεταξύ ενός "χώρου" εντός και ενός "χώρου" εκτός του υποκειμένου ως κριτηρίου για τον προσδιορισμό της συναίσθησης της πατρότητας: η συναίσθηση της πατρότητας δεν είναι πρωταρχικά μια συναίσθηση της εσωτερικότητας, αλλά μια συναίσθηση της ικανότητας του πράττειν. Στην επόμενη ενότητα θα επανέλθουμε στην έννοια της συναίσθησης της ικανότητας του πράττειν και θα δείξουμε ότι διακρίσεις μεταξύ διαφορετικών ειδών προτασιακών στάσεων μπορούν, ως ένα βαθμό, να αναχθούν σε διαφοροποιήσεις της συναίσθησης της ικανότητας του πράττειν.

8.2.2 Άμεση Αντίληψη

Μέχρι αυτό το σημείο εξετάσαμε της συνέπειες της πρώτης μεγάλης διαφοράς που εντοπίσαμε μεταξύ της οπτικής του πρώτου προσώπου και της επιστημονικής οπτικής του τρίτου προσώπου. Πρόκειται για την εξάλειψη του υποκειμενικού "πόλου" των νοητικών φαινομένων όταν περνάμε από την οπτική του πρώτου προσώπου στην επιστημονική οπτική του τρίτου προσώπου. Στη συνέχεια, θα εξετάσουμε μια δεύτερη μεγάλη διαφορά που παρουσιάζει η οπτική του πρώτου προσώπου από την επιστημονική οπτική του τρίτου προσώπου.

8.2.2.1 Μπορεί η άμεση αντίληψη να εξηγηθεί από την επιστημονική οπτική του τρίτου προσώπου;

Θα ξεκινήσουμε και πάλι από την περίπτωση της αντίληψης. Έστω ότι το υποκείμενο Α βλέπει ένα δένδρο. Αν δεχτούμε την ΑΥΘΝ, τότε στον εγκέφαλο του Α δημιουργείται μια νοητική αναπαράσταση ενός

⁴⁴ Αντικτό, επίσης, παραμένει το ερώτημα για το εάν ισχύει το αντίστροφο. Βλ. και υποσημείωση 423.

δένδρου - κάτι που μπορεί να διαπιστωθεί από ένα νευροφυσιολόγο που έχει προσαρμόσει κάποιο εγκεφαλοσκόπιο στον εγκέφαλο του Α. Μια δεύτερη μεγάλη διαφορά των δύο οπτικών είναι ότι ο Α βλέπει ένα δένδρο, ενώ ο νευροφυσιολόγος βλέπει μια *αναπαράσταση* ενός δένδρου. Για την ακρίβεια, μάλιστα, ο νευροφυσιολόγος δεν βλέπει την ίδια την νοητική αναπαράσταση όπως σχηματίζεται στον εγκέφαλο του Α, αλλά μια αναπαράσταση αυτής όπως σχηματίζεται στην οθόνη του εγκεφαλοσκόπιου. Αλλά παραβλέπουμε αυτή την περιπλοκή, γιατί όπως επισημάναμε και κατά την εξέταση της πρώτης διαφοράς, θα μπορούσαμε να φανταστούμε την οριακή περίπτωση που ο νευροφυσιολόγος παρατηρεί "δια γυμνού οφθαλμού" τον εγκέφαλο.

Η διαφορά, λοιπόν, που εντοπίζουμε έγκειται στο *περιεχόμενο των νοητικών φαινομένων*: όταν το υποκείμενο Α βλέπει ένα δένδρο, ο νευροφυσιολόγος βλέπει μια *αναπαράσταση* ενός δένδρου. Αν δεχτούμε ότι κάθε αναπαράσταση του κόσμου είναι μια θεωρητικοποίηση και τυποποίηση του κόσμου, τότε γίνεται φανερό ότι ο νευροφυσιολόγος δεν θα μπορέσει ποτέ- με την παρατήρηση των νοητικών αναπαραστάσεων και μόνο- να δει τον κόσμο όπως τον βλέπει το υποκείμενο Α. Πιο συγκεκριμένα, αυτό που θα διαφεύγει από τον νευροφυσιολόγο είναι η *άμεση* αντίληψη του κόσμου.

Επίσης, ανοιχτό παραμένει το ερώτημα για το πώς το ίδιο το υποκείμενο Α στο οποίο ανήκουν οι νοητικές αναπαραστάσεις μπορεί να βλέπει *άμεσα* τον κόσμο (όπως του εμφανίζεται από την οπτική του πρώτου προσώπου, δηλαδή στο φαινομενολογικό επίπεδο) και όχι απλώς *αναπαραστάσεις* του κόσμου. Με άλλα λόγια, το πρόβλημα είναι: πώς, ενώ στο νευροφυσιολογικό ή λειτουργικό επίπεδο υπάρχουν μόνο αναπαραστάσεις του κόσμου, στο φαινομενολογικό επίπεδο ο κόσμος μπορεί να δίδεται άμεσα.

Τσως, όμως, τελικά, αυτό το πρόβλημα να μην είναι τόσο ανυπέρβλητο όσο φαίνεται. Θα μπορούσε κανείς να αντιτείνει ότι αν, όντως, ο μόνος τρόπος που ο εγκέφαλός μας επικοινωνεί με τον κόσμο είναι με το να σχηματίζει εσωτερικές αναπαραστάσεις του κόσμου, τότε εκείνο που, από την οπτική του πρώτου προσώπου, θεωρούμε ως άμεση αντίληψη του κόσμου δεν είναι τίποτα άλλο από την αντίληψη αυτών των εσωτερικών αναπαραστάσεων. Αν, δηλαδή, ο μόνος τρόπος με τον οποίο ερχόμαστε σε επαφή με τον κόσμο είναι -σύμφωνα πάντα με την ΑΥΘΝ- μέσω νοητικών αναπαραστάσεων, τότε είναι εύλογο να υποθέσουμε ότι έχουμε συνηθίσει να εκλαμβάνουμε τις αναπαραστάσεις του κόσμου ως τον ίδιο τον κόσμο.

Η παρατήρηση αυτή είναι πολλαπλώς προβληματική. Καταρχάς, οι νοητικές αναπαραστάσεις για τις οποίες μιλά η ΑΥΘΝ, εξ ορισμού (βλ. κεφάλαιο 2, ενότητα 2.4.1), δεν είναι προσβάσιμες από τη συνείδηση, δηλαδή δεν δίδονται από την οπτική του πρώτου προσώπου. Μια δεύτερη επισήμανση, πολύ σημαντικότερη, έχει να κάνει με την ίδια τη λογική της έννοιας της αναπαράστασης. Όπως δείξαμε στο τρίτο κεφάλαιο, ένας ελάχιστος ορισμός της έννοιας της αναπαράστασης είναι ο εξής: Ένα

αντικείμενο Α συνιστά αναπαράσταση κάποιου πράγματος Β, δηλαδή έχει ως περιεχόμενο το Β, αν για κάποιον επαρκή χρήση Χ η μορφή Μ του Α αναφέρεται στο Β. Στις συνήθεις αναπαραστάσεις, δηλαδή τις εξωτερικές αναπαραστάσεις, η κατανόσή τους προϋποθέτει τη δυνατότητα άμεσης αντιληπτικής πρόσβασης στο αναπαριστώμενο Β, δηλαδή χωρίς τη βοήθεια της αναπαράστασης Α. Όπως επισημαίνει και ο Peirce (Peirce, 1967, 2.231) η κατανόηση μιας αναπαράστασης προϋποθέτει εξοικείωση με ό,τι αναπαρίσταται: "Το Σημείο μπορεί μόνο να αναπαραστήσει το Αντικείμενο και να μιλήσει περί αυτού. Δεν μπορεί να παράσχει εξοικείωση χωρίς αναγνώριση αυτού του Αντικειμένου ... δηλαδή αυτού με το οποίο προϋποθέτει μια εξοικείωση προκειμένου [το Σημείο] να μεταφέρει κάποιες επιπλέον πληροφορίες αναφορικά με [το Αντικείμενο]". Αντίθετα, στην περίπτωση των νοητικών αναπαραστάσεων, το αναπαριστώμενο δεν μπορεί να δοθεί ανεξάρτητα από αυτές. Αν ισχύει η υπόθεση της ΑΥΘΝ, τότε ο κόσμος μπορεί να δοθεί μόνο μέσω του "πέπλου" των νοητικών αναπαραστάσεων.

Θα μπορούσε, ωστόσο, κανείς να αντιτείνει ότι η θέση πως η κατανόηση μιας αναπαράστασης προϋποθέτει εξοικείωση με ό,τι αναπαρίσταται δεν είναι ορθή διότι, αν ισχυε, τότε δεν θα μπορούσαμε να κατανοούμε εξωτερικές αναπαραστάσεις φανταστικών όντων. Όμως, ακόμα και ένα φανταστικό ον αποτελείται από συστατικά πραγματικών όντων, δηλαδή όντων που μπορούν να δοθούν άμεσα στην αντίληψη. Για παράδειγμα, ένας κένταυρος είναι ένα ον που από τη μέση και πάνω είναι άνθρωπος και από τη μέση και κάτω ίππος.

Ακόμα, όμως, και αν δεχθούμε την δυσμενέστερη εκδοχή ότι υπάρχουν αναπαριστώμενα όντα που δεν αποτελούνται από συστατικά πραγματικών όντων, μια τέτοια παραδοχή σίγουρα δεν οδηγεί στην ανίχνευση της υποστηριζόμενης θέσης, αλλά στη σχετικοποίησή της, δηλαδή στο ότι η κατανόηση μιας εξωτερικής αναπαράστασης προϋποθέτει *ορισμένες φορές* την εξοικείωση με ό,τι αναπαρίσταται.

Εξάλλου, αν η κατανόηση εξωτερικών αναπαραστάσεων δεν προϋπέθετε *ποτέ* την εξοικείωση με ό,τι αναπαρίσταται, δηλαδή την δυνατότητα άμεσης αντίληψης των αναπαριστώμενων (χωρίς τη βοήθεια των εξωτερικών αναπαραστάσεων), τότε θα ετίθετο το πρόβλημα πώς είναι δυνατή η κατανόηση της διάκρισης μεταξύ αναπαριστώμενων και άμεσα αντιληπτών όντων. Αυτή η διάκριση προϋποτίθεται, άλλωστε, στην ίδια τη λογική της έννοιας της εξωτερικής αναπαράστασης: η διάκριση μεταξύ αναπαραστασιακού φορέα και αναπαριστώμενου είναι ουσιαστικά η διάκριση μεταξύ ενός αντικειμένου Α που δίδεται *άμεσα* στην αντίληψη και ενός αντικειμένου Β που δεν δίνεται άμεσα αλλά αναπαρίσταται μέσω του Α. Στις εξωτερικές αναπαραστάσεις είναι αδύνατον να διακρίνει κανείς το Β ως αναπαριστώμενο, δηλαδή ως μη σωματικά παρόν, αν

ταυτόχρονα δεν διακρίνει το Α ως σωματικά παρόν, δηλαδή ως άμεσα αντιληπτό⁴⁴⁵.

Για να συνοψίσουμε, η θέση στην οποία καταλήγουμε είναι η εξής: η κατανόηση αναπαραστάσεων προϋποθέτει τη δυνατότητα άμεσης αντίληψης αν όχι πάντα του αναπαριστώμενου, τουλάχιστον του αναπαραστασιακού φορέα. Ή, αν διατυπώναμε την ίδια θέση περισσότερο "φυσιοκρατικά": η χρήση αναπαραστάσεων προϋποθέτει τη δυνατότητα άμεσης πρόσβασης αν όχι πάντα στο αναπαριστώμενο, τουλάχιστον στον αναπαραστασιακό φορέα.

Ωστόσο, οι νοητικές αναπαραστάσεις, όπως προείπαμε, δεν είναι εξ' ορισμού προσβάσιμες από την οπτική του πρώτου προσώπου. Με άλλα λόγια, αν δεχθούμε την ΑΥΘΝ, το υποκείμενο δεν έχει άμεση αντιληπτική πρόσβαση στον φορέα των νοητικών αναπαραστάσεων. Κατά συνέπεια, το υποκείμενο δεν μπορεί να εκληφθεί ως χρήστης των νοητικών αναπαραστάσεων, διότι η χρήση αναπαραστάσεων προϋποθέτει τη δυνατότητα άμεσης πρόσβασης στον αναπαραστασιακό φορέα.

Ούτε, όμως σύμφωνα πάντα με την ΑΥΘΝ, οι γνωσιακοί υπομηχανισμοί του εγκεφάλου έχουν ποτέ άμεση πρόσβαση στα αντικείμενα του κόσμου, εφόσον ο μόνος τρόπος που μπορούν να "επικοινωνούν" με τον κόσμο είναι πάντα μέσω νοητικών αναπαραστάσεων. Ωστόσο, όπως δείξαμε παραπάνω, η χρήση αναπαραστάσεων προϋποθέτει τη δυνατότητα άμεσης πρόσβασης σε κάποια από τα αναπαριστώμενα, δηλαδή χωρίς τη βοήθεια αναπαραστάσεων. Κατά συνέπεια, ούτε οι γνωσιακοί υπομηχανισμοί του εγκεφάλου μπορούν να θεωρηθούν χρήστες των νοητικών αναπαραστάσεων.

Αν, όμως, ούτε το υποκείμενο ούτε οι γνωσιακοί του υπομηχανισμοί συνιστούν χρήστες των νοητικών αναπαραστάσεων, τότε δεν απομένει κανείς άλλος "υποψήφιος" που να μπορεί να παίξει το ρόλο του χρήστη. Κατά συνέπεια, αυτό που οι υποστηρικτές της ΑΥΘΝ ονομάζουν "νοητική αναπαράσταση" δεν μπορεί να είναι πράγματι

⁴⁴⁵ Ειδικότερα, στις εικονιστικές αναπαραστάσεις αν κανείς δεν καταφέρει να διακρίνει τον αναπαραστασιακό φορέα από το αναπαραστασιακό περιεχόμενο έχει την ψευδαίσθηση ότι αυτό που βλέπει είναι σωματικά παρόν και όχι απλώς μια απεικόνιση. Για παράδειγμα, στο πίνακα του Margitte που εικονίζεται στο εξώφυλλο της διατριβής υπάρχει ένας δεύτερος πίνακας μέσα στον πίνακα. Ωστόσο, όταν κάποιος δει για πρώτη φορά τον πίνακα του Magritte, δυσκολεύεται να διακρίνει τον δεύτερο πίνακα. Ο λόγος που συμβαίνει αυτό είναι ότι ο Magritte έχει περιορίσει δραστικά τα στοιχεία που απεικονίζουν τον φορέα του δεύτερου πίνακα. Έτσι, ο θεατής μόνο μετά από προσεκτική εξέταση αντιλαμβάνεται ή συνάγει κάποια στοιχεία που απεικονίζουν τον φορέα του δεύτερου πίνακα και έτσι διακρίνει αυτόν τον πίνακα. Τέτοια στοιχεία είναι η δεξιά πλαϊνή επιφάνεια του πίνακα, η εγκοπή στην αριστερή κourτίνα και, τέλος, το καβαλέτο που συνήθως χρησιμοποιείται για την στήριξη πινάκων.

αναπαράσταση διότι η έννοια της αναπαράστασης προϋποθέτει την ύπαρξη χρήστη της αναπαράστασης.

Η δεύτερη μεγάλη διαφορά που εντοπίσαμε μεταξύ της οπτικής του πρώτου προσώπου και της επιστημονικής οπτικής του τρίτου προσώπου οδηγεί, λοιπόν, σε μια ριζική κριτική της έννοιας της νοητικής αναπαράστασης, μια κριτική που έρχεται να προστεθεί σε όλες τις υπόλοιπες που διατυπώθηκαν στα προηγούμενα κεφάλαια. Πρόκειται για μια κριτική που δείχνει ότι η έννοια της νοητικής αναπαράστασης είναι ασυνάρτητη (incoherent).

8.2.2.2 Υπάρχει δυνατότητα άμεσης αντίληψης από την οπτική του πρώτου προσώπου;

Η συγκεκριμένη κριτική βασίζεται σε δύο παραδοχές. Η πρώτη παραδοχή, ως την ονομάσουμε παραδοχή Α, είναι ότι η έννοια της αναπαράστασης προϋποθέτει αναγκαστικά την έννοια του χρήστη. Η ορθότητα της παραδοχής Α υποστηρίχθηκε καταρχάς στο τρίτο κεφάλαιο. Ειδικότερα, σε ό,τι αφορά τις νοητικές αναπαραστάσεις, στο έβδομο κεφάλαιο είδαμε ότι όλες οι προσπάθειες προσδιορισμού της έννοιας της νοητικής αναπαράστασης χωρίς προσφυγή στην έννοια του χρήστη αποδείχτηκαν αποτυχημένες. Κατά συνέπεια, θεωρούμε δικαιολογημένη την παραδοχή Α.

Ωστόσο, υπάρχει και μια δεύτερη παραδοχή -θα την ονομάσουμε παραδοχή Β- η οποία υποστηρίχθηκε δογματικά και δεν έχει δικαιολογηθεί. Πρόκειται για την παραδοχή ότι από την οπτική του πρώτου προσώπου ο κόσμος δίδεται άμεσα. Ευσταθεί, άραγε, η παραδοχή Β; Πολλοί φιλόσοφοι απαντούν αρνητικά. Καταρχάς, τα πράγματα δεν εμφανίζονται πάντα όπως πραγματικά είναι. Για παράδειγμα, όταν βλέπω ένα μολύβι μισοβυθισμένο σε ένα ποτήρι νερό, το βλέπω λυγισμένο, δηλαδή έχω τη ψευδαίσθηση (illusion) ότι είναι λυγισμένο. Ωστόσο το μολύβι δεν είναι όντως λυγισμένο, αλλά φαίνεται λυγισμένο λόγω της διάθλασης του φωτός. Άρα, λοιπόν, σε αυτή την περίπτωση δεν μπορεί να είναι το πραγματικό μολύβι αυτό που βλέπω ως λυγισμένο, αλλά κάτι άλλο, δηλαδή κάτι που δεν υφίσταται στον εξωτερικό κόσμο αλλά είναι μια εσωτερική αναπαράσταση μέσα στο νου μου⁴⁴⁶. Με άλλα λόγια, αυτό που πραγματικά βλέπω από την οπτική του πρώτου προσώπου, δηλαδή του φαινομενολογικό επίπεδο, δεν είναι το πραγματικό μολύβι, αλλά μια νοητική αναπαράσταση του μολυβιού. Αυτή η υπόθεση, εξάλλου, ενισχύεται και από την περίπτωση των παραισθήσεων (hallucinations): όταν έχω παραισθήσεις βλέπω πράγματα που δεν υφίστανται καν στον

⁴⁴⁶ Πρόκειται, ουσιαστικά, για τις παραστάσεις (Vorstellungen) ή ιδέες που συναντάμε στη νεώτερη φιλοσοφία (βλέπε κεφάλαιο 3, ενότητα 3.1). Για λόγους απλότητας του κειμένου θα εξακολουθήσουμε να χρησιμοποιούμε τον όρο "αναπαράσταση" με τον προσδιορισμό ότι αφορούν το φαινομενολογικό επίπεδο.

εξωτερικό κόσμο. Έτσι, για παράδειγμα, με τη βοήθεια παραισθησιογόνων μπορεί κάποιος να βλέπει πράγματα που κανείς άλλος δεν βλέπει. Τι άλλο θα μπορούσε, λοιπόν, να δικαιολογεί τις παραισθήσεις αν όχι η υπόθεση ότι αυτοί που βλέπουμε στις συγκεκριμένες περιπτώσεις είναι εσωτερικές αναπαραστάσεις; Αν, τώρα, λάβουμε υπόψη μας ότι οι παραισθήσεις και οι ψευδαισθήσεις είναι φαινομενολογικά ίδιες με την κανονική αντίληψη⁴⁴⁷, τότε μπορούμε να συμπεράνουμε ότι και στην κανονική αντίληψη αυτό που πραγματικά αντιλαμβανόμαστε δεν είναι τα αντικείμενα του εξωτερικού κόσμου αλλά νοητικές αναπαραστάσεις τους. Κατά συνέπεια, στο φαινομενολογικό επίπεδο η πρόσβασή μας στον κόσμο δεν είναι άμεση αλλά επιτυγχάνεται μέσω νοητικών αναπαραστάσεων⁴⁴⁸. Με άλλα λόγια, δεν έχουμε άμεση πρόσβαση στον κόσμο ούτε από την οπτική του πρώτου προσώπου⁴⁴⁹.

⁴⁴⁷ Με την έκφραση "κανονική αντίληψη" μεταφράζουμε, κάπως ελεύθερα, την αγγλική έκφραση "veridical perception".

⁴⁴⁸ Πρόκειται για το "περιβόητο" επιχείρημα με βάση την ψευδαίσθηση (argument from illusion).

⁴⁴⁹ Οι εμπειριστές του 17ου αιώνα, για παράδειγμα, υποστήριζαν ότι αυτό στο οποίο έχουμε άμεση πρόσβαση είναι εντυπώσεις ή ιδέες (βλ. κεφάλαιο 3, ενότητα 3.1). Για παράδειγμα, ο Locke (1984, βιβλίο II, κεφάλαιο VIII, σελ. 8) δηλώνει: "Οτιδήποτε αντιλαμβάνεται ο νους ή οτιδήποτε αποτελεί άμεσο αντικείμενο αισθητηριακής αντίληψης, σκέψης ή νόησης, το αποκαλώ *ιδέα*". Σύμφωνα με τον Thomas Reid (1785/1969, σελ. 124), "όλοι οι φιλόσοφοι, από τον Πλάτωνα ως τον κ. Hume, συμφωνούν σε αυτό, ότι το άμεσο αντικείμενο της αντίληψης πρέπει να είναι κάποιο είδωλο που παρουσιάζεται στο νου". Αντίστοιχα, οι εμπειριστές του 20ου αιώνα μίλησαν για αισθητηριακά δεδομένα (sense data): "Έτσι, καθίσταται φανερό ότι το πραγματικό τραπέζι, αν υπάρχει κάτι τέτοιο, δεν είναι το ίδιο με αυτό του οποίου έχουμε άμεσα την εμπειρία μέσω της όρασης ή της αφής ή της ακοής. Το πραγματικό τραπέζι, αν υπάρχει τέτοιο, δεν μας γίνεται με κανέναν τρόπο *άμεσα* γνωστό, αλλά πρέπει να το συναγάγουμε από ό,τι μας είναι άμεσα γνωστό ... Ας δώσουμε το όνομα 'αισθητηριακά δεδομένα' στα πράγματα εκείνα που μας είναι άμεσα γνωστά στις αισθήσεις: στα χρώματα, τους ήχους, τις οσμές, τη σκληρότητα, την τραχύτητα κ.ο.κ." (Ράσσελ 1963, σελ. 11). Ο Ayer (1987, σελ. 309) επισημαίνει ανφορικά με το βιβλίο του Russell *Τα Προβλήματα της Φιλοσοφίας* ότι "...απλώς συνεχίζει την βρετανική εμπειριστική παράδοση: είναι ένα βιβλίο που απορρέει κατευθείαν από τους Locke, Berkeley και Hume ξεκινώντας από μια θεωρία αντίληψης που συγγενεύει πολύ με αυτή του Berkeley. Η θεωρία προεβδεί ότι αυτοί που αντιλαμβανόμαστε δεν είναι τραπέζια και καρέκλες και πράγματα αυτού του τύπου, αλλά ό,τι ο Locke αποκαλεί '*απλές ιδέες*', και ό,τι ο Russell, ακολουθώντας τον φίλο του Moore, αποκαλεί '*αισθητηριακά δεδομένα*'. Βλ., επίσης, Ayer 1956, σελ. 89: "[Αυτό που] δίδεται άμεσα στην αντίληψη είναι ένα φευγαλέο [evanescent] αντικείμενο που ονομάζεται *ιδέα*". Για μια συστηματική διερεύνηση της παραστασιοκρατίας (representationalism) στον Καρτέσιο και τον Leibniz βλ. Anapolitanos 1998, κεφάλαιο πρώτο.

8.2.2.2.1 Στο φαινομενολογικό επίπεδο δεν εμφανίζονται νοητικές αναπαράστάσεις

8.2.2.2.1.1 Κατά την κανονική αντίληψη δεν εμφανίζονται νοητικές αναπαράστάσεις

Το συγκεκριμένο επιχείρημα είναι προβληματικό από πολλές απόψεις. Για το λόγο αυτό θα επιχειρήσουμε να διακρίνουμε σταδιακά τα διάφορα προβλήματα που παρουσιάζει. Καταρχάς, αν δεχθεί κανείς ότι η κανονική αντίληψη ενός αντικειμένου απαιτεί την αντίληψη μιας εσωτερικής αναπαράστασης του αντικειμένου στο φαινομενολογικό επίπεδο, τότε η αντίληψη της εσωτερικής αναπαράστασης του αντικειμένου απαιτεί, με τη σειρά της, την αντίληψη μιας εσωτερικής αναπαράστασης της εσωτερικής αναπαράστασης του αντικειμένου κ.ο.κ. Έτσι, η υιοθέτηση εσωτερικών αναπαράστάσεων στο φαινομενολογικό επίπεδο οδηγεί σε επ' άπειρον αναδρομή⁴⁵⁰. Άρα κατά την αντίληψη ενός αντικειμένου δεν έχουμε πρόσβαση σε εσωτερικές αναπαράστάσεις⁴⁵¹. Γενικότερα, αυτό που δείχνει το επιχείρημα είναι ότι η αντίληψη- με την έννοια της κατανόησης- δεν μπορεί να εξηγηθεί μέσω της έννοιας της αναπαράστασης: κατοχή μιας αναπαράστασης δεν σημαίνει κατανόησή της⁴⁵². Θα επανέλθουμε, όμως, σε αυτό το σημείο στον Επίλογο. Προς το παρόν, εκείνο που θέλουμε να κρατήσουμε από το παραπάνω επιχείρημα είναι ότι, τουλάχιστον όσον αφορά την κανονική αντίληψη, έχουμε άμεση πρόσβαση στα ίδια τα αντικείμενα και όχι σε εσωτερικές αναπαράστάσεις των αντικειμένων, διότι διαφορετικά οδηγούμαστέ σε επ' άπειρον αναδρομή: Αν για να αντιληφθώ ένα εξωτερικό αντικείμενο πρέπει να σχηματίσω μια εσωτερική αναπαράσταση αυτού, τότε για να αντιληφθώ την εσωτερική αναπαράσταση πρέπει να σχηματίσω μια νέα εσωτερική αναπαράσταση κ.ο.κ.

Επίσης, αν σταθούμε στην ίδια τη φαινομενολογία της αντίληψης, καθίσταται φανερό ότι κατά την αντίληψη ενός αντικειμένου όσο κι' αν

⁴⁵⁰ "Εάν η γνώση εν γένει είναι η σύλληψη ενός αντικειμένου-εικόνα ως μιας εμμενούς εικόνας ενός υπερβατικού πράγματος έξω από τη συνείδηση, πώς πρέπει να συλληφθεί τότε το ίδιο το υπερβατικό αντικείμενο; Αν κάθε σύλληψη αντικειμένων είναι συνείδηση εικόνας, τότε χρειάζομαι και για την εμμενή εικόνα εκ νέου ένα πράγμα-εικόνα, που παρουσιάζει μέσα μου το εμμενές πράγμα-εικόνα κ.ο.τ" (Heidegger 1999, σελ. 77). Βλέπε, επίσης, Husserl (1983, σελ. 243) και Wittgenstein (1953).

⁴⁵¹ Προφανώς, αυτό το επιχείρημα είναι γενικό και αφορά κάθε είδους αντίληψη - με τη γενικότερη έννοια του όρου. Τόσο κατά την κανονική αντίληψη, όσο και κατά την παραίσθηση ή τη φαντασία αυτό στο οποίο έχουμε πρόσβαση δεν είναι νοητικές αναπαράστάσεις αντικειμένων αλλά τα ίδια τα αντικείμενα. Θα εξετάσουμε διεξοδικότερα αυτό το ζήτημα στη συνέχεια.

⁴⁵² Διαφορετικά, η διαταγή "φαντάσου ένα πράσινο ποτήρι" δεν θα γινόταν κατανοητή παρά μόνο μετά την εκτέλεσή της (βλ. Glock 1996, σελ. 374).

προσπαθήσουμε δεν θα διαπιστώσουμε κάποια νοητική αναπαράσταση του αντικειμένου -και μιλάμε πάντα για νοητικές αναπαραστάσεις του φαινομενολογικού επιπέδου. Όπως είδαμε στο τρίτο κεφάλαιο, ένα αντικείμενο μπορεί να χαρακτηριστεί αναπαράσταση, αν μπορεί να γίνει σε αυτό η διάκριση μεταξύ φορέα και περιεχομένου. Αυτό το κριτήριο, προσαρμοσμένο στην οπτική του πρώτου προσώπου, σημαίνει ότι αντιλαμβάνομαι ένα πράγμα ως αναπαράσταση, αν μπορώ να εστιάσω την προσοχή μου άλλοτε στον φορέα του πράγματος και άλλοτε στο περιεχόμενο του πράγματος. Για παράδειγμα, όταν αντιλαμβάνομαι ένα πράγμα ως την εικόνα ενός ποτηριού, τότε μπορώ να εστιάσω την προσοχή μου άλλοτε στο φορέα της εικόνας (π.χ. στη ποιότητα του χαρτιού) και άλλοτε στο περιεχόμενο της εικόνας, δηλαδή στο ίδιο το ποτήρι. Ωστόσο, κατά την αντίληψη ενός αντικειμένου κάτι τέτοιο δεν είναι δυνατόν. Και δεν είναι δυνατόν γιατί δεν έχει νόημα η διάκριση ανάμεσα σε φορέα και περιεχόμενο στην περίπτωση της αντίληψης αντικειμένων. Έτσι, μπορώ να εστιάσω την προσοχή μου, για παράδειγμα, σε μια λεπτομέρεια του ποτηριού που έχω απέναντί μου, αλλά γνωρίζω πάντα ότι πρόκειται για κάποια λεπτομέρεια του ποτηριού. Αντίθετα, στην περίπτωση μιας εικόνας του ποτηριού, όπως είδαμε, μπορώ να εστιάσω την προσοχή μου όχι μόνο σε κάποιες λεπτομέρειες του ποτηριού, αλλά και σε λεπτομέρειες του φορέα της αναπαράστασης (π.χ. την ποιότητα του χαρτιού) οι οποίες δεν έχουν καμία σχέση με το ποτήρι.

Άρα, λοιπόν, ούτε από φαινομενολογική άποψη ευσταθεί η θέση ότι κατά την αντίληψη αντικειμένων αυτό που πραγματικά αντιλαμβανόμαστε είναι μια νοητική αναπαράσταση του φαινομενολογικού επιπέδου. Με άλλα λόγια, η συγκεκριμένη θέση δεν υποστηρίζεται φαινομενολογικά.

Όμως γιατί θα έπρεπε να δώσουμε τόση βαρύτητα στη φαινομενολογία της αντίληψης; Δεν θα ήταν εύλογο να θεωρήσουμε ότι, αν αυτό που μας δίδεται από την οπτική του πρώτου προσώπου είναι όντως μια νοητική αναπαράσταση των αντικειμένων, τότε είναι αδύνατον το συγκεκριμένο γεγονός να διαπιστωθεί μέσα από την ίδια την οπτική του πρώτου προσώπου;

Μια τέτοια ένσταση μας ξαναγυρίζει ένα βήμα πίσω: αν οι εν λόγω νοητικές αναπαραστάσεις δεν δίνονται από την οπτική του πρώτου προσώπου, τότε δεν είναι νοητικές αναπαραστάσεις του φαινομενολογικού επιπέδου, αλλά νοητικές αναπαραστάσεις του λειτουργικού ή του φυσιολογικού επιπέδου, δηλαδή νοητικές αναπαραστάσεις που μπορούν να διαπιστωθούν από την επιστημονική οπτική του τρίτου προσώπου. Όπως όμως έχουμε μέχρι στιγμής δείξει, δεν έχει νόημα να μιλάμε για νοητικές αναπαραστάσεις του λειτουργικού ή του φυσιολογικού επιπέδου παρά μόνο αν ισχύει η θέση ότι από την οπτική του πρώτου προσώπου δεν έχουμε άμεση πρόσβαση στον κόσμο, ή, ισοδύναμα, ότι υπάρχουν νοητικές αναπαραστάσεις στο φαινομενολογικό επίπεδο οι οποίες διαμεσολαβούν κάθε είδους αντίληψη. Άρα, λοιπόν, σε αυτό το στάδιο του επιχειρήματος δεν δικαιούται κανείς να οπισθοχωρήσει στην έννοια της

νοητικής αναπαράστασης του λειτουργικού ή του φυσιολογικού επιπέδου. Η μόνη διέξοδος που υπάρχει για να υποστηριχθεί η θέση ότι υπάρχουν νοητικές αναπαραστάσεις στα προαναφερθέντα επίπεδα περιγραφής είναι να υποστηριχθεί αυτόνομα η θέση ότι υπάρχουν νοητικές αναπαραστάσεις στο φαινομενολογικό επίπεδο οι οποίες διαμεσολαβούν κάθε είδους αντίληψη. Εξάλλου, δεν πρέπει να ξεχνάμε ότι το ίδιο το επίχειρημα με βάση την ψευδαίσθηση στηρίζεται καθοριστικά στη φαινομενολογία της κανονικής αντίληψης και της ψευδαίσθησης, ενώ δεν βασίζεται σε τεκμήρια που προκύπτουν από το νευροφυσιολογικό ή το λειτουργικό επίπεδο.

8.2.2.1.2 Κατά την παραίσθηση δεν εμφανίζονται νοητικές αναπαραστάσεις

Μέχρι στιγμής, λοιπόν, έχουμε δείξει ότι, τουλάχιστον όσον αφορά την περίπτωση της κανονικής αντίληψης, δεν υπάρχουν νοητικές αναπαραστάσεις στο φαινομενολογικό επίπεδο. Με δεδομένο αυτό το συμπέρασμα, μπορούμε τώρα να δεξούμε ότι η φαινομενολογία των ψευδαισθήσεων και των παραισθήσεων που προϋποτίθεται στο επίχειρημα με βάση την ψευδαίσθηση είναι λανθασμένη. Θα ξεκινήσουμε με μια επισήμανση του Heidegger (1982, σελ. 63):

"Εάν στο σκοτάδι εκλάβω λανθασμένα ένα δένδρο για άνθρωπο, θα ήταν λάθος να πω ότι αυτή η αντίληψη κατευθύνεται προς ένα δένδρο αλλά το εκλαμβάνει ως άνθρωπο, ούτε ότι ο άνθρωπος είναι μια απλή αναπαράσταση και ότι, κατά συνέπεια, σε αυτή την ψευδαίσθηση κατευθύνομαι προς μια αναπαράσταση. Αντίθετα, το νόημα της ψευδαίσθησης είναι ακριβώς ότι με το να εκλαμβάνω το δένδρο για άνθρωπο συλλαμβάνω ό,τι αντιλαμβάνομαι και ό,τι πιστεύω ότι αντιλαμβάνομαι ως ένα παρευρισκόμενο ον. Σε αυτή την αντιληπτική ψευδαίσθηση μου δίδεται ο ίδιος ο άνθρωπος και όχι, ως πούμε, μια αναπαράσταση του ανθρώπου."⁴⁵³

Θεωρούμε ότι αν κανείς παραμείνει προσηλωμένος στην ίδια την φαινομενολογία της ψευδαίσθησης, τότε η παραπάνω επισήμανση του Heidegger είναι προφανής. Ωστόσο, επειδή συχνά οι θεωρητικές μας προκαταλήψεις συσκοτίζουν τα φαινομενολογικά δεδομένα, θα ενισχύσουμε την συγκεκριμένη παρατήρηση με ένα επίχειρημα ανάλογο με εκείνο που χρησιμοποιήσαμε παραπάνω για την περίπτωση της κανονικής αντίληψης. Όταν έχω την παραίσθηση ότι βλέπω έναν άνθρωπο δεν έρχομαι σε "επαφή" με μια εικόνα του ανθρώπου που εμφανίζεται στο νου μου. Διότι, αντίθετα με ό,τι συμβαίνει κατά την αντίληψη εξωτερικών εικόνων, στην περίπτωση της παραισθήσης αλλά και, γενικότερα, της

⁴⁵³ Βλ., επίσης, Heidegger 1982, σελ. 60.

φαντασίας⁴⁵⁴, δεν μπορώ κατά βούληση να εστιάσω την προσοχή μου άλλοτε στο περιεχόμενο και άλλοτε στον φορέα της υποτιθέμενης εσωτερικής εικόνας. Αυτό που βλέπω ή φαντάζομαι είναι τον ίδιο τον άνθρωπο όπως αν τον έβλεπα στην πραγματικότητα. Είναι αδύνατον να εστιάσω την προσοχή μου στον φορέα της υποτιθέμενης εσωτερικής εικόνας. Άρα, λοιπόν, στο φαινομενολογικό επίπεδο οι υποτιθέμενες (νοητικές) εικόνες δεν έχουν κάποιο φορέα που μπορεί να διαχωριστεί από το περιεχόμενό τους. Γι' αυτό, άλλωστε, και δεν μπορούν να μας εμφανισθούν στην φαντασία αμφίσημα σχήματα ως τέτοια, δηλαδή σχήματα που λόγω της μορφής του φορέα τους φαίνονται να εικονίζουν άλλοτε ένα πράγμα και άλλοτε κάτι άλλο⁴⁵⁵. Για παράδειγμα, δεν μπορούμε να φανταστούμε το γνωστό αμφίσημο σχήμα λαγού/πάπιας ως απεικονίζον αποκλειστικά μια πάπια και στη συνέχεια ως απεικονίζον αποκλειστικά ένα λαγό. Ο μόνος τρόπος που μπορούμε να το φανταστούμε είναι ως το αμφίσημο σχήμα λαγού/πάπιας⁴⁵⁶. Άρα, λοιπόν, από φαινομενολογική σκοπιά, η ψευδαισθήση, η παραίσθηση και η φαντασία συγγενεύουν περισσότερο με την κανονική αντίληψη αντικειμένων, παρά με την κανονική αντίληψη εξωτερικών αναπαραστάσεων αντικειμένων⁴⁵⁷.

Με το παραπάνω επιχειρήμα δείξαμε ότι κατά την ψευδαισθήση, την παραίσθηση και, γενικότερα, την φαντασία συμβαίνει ό,τι και κατά την κανονική αντίληψη: δεν έχουμε πρόσβαση σε νοητικές αναπαραστάσεις. Όταν έχουμε ψευδαισθήσεις, παραισθήσεις ή φανταζόμαστε κάτι, ο κόσμος μάς εμφανίζεται εξίσου άμεσα, όπως συμβαίνει στην κανονική αντίληψη. Συνολικά, λοιπόν, δείξαμε ότι δεν υπάρχουν νοητικές αναπαραστάσεις στο φαινομενολογικό επίπεδο.

8.2.2.2.2 Στο φαινομενολογικό επίπεδο δεν εμφανίζονται εσωτερικά αντικείμενα

⁴⁵⁴ Φυσικά η φαντασία δεν συνοδεύεται πάντα από κάποια εποπτεία, ωστόσο στο παρόν κεφάλαιο όταν χρησιμοποιούμε τον όρο "φαντασία" θα αναφερόμαστε στενά στην εποπτική φαντασία. Πρόκειται για μια ικανότητα που εμπίπτει σε ό,τι στο πέμπτο κεφάλαιο, αποδεχόμενοι προσωρινά την αναπαραστασιακή θεωρία του νου, ονομάσαμε "χρήση νοητικών ειδώλων" (imagery).

⁴⁵⁵ Βλ. Reisberg & Chambers 1991. Για μια σύνοψη των ερευνητικών αποτελεσμάτων γύρω από τη δυνατότητα εμφάνισης αμφίσημων σχημάτων στη φαντασία βλ. Pylyshyn 2001.

⁴⁵⁶ Πρβλ. Putnam 1994a, σελ. 490.

⁴⁵⁷ Το ίδιο ισχύει και για την ανάμνηση (που συνοδεύεται από εποπτεία). Για παράδειγμα ο Ryle, αναφερόμενος σε ένα σκακιστή που παίζει έχοντας δεμένα τα μάτια του, επισημαίνει "...όταν, αφού πασχίσει να θυμηθεί τις θέσεις των πιονιών, ο σκακιστής καταφέρει τελικά να θυμηθεί, τότε η θέαση αυτών με τα μάτια του νου του, εάν όντως κάνει κάτι τέτοιο, δεν είναι κάτι μέσω του οποίου αρχίζει να θυμάται. Είναι ο στόχος και όχι το μέσον της προσπάθειάς του να θυμηθεί" (Ryle 1971, σελ. 398). Βλ., επίσης, Heidegger 1979, σελ. 35.

Μία ένσταση που μπορεί να εγερθεί σε αυτό το σημείο είναι η εξής: το γεγονός ότι δεν έχουμε πρόσβαση σε εσωτερικές αναπαραστάσεις αντικειμένων δεν αποκλείει το ενδεχόμενο ότι αυτό στο οποίο έχουμε άμεση πρόσβαση είναι κάποια εσωτερικά αντικείμενα. Για παράδειγμα, στην περίπτωση της παραίσθησης βλέπω κάτι που δεν είναι ενεργεια υπαρκτό (actual, wirklich), δηλαδή σωματικά παρόν. Σε κάθε τέτοια περίπτωση, αυτό που βλέπω αποκλείεται να ανήκει στον (ενεργεια υπαρκτό) περιβάλλοντα κόσμο. Πού θα μπορούσε, λοιπόν, να βρίσκεται αυτό που βλέπω αν όχι μέσα στο νου μου; Με άλλα λόγια, η ένσταση που εγείρεται είναι ότι με το να δείξουμε ότι η αντίληψη -με τη γενικότερη έννοια του όρου (κανονική αντίληψη, παραίσθηση, ψευδαίσθηση, φαντασία κ.λπ.)⁴⁵⁸ - είναι άμεση, δείξαμε απλώς ότι η αντίληψη δεν διαμεσολαβείται από νοητικές αναπαραστάσεις στο φαινομενολογικό επίπεδο. Ωστόσο, η ύπαρξη άμεσης αντίληψης δεν ισοδυναμεί, κατ' ανάγκη, με την αντίληψη εξωτερικών αντικειμένων. Ενδέχεται, λοιπόν, η άμεση αντίληψη να αποτελεί άμεση αντίληψη εσωτερικών και όχι εξωτερικών αντικειμένων - με άλλα λόγια, ενδέχεται να αποτελεί ένα είδος ενδοσκοπησης.

8.2.2.2.1 Η φαινομενολογία της παραίσθησης διαφέρει από τη φαινομενολογία της κανονικής αντίληψης

Μέχρι τώρα τονίσαμε τις (φαινομενολογικές) ομοιότητες των ψευδαισθήσεων, των παραισθήσεων και της φαντασίας με την κανονική αντίληψη. Ο στόχος ήταν να εκμεταλλευτούμε αυτές τις ομοιότητες για να δείξουμε ότι όλες οι παραπάνω μορφές αντίληψης -με τη γενικότερη έννοια του όρου- δεν διαμεσολαβούνται από εσωτερικές αναπαραστάσεις στο φαινομενολογικό επίπεδο. Ωστόσο, τώρα θα πρέπει να εστιάσουμε την προσοχή μας στις μεταξύ τους διαφορές προκειμένου να αμφισβητήσουμε τη θέση ότι όλες αυτές οι μορφές αντίληψης αποτελούν, ουσιαστικά, αντίληψη εσωτερικών αντικειμένων. Δηλαδή, ο στόχος μας, σε πρώτη φάση, είναι να ανακόψουμε τη δυνατότητα γενίκευσης σε όλα τα είδη αντίληψης (με τη γενικότερη έννοια του όρου) της θέσης ότι η παραίσθηση αποτελεί περίπτωση αντίληψης εσωτερικών αντικειμένων. Αυτή η γενίκευση στηρίζεται αποκλειστικά στη θέση ότι η παραίσθηση είναι φαινομενολογικά ίδια με κάθε άλλο είδος αντίληψης και ιδίως με την κανονική αντίληψη. Εφόσον, λοιπόν, η παραίσθηση, συνεχίζει το επιχειρήμα με βάση την ψευδαίσθηση, είναι απαραίλακτη φαινομενολογικά με την κανονική αντίληψη, θα πρέπει και η τελευταία να αποτελεί αντίληψη κάποιου παρόμοιου εσωτερικού αντικειμένου⁴⁵⁹.

⁴⁵⁸ Δηλαδή, με την έννοια της εποπτικής πρόσβασης στον κόσμο.

⁴⁵⁹ Είναι αξιοσημείωτο ότι όλη η ισχύς του επιχειρήματος με βάση την ψευδαίσθηση αντλείται από "δεδομένα" που δίδονται μόνο από την οπτική του πρώτου προσώπου. Σε μια φιλοσοφική παράδοση όπως η αναλυτική φιλοσοφία,

Μία πρώτη αντίρρηση σε αυτό το επιχείρημα είναι ότι αν, τελικά, δεν υπήρχαν διαφορές μεταξύ παραίσθησης και κανονικής αντίληψης, δεν θα ήταν δυνατή η διάκρισή τους, ούτε θα υπήρχαν οι δύο αυτές διαφορετικές έννοιες και οι αντίστοιχοι γλωσσικοί όροι. Η παραίσθηση μπορεί να γίνει κατανοητή μόνο αν κανείς κατανοεί ήδη το τι σημαίνει κανονική αντίληψη. Διότι, μόνο έτσι μπορεί κανείς να αναγνωρίσει μια αντιληπτική κατάσταση ως παραίσθηση, δηλαδή συγκρίνοντας την με περιπτώσεις κανονικής αντίληψης⁴⁶⁰. Όπως λέει και ο Merleau-Ponty (1962, σελ. 295): "γνωρίζουμε ότι υπάρχουν λάθη μόνο επειδή κατέχουμε αλήθειες, στο όνομα των οποίων διορθώνουμε τα λάθη και τα αναγνωρίζουμε ως λάθη"⁴⁶¹.

Η παραίσθηση και η κανονική αντίληψη έχουν, πράγματι, πολλές και σημαντικές διαφορές. Καταρχάς, όταν έχω την παραίσθηση ενός δένδρου κανείς άλλος δεν μπορεί να δει το δένδρο (και αυτό είναι κάτι που καταρχήν μπορώ να το γνωρίζω από τις αντιδράσεις των άλλων στη συμπεριφορά μου), ενώ όταν βλέπω κανονικά ένα δένδρο τότε μπορούν και άλλοι να δουν το δένδρο. Επίσης, μια συγκεκριμένη παραίσθηση, σε αντίθεση με την κανονική αντίληψη, σπάνια επαναλαμβάνεται και σίγουρα δεν επαναλαμβάνεται κατά βούληση. Για παράδειγμα, όταν υπάρχει ένα δένδρο στην αυλή του σπιτιού μου μπορώ καθημερινά να το βλέπω όποτε θέλω, απλά με το να βγαίνω στην εξώπορτα, κάτι που δεν ισχύει, φυσικά, στην περίπτωση της παραίσθησης⁴⁶².

όπου κάθε αναφορά στην οπτική του πρώτου προσώπου θεωρείται εκ προοιμίου υποοπτη, είναι ενδιαφέρον να διαπιστώνουμε ότι ορισμένοι αναλυτικοί φιλόσοφοι βασίζουν τις θεωρίες τους σε ένα τέτοιου τύπου "δεδομένο".

⁴⁶⁰ Βλ. Austin 1962, σελ. 11 και Kant 1983a, A377: "...χωρίς κατ' αίσθηση αντίληψη ακόμα και η πλαστή επινόηση και το όνειρο δεν είναι δυνατά". Πρβλ, επίσης, Merleau-Ponty 1977, σελ. 34-35: "...αν μπορώ να μιλάω για 'όνειρα' και για 'πραγματικότητα', να διερωτώμαι ως προς τη διάκριση του φαντασιακού και του πραγματικού, και να δυσπιστώ για το 'πραγματικό', τούτο συμβαίνει επειδή αυτή η διάκριση έχει κιόλας γίνει από μένα πριν από την ανάληψη, επειδή έχω μιαν εμπειρία για το πραγματικό καθώς και για το φαντασιακό, και το πρόβλημα είναι τότε όχι ν' αναζητήσω πώς η κριτική σκέψη μπορεί να παρέχει δευτερογενείς αντιστοιχίες τούτης της διάκρισης, αλλά να διαλευκάνω την αρχέγονη γνώση μας του 'πραγματικού', να περιγράψω την αντίληψη του κόσμου που θεμελιώνει δια παντός την ιδέα που έχουμε για την αλήθεια".

⁴⁶¹ Επίσης, αν η παραίσθηση δεν μπορούσε να διακριθεί από την κανονική αντίληψη, τότε δεν θα ήμασταν σίγουροι όχι μόνο για το αν αυτό που αντιλαμβανόμαστε είναι όντως πραγματικό, αλλά και για το αν είναι όντως παραισθητικό. Με άλλα λόγια, η αμφιβολία είναι συμμετρική.

⁴⁶² Ακόμα και αν παίρνω καθημερινά το ίδιο παραισθησιόγonio δεν έχω πάντα την ίδια παραίσθηση.

Σε τι διαφέρει, όμως, η ίδια η *φαινομενολογία* της κανονικής αντίληψης από εκείνη της παραίσθησης⁴⁶³; Ένα λάθος που γίνεται, το οποίο αποκρύπτει τη διαφορά ανάμεσα στην κανονική αντίληψη και την παραίσθηση, είναι ότι αυτές κατανοούνται ως στιγμιακά και στατικά φαινόμενα. Όμως, όταν αντιλαμβάνομαι κάτι δεν έχω την εμπειρία μιας σειράς διαδοχικών αντικειμένων -ένα για κάθε χρονική στιγμή - που μ' ένα πρόσθετο ενέργημα τα εκλαμβάνω ως το ίδιο αντικείμενο⁴⁶⁴. Η στιγμιακή και στατική αντίληψη αποτελεί μια *αφαίρεση* και όχι ένα φαινομενολογικό δεδομένο. Η αντίληψη έχει μια συνέχεια και μια ροή -αυτός είναι ο πρωταρχικός τρόπος που μας δίδεται. Αν η κανονική αντίληψη και η παραίσθηση νοηθούν ως στιγμιακά και στατικά φαινόμενα⁴⁶⁵, τότε

⁴⁶³ Υποστηρίζοντας τη θέση ότι η φαινομενολογία της κανονικής αντίληψης διαφέρει από εκείνη της παραίσθησης διαφοροποιούμαστε από τις δύο αντίταλες μεταξού τους φιλοσοφικές θεωρίες αντίληψης που κυριαρχούν στη σύγχρονη φιλοσοφία του νου: την αιτιακή θεωρία αντίληψης (Grice 1961) και την διαζευκτική θεώρηση των φαινομένων (Snowdon 1981, McDowell 1983). Σύμφωνα με την πρώτη, η αντίληψη ενέχει μια υποκειμενική εμπειρία η οποία είναι λογικά διακριτή από το αντιληπτό αντικείμενο. Αν η υποκειμενική εμπειρία δεν σχετίζεται αξιόπιστα με το εμφανιζόμενο αντικείμενο τότε δεν έχουμε κανονική αντίληψη αλλά παραίσθηση. Συνεπώς, τόσο η κανονική αντίληψη όσο και η παραίσθηση είναι μη διακρίσιμες φαινομενολογικά διότι αυτό στο οποίο έχουμε πρόσβαση από την οπτική του πρώτου προσώπου είναι το ίδιο, δηλαδή η υποκειμενική εμπειρία. Σύμφωνα, με τη διαζευκτική θεώρηση των φαινομένων, η αντιληπτική εμπειρία δεν είναι διακρίσιμη από το αντιληπτό αντικείμενο. Κατά την κανονική αντίληψη, αντιλαμβανόμαστε άμεσα τον κόσμο και όχι κάποια αντιληπτική εμπειρία - όπως συμβαίνει στην περίπτωση της παραίσθησης (βλ. McDowell 1994, σελ. 26). Κατά την κανονική αντίληψη και την παραίσθηση μάς εμφανίζονται δύο διαφορετικές από οντολογική άποψη καταστάσεις πραγμάτων. Ωστόσο, γνωσιολογικά το υποκείμενο δεν μπορεί να τις διακρίνει, διότι η φαινομενολογία της κανονικής αντίληψης και της παραίσθησης είναι η ίδια.

⁴⁶⁴ Όπως, επίσης, σε κάθε χρονική στιγμή δεν έχω την εμπειρία δύο διδιάστατων εικόνων που τις εκλαμβάνω ως ένα τρισδιάστατο αντικείμενο.

⁴⁶⁵ Μια τέτοια παραδοχή, για παράδειγμα, υπόκειται στα πειράματα τυχροσκοπικής παρουσιασής ερεθισμάτων που γίνονται κατά κόρον στο πλαίσιο της κλασικής γνωσιακής επιστήμης. Σ' αυτά τα πειράματα προβάλλονται σχήματα για χρονική περίοδο πολύ μικρότερη εκείνης που απαιτείται για την αντίδραση του ματιού σε ένα ερέθισμα (150ms). Ωστόσο, τέτοια πειράματα "δουλεύουν" μόνο για πολύ οικεία ερεθίσματα. Για παράδειγμα, σε ένα πείραμα των Nazir & O'Regan (1990) ζητήθηκε από τα υποκείμενα να μάθουν να ξεχωρίζουν τρία άγνωστα σύμβολα υπό την προϋπόθεση ότι θα κοιτούν στο μέσον του εκάστοτε εμφανιζόμενου συμβόλου και δεν θα μετακινούν το βλέμμα τους. Η μηχανή προβολής ήταν συνδεδεμένη με ένα μηχάνημα ανίχνευσης της κίνησης των ματιών και είχε κατά τέτοιο τρόπο ρυθμιστεί ώστε να σταματά την προβολή του ερεθίσματος κάθε φορά που ο παρατηρητής κούνησε τα μάτια του. Το αποτέλεσμα του πειράματος ήταν ότι τα υποκείμενα βρήκαν την όλη διαδικασία τρομερά διασπαστική και ενοχλητική, ενώ χρειάστηκαν εκατοντάδες επαναλήψεις για να μάθουν τελικά να

αποκόπτονται από τη συμπεριφορά του υποκειμένου, τους τρόπους που πράττει και αντιδρά. Ειδικότερα, η αποσύνδεση της αντίληψης από τη σωματική δραστηριότητα του υποκειμένου αποκρύπτει την αδιάσπαστη σχέση που έχουν μεταξύ τους, το ότι, δηλαδή, η μια ρυθμίζει την άλλη. Το σημαντικό, εν προκειμένω, δεν είναι ότι η αντίληψη καθοδηγεί τη σωματική δραστηριότητα, αλλά το αντίστροφο: ότι η σωματική δραστηριότητα καθοδηγεί την αντίληψη. Ένα υποκείμενο, ακόμα και όταν αντιλαμβάνεται με τον πιο παθητικό τρόπο τον κόσμο, δεν είναι ποτέ απαλλαγμένο από κάποιες μορφές σωματικής κίνησης. Ακόμα και όταν, βυθισμένος μέσα στη πολυθρόνα μου, "χαζεύω" από το παράθυρο τον έξω κόσμο, τα μάτια μου κινούνται διαρκώς ανάλογα με το που εστιάσω και όταν αυτό δεν είναι αρκετό, με ανεπαισθητες κινήσεις του κεφαλιού μου επιχειρώ να βρω τη βέλτιστη γωνία θέασης. Η συγκεκριμένη παρατήρηση είναι περισσότερο προφανής στην περίπτωση της αφής. Για παράδειγμα, για να καταλάβουμε την υφή μιας επιφάνειας τρίβουμε ελαφρά το χέρι μας πάνω σε αυτή. Χωρίς την κίνηση, η αντίληψη της υφής της επιφάνειας σχεδόν εξαφανίζεται. Καθ' όμοιο τρόπο, η οπτική αντίληψη πρέπει να κατανοηθεί ως μια μορφή οπτικής "ψαύσης": για να αντιληφθούμε ένα αντικείμενο χρειάζεται να το "χειριστούμε" κινώντας τα μάτια μας και το σώμα μας σε σχέση με αυτό. Η αντίληψη δεν είναι μια παθητική κατάσταση του οργανισμού που απλώς συνεπικουρείται από κάποιες μορφές σωματική δραστηριότητα. Η αντίληψη αποτελεί η ίδια μια μορφή σωματικής δραστηριότητας και, υπ' αυτή την έννοια, έχει αναγκαία μια χρονική διάρκεια⁴⁶⁶.

Αν θεωρήσουμε κατ' αυτόν τον τρόπο την αντίληψη, τότε προκύπτουν αρκετές και σημαντικές διαφορές μεταξύ κανονικής

διακρίνουν τα σύμβολα. Επίσης άλλα πειράματα που βασίζονται στην ακινητοποίηση του αμφιβληστροειδή δείχνουν ότι αυτή η ακινητοποίηση έχει ως συνέπεια την απώλεια της αντίληψης χρωμάτων και περιγραμμάτων. Πρόσθετα επιχειρήματα και πειράματα κατά της θέσης ότι η αντίληψη αποτελεί ένα στιγμιαίο και στατικό φαινόμενο παρατίθενται στο O'Regan & Noë 2001. Βλέπε, επίσης, Wheeler 1994 και McClamrock 1995 για πρόσθετα φιλοσοφικά επιχειρήματα υπέρ της θέσης ότι η αντίληψη είναι ενεργητική (active).

⁴⁶⁶ Υπάρχουν αρκετά ευρήματα που ενισχύουν από γενετική άποψη αυτή τη θέση. Καταρχάς έχει παρατηρηθεί ότι γάτες οι οποίες αμέσως μετά τη γέννησή τους στερφήκαν τη δυνατότητα ενεργητικής οπτικής εξερεύνησης του περιβάλλοντος δεν κατάφεραν να αναπτύξουν τις πιο βασικές μορφές οπτικο-κινητικού συντονισμού και αυτό θεωρείται μια μορφή τύφλωσης (Held & Hein 1963). Επίσης, ένα άλλο εύρημα αφορά τη συμπεριφορά εκ γεννητής τυφλών που με εγχείρηση αποκτούν την ικανότητα να έχουν οπτική επαφή με το περιβάλλον (βλ. Sacks 1995a). Σε αυτές τις περιπτώσεις τα υποκείμενα καταφέρνουν μετά από πολύ χρόνο και προσπάθεια να διακρίνουν και να αναγνωρίζουν αντικείμενα στο οπτικό τους πεδίο. Τα μάτια τους, όντας προηγουμένως αδρανολοποιημένα, στην αρχή κινούνται εντελώς ακανόνιστα κατά τρόπο που δεν τους επιτρέπει να διακρίνουν με σαφήνεια τα όρια των αντικειμένων.

αντίληψης και παραίσθησης. Καταρχάς, στην περίπτωση της κανονικής αντίληψης αυτό που αντιλαμβάνομαι έχει έναν σχεδόν ανεξάντλητο πλούτο τον οποίο μπορώ να γνωρίζω μέσω της παρατήρησης, δηλαδή μιας εξερεύνησης που προχωρά σταδιακά με τη βοήθεια μεταβολών της στάσης μου. Αντίθετα, στην περίπτωση της παραίσθησης δεν μπορώ να μάθω τίποτα από αυτήν που δεν το γνωρίζω ήδη από τη στιγμή που μου εμφανίζεται η παραίσθηση. Γι' αυτό, ίσως, δεν έχει νόημα να πούμε ότι κατά την παραίσθηση υπάρχει δυνατότητα να πέσω θύμα *γευδαίσθησης*. Τα αντικείμενα που εμφανίζονται κατά την παραίσθηση δεν είναι, κάθε στιγμή, τίποτε παραπάνω απ' ό,τι τα εκλαμβάνουμε ότι είναι. Έτσι, δεν υπάρχει περιθώριο λάθους.

Επίσης, μια παραίσθηση μόνο σε ελάχιστες περιπτώσεις μπορεί να διατηρηθεί για πολύ. Συνήθως, αυτό που αντιλαμβανόμαστε μέσω μιας από τις αισθήσεις, έστω της όρασης, μετά από λίγο τίθεται υπό αμφισβήτηση, λόγω πληροφοριών που παίρνουμε από τις άλλες τροπικότητες (modalities). Για παράδειγμα, αν έχω την παραίσθηση ότι βλέπω ένα δένδρο, τότε η αντίληψη αυτή κλονίζεται αμέσως μόλις επιχειρήσω να το πιάσω γιατί αυτό που μου δίδεται από την αίσθηση της αφής δεν συμβιβάζεται με ό,τι βλέπω όπως, επίσης, κλονίζεται από τα λεγόμενα των άλλων οι οποίοι δεν βλέπουν το δένδρο. Ακόμα, όμως, και αν δεν ανατρέξω στη βοήθεια των άλλων αισθήσεων, πολύ σπάνια μια παραίσθηση έχει τη σταθερότητα και τη διάρκεια που διαθέτει η αντίληψη ενός πραγματικού αντικειμένου. Όταν βλέπω κανονικά ένα δένδρο, η αντίληψη που έχω μεταβάλλεται με έναν ομαλό τρόπο ανάλογα με το πόσο γρήγορα πλησιάζω ή απομακρύνομαι από αυτό. Το αντικείμενο και το υπόβαθρο εντός του οποίου εμφανίζεται, όντας ανεξάρτητα από εμένα, μεταβάλλονται με συγκεκριμένο τρόπο συναρτήσει της κίνησής μου. Αντίθετα, στην παραίσθηση το δένδρο δεν μεταβάλλεται με τον ίδιο ομαλό τρόπο σε σχέση με την κίνηση του σώματός μου. Γενικότερα, η παραίσθηση δεν αποτελεί μια εμπειρία που υποκαθιστά *πλήρως* την κανονική αντίληψη, αλλά μάλλον *υπερτίθεται* στην τελευταία. Όπως επισημαίνει ο Merleau-Ponty (1962, σελ. 339):

"Οι παραισθήσεις παίζονται σε μια διαφορετική σκηνή από εκείνη του αντιληπτού κόσμου και κατά ένα τρόπο υπερτίθενται: 'Ξέρεις', λέει ένας ασθενής, 'ενόσω συζητάμε, κάποιος μου λέει μια το ένα και μια το άλλο. Από πού να προέρχονται όλα αυτά;' Το γεγονός ότι η παραίσθηση δεν λαμβάνει κάποια θέση στον σταθερό και διυποκειμενικό κόσμο σημαίνει ότι στερείται της πληρότητας, της εσωτερικής άρθρωσης που κάνει το πραγματικό πράγμα να αναπαύεται / διαμένει (αγγλ. *reside*, γαλ. *repose*) 'εν εαυτώ', και να ενεργεί και να υπάρχει από μόνο του"

Ο ασθενής παράλληλα με την παραίσθηση ότι ακούει φωνές, ακούει και το συνομιλητή του. Και όχι μόνο αυτό. Βλέπει ότι βρίσκεται σε ένα δωμάτιο και ότι έχει απέναντί του το γιατρό, αισθάνεται την καρέκλα πάνω στην οποία κάθεται, οσφραίνεται τη μυρωδιά του δωματίου. Το ίδιο συμβαίνει και στις οπτικές παραισθήσεις: ο ασθενής που βλέπει ροζ

ποντική στον τοίχο έχει προφανώς τη δυνατότητα να βλέπει και τον τοίχο που βλέπουμε όλοι μας. Ο ασθενής, με άλλα λόγια, δεν είναι αποκομμένος αντιληπτικά από τον κόσμο, οι παραισθήσεις του απλώς επικαθόνται στον πραγματικό κόσμο⁴⁶⁷ και δεν τον υποκαθιστούν⁴⁶⁸. Έτσι η παραισθήση συνυπάρχει με την κανονική αντίληψη, αλλά με έναν τρόπο που δεν την καθιστά ενσωματώσιμη στη τελευταία. Η παραισθήση, παρότι ορισμένες φορές εμφανίζεται ως εξίσου "πραγματική"⁴⁶⁹ με τον αντιληπτό κόσμο, παραμένει, για τους λόγους που αναφέραμε στην προηγούμενη παράγραφο, εκτός του αντιληπτού κόσμου. Ο ασθενής, παρότι επιχειρεί να την ενσωματώσει μεταβάλλοντας τις πεποιθήσεις του για τον κόσμο, δεν κατορθώνει τελικά να ενσωματώσει την παραισθήση στον άμεσα αντιληπτό κόσμο αλλά μόνο στις θεωρίες του για τον κόσμο⁴⁷⁰.

Τέλος, πρέπει να λάβουμε υπόψη δύο ακόμη στοιχεία. Το πρώτο είναι ότι οι παραισθήσεις αποτελούν συνήθως παθολογικές καταστάσεις και υπ' αυτή την έννοια δεν αφορούν άμεσα τα κανονικά υποκείμενα. Φυσικά υπάρχουν και παραισθήσεις που δεν θεωρούνται παθολογικές όπως το μετείκασμα. Ωστόσο, σε αυτές τις περιπτώσεις το υποκείμενο δεν έχει καμιά αμφιβολία ότι αυτό που βλέπει δεν ανήκει στον κόσμο. Γι' αυτό τα παραδείγματα που χρησιμοποιούνται από τους φιλοσόφους όταν επιχειρούν να υποστηρίξουν το επιχείρημα με βάση τη ψευδαίσθηση είναι παραδείγματα παθολογικών παραισθήσεων. Σ' αυτές τις περιπτώσεις ωστόσο, δεν νομιμοποιείται η εξαγωγή συμπερασμάτων για το κανονικό υποκείμενο χωρίς να λαμβάνεται υπόψη ότι οι (παθολογικές) παραισθήσεις δεν χαρακτηρίζουν το κανονικό υποκείμενο.

⁴⁶⁷ Αυτή η θεώρηση πιστεύουμε ότι δικαιολογεί πλήρως την μετάφραση του όρου "hallucination" με τον όρο "παραισθήση". Η πρόθεση "παρά" τονίζει ότι αυτή η δραστηριότητα λαμβάνει χώρα δίπλα και, άρα, εκτός της αντίληψης.

⁴⁶⁸ Βέβαια, υπάρχουν και περιπτώσεις ακραίου παραληρήματος όπου ο ασθενής φαίνεται να ζει εντελώς αποκομμένος από τον αντιληπτό κόσμο. Το ίδιο, ενδεχομένως, συμβαίνει ορισμένες φορές κατά τη λήψη ισχυρών παραισθησιογόνων ουσιών. Η συγκεκριμένη περιγραφή, ωστόσο, γίνεται από την οπτική του τρίτου προσώπου και δεν είναι δυνατόν να επαληθευθεί, διότι ο ασθενής/χρήστης ουσιών είναι τόσο απομονωμένος στον εαυτό του που κάθε επικοινωνία μαζί του καθίσταται αδύνατη. Εξάλλου, αν μπορούσε να επικοινωνεί με άλλους ανθρώπους κατά τη διάρκεια του παραληρήματος, τότε η παραισθήση δεν θα ήταν καθολική γιατί θα είχε πρόσβαση στον αντιληπτό κόσμο, δηλαδή στα λόγια του συνομιλητή του. Ούτε, φυσικά, η μαρτυρία αυτού του ασθενούς/χρήστη ουσιών όταν "επανεέλθει" μπορεί να θεωρηθεί αξιόπιστη, διότι η μνήμη του δεν είναι, εν προκειμένω, περισσότερο ασφαλής.

⁴⁶⁹ "Πραγματική" με μια προθεωρητική έννοια του όρου, δηλαδή όπως θα χρησιμοποιούταν από έναν ασθενή. "Πραγματική" σε ένα τέτοιο πλαίσιο θα μπορούσε να σημαίνει "εξίσου επιβλητική με τον αντιληπτό κόσμο".

⁴⁷⁰ Για την αναγκαιότητα υποστήριξης της διάκρισης μεταξύ αντιληπτικών και μη αντιληπτικών πεποιθήσεων βλέπε McDowell 2000.

Το δεύτερο στοιχείο είναι πραγματολογικό: οι παραισθήσεις δεν είναι καθόλου συχνές. Πολύ σπάνια πάσχουμε από παραισθήσεις και όποτε συμβαίνει κάτι τέτοιο αποδίδεται, όπως είπαμε, σε κάποιο πρόβλημα υγείας: κάτι δεν παγαίνει καλά με εμάς και όχι με τον κόσμο. Εξάλλου, πόσοι από 'μας έχουμε πραγματικά πέσει θύματα παραισθήσεων; Τις περισσότερες φορές μιλάμε για την εμπειρία των παραισθήσεων βασιζόμενοι σε μια στερεοτυπική ιδέα που έχουμε για αυτές- μια ιδέα η οποία αντλείται από την εμπειρία φαινομένων που τα θεωρούμε παρεμφερή, όπως οι ψευδαισθήσεις και τα όνειρα⁴⁷¹. Ειδικότερα, οι ψευδαισθήσεις είναι πράγματι ένα πολύ συχνό φαινόμενο. Όλο για πέφτουμε καθημερινά θύματα τέτοιων ειδών πλάνης (βλέπε, για παράδειγμα, τις οφθαλμαπάτες). Όμως, οι ψευδαισθήσεις δεν πρέπει να συγχέονται με τις παραισθήσεις⁴⁷². Οι ψευδαισθήσεις είναι κάτι αντικειμενικό, υπό την έννοια ότι όταν έχω μια ψευδαισθηση βλέπω κάτι που και οι άλλοι μπορούν να δουν, έστω κι αν αυτό που βλέπω διαφέρει από το πώς είναι όντως τα πράγματα. Η ψευδαισθηση Mueller-Lyer είναι κάτι που οποιoσδήποτε (τουλάχιστον του δυτικού πολιτισμού) μπορεί να

⁴⁷¹ Η περίπτωση των ονείρων απαιτεί από μόνη της μια ξεχωριστή ανάλυση καθώς θέτει πολύ σημαντικά προβλήματα σε κάθε θεωρία άμεσης αντίληψης: Αν η φαινομενολογία της κανονικής αντίληψης δεν διαφέρει από εκείνη του ονείρου, τότε είναι αδύνατον να διακρίνουμε αν ο κόσμος που αντιλαμβανόμαστε είναι πραγματικός ή ονειρικός. Στην παρούσα εργασία δεν θα εξετάσουμε την ορθότητα αυτής της θέσης. Θα παραπέμψουμε απλά στο άρθρο του Jeffrey Gordon (1985), ο οποίος δείχνει, αρκετά πειστικά, ότι η φαινομενολογία της κανονικής αντίληψης διαφέρει τελικά από εκείνη του ονείρου: ο ονειρευόμενος "στερείται κάθε αίσθησης μιας άμεσης επαφής με τις δυνάμεις του σώματός του" αισθάνεται, αντίθετα, αποκομμένος από αυτές με μια αγεφύρωτη απόσταση, και έτσι το εάν ή όχι αυτές υλοποιούν την θέλησή του αποτελεί εξ' ολοκλήρου ένα ζήτημα εκ περισσοῦ. Ο ονειρευόμενος ποτέ δεν επιφέρει μια κατάσταση πραγμάτων που επιθυμεί, αλλά μάλλον επιθυμεί και η κατάσταση πραγμάτων παράγεται (ή δεν παράγεται)" (Gordon 1985, σελ. 187). Η παρατήρηση αυτή ενισχύει την άποψη ότι η διαφορά της φαινομενολογίας της κανονικής αντίληψης από τη φαινομενολογία του ονείρου μπορεί, ως ένα βαθμό, να εντοπιστεί σε μια διαφοροποίηση της συναίσθησης της ικανότητας του πράττειν. Ένας λόγος που η συναίσθηση της ικανότητας του πράττειν διαφοροποιείται στην περίπτωση του ονείρου είναι, ενδεχομένως, ότι το σώμα του ονειρευόμενου βρίσκεται σε κατάσταση παράλυσης. Επίσης, στο ίδιο άρθρο ο Gordon δείχνει ότι η φαινομενολογία του ονείρου διαφέρει και από την φαινομενολογία της παραισθήσης: "Στην παραισθήση εγώ ο ίδιος διακρίνω τις φωνές ή τις μορφές μου από εκείνες του αντιληπτού κόσμου. Είμαι αρκετά πρόθυμος να αναγνωρίσω ότι αυτές υπάρχουν μόνο για μένα, ενώ τα ονειρικά αντικείμενα δεν προβάλλουν έναντι κάποιας αντιπαθέμενης τάξης" (Gordon 1985, σελ. 178).

⁴⁷² Βλ. Austin 1962, σελ. 23.

δει και, μάλιστα, κατ' επανάληψη⁴⁷³. Σε μια ψευδαισθήση δεν βλέπουμε κάτι που δεν υπάρχει, αλλά βλέπουμε κάτι που όντως υπάρχει, αλλά του αποδίδουμε ιδιότητες που δεν έχει. Γι' αυτό και η φαινομενολογία των ψευδαισθήσεων παρουσιάζει μια σταθερότητα που προσεγγίζει πολύ τη σταθερότητα της κανονικής αντίληψης. Μια σταθερότητα που, για παράδειγμα, επιτρέπει την παρατήρηση, δηλαδή μια εξερεύνηση που προχωρά σταδιακά με τη βοήθεια μεταβολών της στάσης μου. Γενικότερα, οι παραισθήσεις, σε αντίθεση με τις ψευδαισθήσεις, δεν συναρμολόγουν με τις πράξεις μας κατά έναν τρόπο που να τις καθιστά τμήμα του αντίληπτου κόσμου - ένα γεγονός που μπορεί κανείς να παραβλέψει αν, απερίσκεπτα, επιχειρήσει να περιγράψει την φαινομενολογία των παραισθήσεων στηριζόμενος σε εκείνη των ψευδαισθήσεων. Δεν είναι τυχαίο, εξάλλου, ότι το βασικό επίχειρμα με το οποίο αμφισβητείται η αμεσότητα της αντίληψης, παρότι βασίζεται στο φαινόμενο της παραισθήσης, ονομάζεται "επίχειρμα με βάση την ψευδαισθήση" και όχι "επίχειρμα με βάση την παραισθήση".

Εφόσον, λοιπόν, για να ανακεφαλαιώσουμε, η φαινομενολογία της κανονικής αντίληψης (και της ψευδαισθήσης) διαφέρει από εκείνη της παραισθήσης, δεν είναι δικαιολογημένη η επέκταση και στην περίπτωση της κανονικής αντίληψης της θέσης ότι κατά την παραισθήση έχουμε άμεση αντίληψη εσωτερικών αντικειμένων. Η κανονική αντίληψη αποτελεί μια άμεση πρόσβαση στα ενεργεία υπαρκτά αντικείμενα του εξωτερικού κόσμου.

Ωστόσο, μια εύλογη ένσταση σε αυτό το σημείο είναι ότι από το γεγονός ότι η φαινομενολογία της κανονικής αντίληψης διαφέρει από εκείνη της παραισθήσης δεν συνεπάγεται άμεσα ότι η κανονική αντίληψη είναι αντίληψη εξωτερικών αντικειμένων. Θα μπορούσε κάλλιστα να είναι αντίληψη διαφορετικού είδους εσωτερικών αντικειμένων. Τι είναι, λοιπόν, εκείνο που καθιστά την κανονική αντίληψη αντίληψη εξωτερικών αντικειμένων;

Θεωρούμε ότι αυτό το ερώτημα έχει απαντηθεί έμμεσα κατά την παρουσίαση των διαφορών που έχει η φαινομενολογία της κανονικής αντίληψης από εκείνη της παραισθήσης: Η κανονική αντίληψη αποτελεί αντίληψη εξωτερικών αντικειμένων διότι τα συγκεκριμένα αντικείμενα παρουσιάζουν μια αντίσταση, μια αυτονομία και μια σταθερότητα μέσα στο χρόνο η οποία εκδηλώνεται κατά την συναρμολόγηση της αντίληψης με την πράξη. Έτσι, καθίσταται δυνατή η εξερεύνηση των εμφανιζόμενων αντικειμένων και τα αντίληπτικά λάθη.

Με τα παραπάνω επίχειρήματα έχουμε δείξει ότι κατά την κανονική αντίληψη δεν μας δίνονται ούτε νοητικές αναπαραστάσεις του

⁴⁷³ Γι' αυτό το λόγο οι ψευδαισθήσεις δεν θεωρούνται ζήτημα υγείας: κάτι δεν πηγαίνει καλά με τον κόσμο και όχι με 'μας.

φαινομενολογικού επιπέδου ούτε εσωτερικά αντικείμενα, αλλά τα ίδια τα εξωτερικά αντικείμενα. Η θέση αυτή ισοδυναμεί με την παραδοχή Β που θέλαμε να αποδείξουμε ότι, δηλαδή, από την οπτική του πράτου προσώπου ο κόσμος δίδεται άμεσα (βλ. ενότητα 8.2.2.2). Κατά συνέπεια, θεωρούμε δικαιολογημένη τη θέση ότι η χρήση αναπαραστάσεων προϋποθέτει τη δυνατότητα άμεσης πρόσβασης αν όχι πάντα στο αναπαριστώμενο, τουλάχιστον στον αναπαραστασιακό φορέα. Το ίδιο δικαιολογημένη είναι επομένως και η κριτική που, στηριζόμενοι σε αυτή τη θέση, ασκήσαμε στην έννοια της νοητικής αναπαράστασης του λειτουργικού ή του νευροφυσιολογικού επιπέδου: η εν λόγω έννοια είναι ασυνάρτητη. Πρόκειται για μια ριζική κριτική η οποία, προστιθέμενη σε όλες τις υπόλοιπες που διατυπώθηκαν στα προηγούμενα κεφάλαια, καθιστά την έννοια της νοητικής αναπαράστασης εντελώς ανυπόστατη.

8.2.2.2.2 Κατά την παραίσθηση δεν εμφανίζονται εσωτερικά αντικείμενα

Παράλληλα με την κριτική που ασκήσαμε στο επιχείρημα με βάση την ψευδαίσθηση προχωρήσαμε και σε μια θετική προσέγγιση των νοητικών φαινομένων. Προκειμένου να συνεχίσουμε προς αυτήν την κατεύθυνση, θα εξετάσουμε τώρα το πρόβλημα του πώς πρέπει να νοηθούν τα αντικείμενα που εμφανίζονται στην παραίσθηση και στη φαντασία. Καταρχάς, όσον αφορά το πρόβλημα της παραίσθησης, θεωρούμε ότι έχουμε πει αρκετά κατά την εξέταση της πρώτης μεγάλης διαφοράς μεταξύ των δύο οπτικών, ώστε να δεχθούμε τη θέση ότι η παραίσθηση είναι μια μορφή φαντασίας από την οποία έχει εκπέσει η συναίσθηση της ικανότητας του πράττειν. Αυτό είναι σχετικά προφανές στην περίπτωση περιπτώσεων κατά τις οποίες εμφανίζονται (εξωτερικές) γλωσσικές αναπαραστάσεις (συνήθως ηχητικές) που αναφέρονται σε ομιλιακά ενεργήματα. Για παράδειγμα, όταν "ακούω" μια προσταγή αυτή μπορεί να είναι μια δικιά μου πρόθεση από την οποία έχει "απολεσθεί" η συναίσθηση της ικανότητας του πράττειν. Αντιμέτωπος με ένα τέτοιο φαινομενολογικό δεδομένο, ένας τρόπος να το εκλογικεύσω είναι να το αποδώσω σε κάποιον άλλον. Αυτό σημαίνει ότι το κατανοώ ως μια αντίληψη κάποιας φωνής που ακούω, δηλαδή ως κάτι έτερο προς εμένα το οποίο προέρχεται από τον εξωτερικό κόσμο, από κάποιον άλλο.

Τι γίνεται, όμως, με την περίπτωση πιο απλών παραισθήσεων; Όταν βλέπω ροζ ποντίκια στον τοίχο τότε δεν πρόκειται για μια νοητική κατάσταση από την οποία έχει "απολεσθεί" η συναίσθηση της ικανότητας του πράττειν, διότι είμαι βέβαιος ότι εγώ είμαι αυτός που βλέπει τα ροζ ποντίκια. Ωστόσο, θεωρούμε ότι και σε αυτή την περίπτωση μπορεί να εφαρμοστεί η υπόθεση ότι η παραίσθηση είναι ένα είδος φαντασίας από το οποίο έχει εκπέσει η συναίσθηση της ικανότητας του πράττειν. Ιδού πώς: όταν φαντάζομαι ένα ποντίκι χωρίς να έχω συναίσθηση ότι εγώ είμαι αυτός που το φαντάζεται, τότε αυτό που απομένει είναι ότι μου εμφανίζεται ένα ποντίκι που δεν το κατανοώ ως δικό μου κατασκεύασμα. Πώς αλλιώς, λοιπόν, μπορεί να ερμηνευτεί (και να εκλογικευτεί) αυτό το

φαινομενολογικό δεδομένο αν όχι ως μια κανονική αντίληψη του ποντικιού, δεδομένου ότι μόνο κατά την κανονική αντίληψη τα εμφανιζόμενα όντα κατανοούνται ως μη-"παραγόμενα" από το υποκείμενο;

Αν δεχθούμε την παραπάνω υπόθεση, η παραίσθηση μπορεί να αναχθεί στην φαντασία. Εξάλλου, η σχέση που ισχυριζόμαστε ότι υφίσταται μεταξύ παραίσθησης και φαντασίας δικαιολογεί και την αστάθεια που παρουσιάζουν οι παραισθήσεις, μια αστάθεια που χαρακτηρίζει, κατ' εξοχήν, τη φαντασία. Επίσης, δικαιολογεί και τη μη ταύτιση της φαινομενολογίας των παραισθήσεων με εκείνη της κανονικής αντίληψης: στην περίπτωση των παραισθήσεων η συνείδηση της ικανότητας του πράττειν προστίθεται δευτερογενώς και δεν έχει την αμεσότητα που χαρακτηρίζει την κανονική αντίληψη.

8.2.2.2.3 Στην φαντασία δεν εμφανίζονται εσωτερικά αντικείμενα

Ένα βασικό πρόβλημα που απομένει, παραμένοντας πάντα σε μια πρώτη προσέγγιση, είναι πώς πρέπει να κατανοηθούν τα όντα που μας εμφανίζονται στην φαντασία⁴⁷⁴. Μια άποψη θα μπορούσε να είναι ότι όταν φαντάζομαι ένα αντικείμενο, φαντάζομαι κάποιο ον που δεν είναι ενεργεία υπαρκτό, δηλαδή σωματικά παρόν. Κατά συνέπεια, το συγκεκριμένο ον δεν ανήκει στον ενεργεία υπαρκτό περιβάλλοντα κόσμο, αλλά είναι ένα εσωτερικό αντικείμενο.

Όμως εσωτερικό ως προς τι; Σίγουρα όχι ως προς τον εγκέφαλο, διότι όσο και αν προσπαθήσουμε δεν μπορούμε να δούμε ή, γενικότερα, να αισθανθούμε τον εγκέφαλό μας. Εξάλλου, αυτό που εξετάζουμε τώρα είναι το πώς μας εμφανίζονται τα όντα στο φαινομενολογικό επίπεδο, δηλαδή από την οπτική του πρώτου προσώπου. Πρέπει, λοιπόν, να ρωτήσουμε: υπό ποια έννοια θα μπορούσαν τα όντα που μας εμφανίζονται στη φαντασία να είναι εσωτερικά; Εσωτερικά ως προς τι; Μία δεύτερη απάντηση, ενδεχομένως πιο εύλογη, είναι ότι βρίσκονται μέσα στο νου μας. Αλλά τι πραγματικά μπορεί να σημαίνει αυτό; Πώς είναι δυνατόν να γίνονται αντιληπτά αυτά τα εσωτερικά αντικείμενα; Με ποια μάτια τα βλέπουμε, πώς φωτίζονται, που ακριβώς βρίσκονται;

Όταν μας εμφανίζεται, με τη γενικότερη έννοια του όρου, κάποιο αντικείμενο⁴⁷⁵, δηλαδή όταν αντιλαμβανόμαστε, θυμόμαστε, προσδοκούμε, φανταζόμαστε κ.λπ. (εποπτικά) κάποιο αντικείμενο, αυτό το αντικείμενο μπορεί να είναι, μεταξύ άλλων, ενεργεία υπαρκτό ή φανταστικό. Σε κάθε περίπτωση, πάντως, τα εμφανιζόμενα αντικείμενα

⁴⁷⁴ Ένα άλλο μεγάλο πρόβλημα είναι το πώς πρέπει να κατανοηθούν τα όντα που μας εμφανίζονται στη μνήμη. Για περισσότερα επ' αυτού βλ. Casey 1983.

⁴⁷⁵ Δεν υπονοούμε ότι μπορούν να μας εμφανίζονται μόνο αντικείμενα, απλώς για λόγους απλότητας δεν εξετάζουμε την περίπτωση εμφάνισης καταστάσεων πραγμάτων.

νοούνται ως αντικείμενα του εξωτερικού κόσμου –ανεξάρτητα από το αν πρόκειται για έναν ενεργεια υπαρκτό ή φανταστικό κόσμο. Με άλλα λόγια, η σχέση που έχω με τα όντα που μου εμφανίζονται κατά την κανονική αντίληψη ή τη φαντασία είναι μια σχέση εξωτερικότητας - ετερότητας, γι' αυτό και δεν μπορούν να είναι δικά μου κατά τον τρόπο που ένα νοητικό φαινόμενο είναι δικό μου. Όπως δείξαμε και στην ενότητα 8.2.1.1, η συνείδηση της πατρότητας δεν χαρακτηρίζει τα αντικείμενα που μου εμφανίζονται, αλλά τα νοητικά φαινόμενα μέσω των οποίων τα αντικείμενα εμφανίζονται. Τα παρουσιαζόμενα όντα μου εμφανίζονται πάντα σε ένα πλαίσιο άλλων (του ίδιου τύπου από οντολογική άποψη) όντων, έναντι ενός υποβάθρου⁴⁷⁶. Αυτός είναι ο "ορίζοντας" μέσα στον οποίον μου εμφανίζονται τα πράγματα. Ακόμα και στην περίπτωση που φαντάζομαι ένα κένταυρο, φαντάζομαι ένα ον που ζει σε κάποιο φανταστικό κόσμο και όχι μέσα στο νου μου. Με δυο λόγια, όταν αναφερόμαστε (με την γενικότερη έννοια του όρου) σε κάποιο αντικείμενο, αυτό δεν είναι ποτέ αντικείμενο του εσωτερικού μας "κόσμου", γιατί ο εσωτερικός μας κόσμος (αν υιοθετήσουμε, προς στιγμήν, αυτόν τον όρο) δεν αποτελείται από αντικείμενα. Δεν υπάρχουν εσωτερικά αντικείμενα⁴⁷⁷. Υπ' αυτή την έννοια είναι πλεονασμός- ίσως και παραπλανητικό- το να μιλάμε για "εξωτερικά" αντικείμενα⁴⁷⁸.

⁴⁷⁶ Εδώ αναφερόμαστε ακροθιγώς σε ένα σημαντικό ζήτημα όσον αφορά την φαινομενολογία της αντίληψης: η πιο απλή μορφή αντίληψης έχει τη δομή σχήμα - υπόβαθρο. Ο Merleau-Ponty (1962, σελ. 4) σημειώνει χαρακτηριστικά ότι "[το σχήμα πάνω σε υπόβαθρο] δεν αποτελεί ένα τυχαίο χαρακτηριστικό της γεγονικής αντίληψης... Είναι ακριβώς ο ορισμός του φαινομένου της αντίληψης, αυτό χωρίς το οποίο ένα φαινόμενο δεν μπορεί να λεχθεί ότι συνιστά καν αντίληψη". Με άλλα λόγια, υπάρχει μια ακαθοριστία που αποτελεί ουσιαστικό χαρακτηριστικό κάθε αντίληψης: "Η περιοχή που περικλείει το οπτικό πεδίο δεν είναι εύκολο να περιγραφεί, αλλά εκείνο που είναι σίγουρο είναι ότι δεν είναι ούτε μαύρη ούτε γκριζα. Εδώ λαμβάνει χώρα μια ακαθόριστη όραση..." (Merleau-Ponty 1962, σελ. 6). Το σημαντικό εν προκειμένω είναι ότι η ακαθοριστία που παρουσιάζει η αντίληψη δεν είναι αναπαραστάσιμη: κάθε απόπειρα απεικόνισης του αντιληπτικού υποβάθρου αποτελεί απεικόνιση τού τι υπάρχει στο περιβάριο του αντιληπτικού πεδίου, όχι όμως και μια απεικόνιση του τρόπου με τον οποίο εμφανίζεται στην αντίληψη. Επίσης, ένα άλλο απαραίτητο στοιχείο της δομής της αντίληψης είναι ότι σε κάθε αντίληψη προποσιθεται πάντα, κατά ένα σιωπηρό τρόπο, η επίγνωση του σώματος του αντιληπτικού υποκειμένου. Πάλι με τα λόγια του Merleau-Ponty (1962, σελ. 101): "το ίδιο το σώμα [του υποκειμένου] είναι ο τρίτος όρος, πάντα σιωπηρά κατανοημένος, στη δομή σχήμα-υπόβαθρο, και κάθε σχήμα προβάλλει έναντι του διπλού ορίζοντα του εξωτερικού και του εσωτερικού χώρου" (βλ. κεφάλαιο 4, ενότητα 4.3.3). Στο παρόν κεφάλαιο, προκειμένου να διατηρήσουμε όσο το δυνατόν απλούστερη την ανάλυση, δεν θεματοποιούμε ρητά αυτά τα δύο στοιχεία.

⁴⁷⁷ Όσον αφορά την αντίληψη ο Husserl επισημαίνει χαρακτηριστικά: "Αντιλαμβανόμαστε το πράγμα, το αντικείμενο της φύσης, το δένδρο εκεί στον κήπο·

Το όλο πρόβλημα που δημιουργείται οφείλεται, ενδεχομένως, και στον τρόπο με τον οποίο αναφερόμαστε στα αντικείμενα της αντίληψης και της φαντασίας. Λέμε, για παράδειγμα, ότι, όταν φαντάζομαι κάτι, μου εμφανίζεται ένα φανταστικό αντικείμενο, ενώ, όταν αντιλαμβάνομαι κάτι, μου εμφανίζεται ένα πραγματικό (real) αντικείμενο. Εκφραζόμενοι κατ' αυτόν τον τρόπο, καθιστούμε την εξεταζόμενη διαφορά διαφορά των ιδίων των αντικειμένων, ενώ πρωταρχικά αποτελεί διαφορά του τρόπου με τον οποίον σχετιζόμαστε με τα αντικείμενα. Μόνο αν θεωρήσουμε κατά τον πρώτο τρόπο τη διαφορά αρχίζει να φαίνεται εύλογη η διολίσθηση από τη διάκριση μεταξύ πραγματικών και φανταστικών αντικειμένων στη διάκριση μεταξύ εξωτερικών και εσωτερικών αντικειμένων. Καταρχάς, αν ακολουθήσουμε τον Kant⁴⁷⁹, αντί του χαρακτηρισμού "πραγματικό" πρέπει να χρησιμοποιούμε τον χαρακτηρισμό "ενεργεία υπαρκτό" (actual, wirklich), διότι το "ενεργεία υπαρκτό" δεν ονομάζει μια ιδιότητα του ίδιου του αντικειμένου (όπως κάνει το "πραγματικό"), αλλά μια σχέση μεταξύ του υποκειμένου και του αντικειμένου. Θεωρούμε, ωστόσο, ότι ένας ορθότερος τρόπος να εκφράσουμε αυτή τη σχέση είναι με τη χρήση επιρρηματικού προσδιορισμού: αντί να λέμε ότι "μας εμφανίζεται ένα ενεργεία υπαρκτό ή ένα φανταστικό αντικείμενο", να λέμε ότι "μας εμφανίζεται ενεργεία ή φανταστικά ένα αντικείμενο". Υπ' αυτή την έννοια, και η ενδοσκοπήση-η υποτιθέμενη εσωτερική αντίληψη- αντικειμένων δεν πρέπει να θεωρηθεί ως μια πρόσβαση σε εσωτερικά αντικείμενα, αλλά ως πρόσβαση σε αντικείμενα με μη ενεργεία τρόπο.

Συνοψίζοντας, θεωρούμε ότι η διαφορά ανάμεσα στην κανονική αντίληψη και τη φαντασία δεν έγκειται κυρίως στην διαφορά των ιδίων των εμφανιζόμενων αντικειμένων αλλά στον τρόπο που σχετιζόμαστε μαζί τους. Για παράδειγμα, όταν αντιλαμβάνομαι ενεργεία ένα λιοντάρι απέναντί μου, υιοθετώ μια στάση, μια συμπεριφορά και, γενικότερα, ένα σύνολο δεσμεύσεων τελείως διαφορετικό απ' ό,τι όταν αντιλαμβάνομαι φανταστικά ένα λιοντάρι. Στη δεύτερη περίπτωση δεν θα φοβηθώ ούτε θα το βάλω στα πόδια. Γενικότερα, θα έχω ένα πολύ διαφορετικό σύνολο αναμονών για το τι πρόκειται να συμβεί. Αυτό γίνεται φανερό και από την περίπτωση της παραίσθησης. Αν ισχύει η εξήγηση που δώσαμε ότι η παραίσθηση είναι μια φαντασία από την οποία έχει εκπέσει η συναίσθηση της ικανότητας του πράττειν, τότε γίνεται φανερό πόσο δραστικά

ένα δεύτερο εμμενές δένδρο, ή ακόμα μια 'εσωτερική εικόνα' του πραγματικού δένδρου που στέκεται εκεί έξω μπροστά μου, με κανέναν τρόπο δεν δίνεται, και το να υποθέσει κανείς κάτι τέτοιο υπό τη μορφή παραδοχής οδηγεί μόνο σε παραδοξολογίες" (Husserl 1962, σελ. 243). Βλ., επίσης, McDowell 1994a, σελ. 204.

⁴⁷⁸ Εξίσου παραπλανητικό είναι επίσης να μιλάμε για άμεση αντίληψη αν δεν υπάρχουν νοητικές αναπαραστάσεις στο φαινομενολογικό επίπεδο. Πρβλ. Austin 1962.

⁴⁷⁹ Βλέπε, Kant 1979, ειδικότερα τα "Αιτήματα της Εμπειρικής Νοήσεως εν γένει".

μεταβάλλει τον τρόπο που σχετιζόμαστε με το ίδιο περιεχόμενο μια τροποποίηση της συναίσθησης της ικανότητας του πράττειν. Όταν φαντάζομαι ένα λιοντάρι και όταν έχω την παραίτηση ενός λιονταριού σίγουρα αναλαμβάνω ένα πολύ διαφορετικό σύνολο δεσμεύσεων και προδιαθέσεων για συμπεριφορά.

Είναι οι δεσμεύσεις και οι προδιαθέσεις για πράξη που δημιουργούνται κατά την εκάστοτε αντίληψη εκείνες που καθορίζουν αν αντιλαμβάνομαι κάτι ενεργεία ή φανταστικά. Και επειδή οι δεσμεύσεις και οι προδιαθέσεις για πράξη προσδιορίζουν τη συναίσθηση της ικανότητας του πράττειν, μπορούμε να ισχυριστούμε ότι η διαφοροποίηση στη φαινομενολογία της κανονικής αντίληψης από τη φαινομενολογία της φαντασίας μπορεί να εντοπιστεί, σε ένα βαθμό, στην αντίστοιχη διαφοροποίηση της συναίσθησης της ικανότητας του πράττειν.

Πάντως, παρά την τεράστια διαφορά που παρουσιάζει η φαντασία από την κανονική αντίληψη ως προς αυτή τη διάσταση, και στις δύο περιπτώσεις αντιλαμβάνομαι ένα ον έτερο προς εμένα, ένα ον που ανήκει στον κόσμο και όχι ένα εσωτερικό ον. Η διαφορά της φαντασίας από την κανονική αντίληψη δεν πρέπει να αναζητηθεί, λοιπόν, σε κάποια εσωτερικότητα που υποτίθεται ότι χαρακτηρίζει την πρώτη, αλλά στις δεσμεύσεις και τους τρόπους πράξης που συναρμολόγουν με αυτές τις δύο ικανότητες και στις αντίστοιχες διαφοροποιήσεις της συναίσθησης της ικανότητας του πράττειν.

8.2.2.3 Τα νοητικά φαινόμενα ως τρόποι φανέρωσης του κόσμου στον άνθρωπο

Με τα παραπάνω δείξαμε τη γενικότερη θέση ότι άμεση πρόσβαση στον κόσμο δεν έχουμε μόνο κατά την κανονική αντίληψη αλλά και κατά την ψευδαίσθηση, την παραίτηση και τη φαντασία. Πιο συγκεκριμένα, έγινε φανερό ότι με τα νοητικά φαινόμενα ερχόμαστε σε επαφή με τα ίδια τα πράγματα και όχι με νοητικές αναπαραστάσεις τους. Είδαμε, επίσης, ότι τα πράγματα με τα οποία ερχόμαστε σε επαφή δεν δίδονται ως εσωτερικά αντικείμενα αλλά ως πράγματα του κόσμου. Αφού, όμως, τα πράγματα δεν μπορούν παρά να δίδονται ως πράγματα του κόσμου, τα νοητικά φαινόμενα δεν πρέπει να νοούνται ως ένας εσωτερικός αναδιπλασιασμός του κόσμου και των πραγμάτων του, αλλά ως μια δάνοιξη του ανθρώπου στον κόσμο. Τα νοητικά φαινόμενα δεν είναι κάποιος "μεσάζων" μεταξύ του ανθρώπου και του κόσμου, αλλά είναι οι τρόποι φανέρωσης του κόσμου στον άνθρωπο. Ο άνθρωπος, για να χρησιμοποιήσουμε χαιντεγκεριανή ορολογία, είναι πάντα ήδη "έξω", παράπλευρα στα όντα.⁴⁸⁰

⁴⁸⁰ Βλέπε, Heidegger (1978, σελ. 114): "Όταν το Dasein κατευθύνεται προς κάτι και το συλλαμβάνει, δεν βγαίνει πρώτα έξω από την εσωτερική του σφαίρα, όπου τάχα είναι αρχικά κλεισμένο, παρά σύμφωνα με το πρωταρχικό είδος του Είναι

Φυσικά, ο κόσμος που μας φανερώνεται με την αντίληψη δεν είναι, με καντιανούς όρους, ο κόσμος καθεαυτόν. Αυτή η αντίληψη περί πραγματικότητας προϋποτίθεται από την αναπαραστασιακή θεωρία περί αντίληψης υπό την έννοια ότι την καθιστά δυνατή, γιατί προσδιορίζει το όριο που δεν μπορεί να φτάσει καμιά αναπαράσταση. Όταν, λοιπόν, απορρίπτουμε την αναπαραστασιακή θεωρία περί αντίληψης οφείλουμε εξίσου να απορρίψουμε και την ιδέα ενός κόσμου καθεαυτόν⁴⁸¹. Ο κόσμος που μας δίδεται άμεσα στην αντίληψη είναι ο ανθρώπινος κόσμος, δηλαδή ο κόσμος που μπορεί να δοθεί στις εκάστοτε ικανότητες του ανθρώπου, χωρίς αυτό να σημαίνει ότι πρόκειται για έναν υποκειμενικό ή έναν κατασκευασμένο κόσμο. Ο ανθρώπινος κόσμος συγκροτείται μέσα από ένα σύνολο πρακτικών που, όπως είδαμε στο τέταρτο κεφάλαιο (ενότητα 4.3), στηρίζονται στη πρακτική γνώση δεξιοτήτων, δηλαδή μια γνώση που δεν είναι αναπαραστασιακή αλλά ενσώματη (embodied).

Εφόσον κατά την αντίληψη -με τη γενικότερη έννοια του όρου- δεν μας αποκαλύπτονται ούτε νοητικές αναπαραστάσεις ούτε εσωτερικά αντικείμενα αλλά ο ίδιος ο κόσμος (με την προαναφερθείσα έννοια του όρου), η οπτική της αντίληψης δεν διαφέρει, ως προς αυτή τη διάσταση, από την καθημερινή οπτική του τρίτου προσώπου. Η αντίληψη, λοιπόν, δεν αποτελεί παράδειγμα οπτικής του πρώτου προσώπου, αν με την τελευταία εννοούμε πρόσβαση σε εσωτερικές οντότητες που μας αποκαλύπτονται στο φαινομενολογικό επίπεδο, γιατί σε αυτό το επίπεδο δεν εμφανίζονται εσωτερικές οντότητες. Από την άλλη μεριά ωστόσο, η αντίληψη δεν πρέπει να θεωρηθεί, ως προς όλες τις διαστάσεις της, παράδειγμα καθημερινής οπτικής του τρίτου προσώπου, γιατί κάθε αντίληψη χαρακτηρίζεται, μεταξύ άλλων, από μια συναίσθηση της ικανότητας του πράττειν. Όμως η καθημερινή οπτική του τρίτου προσώπου, εξ' ορισμού, δεν μπορεί να συμπεριλάβει αυτό το ανεξάλειπτο ίχνος του υποκειμενικού "πόλου" των νοητικών φαινομένων.

8.3 Δεύτερη κοινή παραδοχή: ταυτότητα μεταξύ νοητικών και λειτουργικών/νευροφυσιολογικών φαινομένων

Ας έρθουμε τώρα σε μια δεύτερη κοινή παραδοχή που υπάρχει μεταξύ της ιντερναλιστικής και της εξτερναλιστικής εκδοχής της ΑΥΘΝ. Η αμφισβήτηση αυτής της παραδοχής δεν έχει πλέον ως στόχο την υπονόμευση της έννοιας της νοητικής αναπαράστασης, αλλά των λόγων για τους οποίους πολλοί θεωρούν ως αυτονόητη και σχεδόν αναγκαστική την ύπαρξη κάποιου είδους νοητικών αναπαραστάσεων.

του είναι πάντα 'έξω', παράπλευρα σε συναντώμενα όντα του εκάστοτε ήδη αποκαλυπτόμενου κόσμου".

⁴⁸¹ Βλ. Putnam 1994, 1994a και την έννοια του αφελούς (naïve) ή φυσικού (natural) ρεαλισμού.

Η δεύτερη κοινή παραδοχή των δύο εκδοχών της ΑΥΘΝ είναι η υποστήριξη κάποιου είδους ταυτότητας μεταξύ των νοητικών φαινομένων και των λειτουργικών ή των νευροφυσιολογικών φαινομένων. Στη σύγχρονη φιλοσοφία του νου έχει επικρατήσει, κυρίως, η αρχή της ταυτότητας εκτύπου, καθώς ασθενέστερη από την αρχή της ταυτότητας τύπου (βλ. κεφάλαιο 1, ενότητα 1.4.2). Σύμφωνα με την αρχή της ταυτότητας εκτύπου, τα έκτυπα νοητικών φαινομένων ταυτίζονται με έκτυπα φυσικών φαινομένων (συνήθως, εγκεφαλικών φαινομένων). Για παράδειγμα, κάθε φορά που κάποιος βρίσκεται σε μια κατάσταση πεποιθήσης, αυτή η κατάσταση δεν είναι τίποτα άλλο από μια φυσική κατάσταση (συνήθως κάποια εγκεφαλική κατάσταση). Η εκάστοτε πεποιθήση, νοημένη ως έκτυπο, ταυτίζεται με κάποιο έκτυπο φυσικής κατάστασης. Όμως, αν εξετάσουμε δύο έκτυπα νοητικών καταστάσεων που υπάγονται στον ίδιο τύπο (π.χ. τον τύπο επιθυμία), τότε τα αντιστοιχα δύο έκτυπα φυσικών καταστάσεων, με τα οποία ταυτίζονται αυτά τα έκτυπα νοητικών καταστάσεων, δεν υπάγονται, κατ' ανάγκη, στον ίδιο τύπο φυσικών καταστάσεων. Ένας άλλος τρόπος να διατυπωθεί η ίδια θέση είναι ότι δεν υπάρχουν ψυχοφυσικοί νόμοι, δηλαδή νόμοι που συνδέουν τύπους νοητικών φαινομένων με τύπους φυσικών φαινομένων⁴⁸². Εφόσον, όμως, τα νοητικά φαινόμενα έχουν αιτιακή δράση δεν μπορούν παρά να είναι φυσικά φαινόμενα και ως τέτοια, πρέπει αναγκαστικά να υπάγονται σε κάποιον φυσικό νόμο που περιγράφει τη συγκεκριμένη αιτιακή δράση⁴⁸³.

Η αρχή της ταυτότητας εκτύπου επικράτησε στη σύγχρονη φιλοσοφία του νου γιατί εξυπηρετούσε, επιπλέον, μια φαινομενικά ασύμβατη παραδοχή, ότι, δηλαδή, υπάρχει αυτονομία μεταξύ των επιστημών. Η ταυτότητα εκτύπου, αντίθετα από την ταυτότητα τύπου, συνεπάγεται ότι οι οντότητες της ψυχολογίας δεν είναι δυνατόν να αναχθούν στις οντότητες της φυσικής (ή έστω της νευροφυσιολογίας). Η αρχή της ταυτότητας εκτύπου θεωρήθηκε το έσχατο σημείο υποχώρησης προς την κατεύθυνση αυτονόμησης της ψυχολογίας από τη νευροφυσιολογία, χωρίς να χάνεται ταυτόχρονα ο φυσιοκρατικός χαρακτήρας της ψυχολογίας. "Φυσιοκρατικός χαρακτήρας", εν προκειμένω, σημαίνει ότι οι οντότητες στις οποίες αναφέρεται η ψυχολογία έχουν αιτιακή δράση. Τα νοητικά φαινόμενα, οι πεποιθήσεις, οι

⁴⁸² Βλ. Davidson (1980) και τη θεωρία του περί ανώμαλου μονισμού, η οποία αποτελεί ίσως την καλύτερη διερεύνηση της αρχής της ταυτότητας εκτύπου.

⁴⁸³ Πρόκειται για αυτό που ο Davidson (1980) ονομάζει "αρχή του νομολογικού χαρακτήρα της αιτιότητας".

επιθυμίες, τα αισθήματα προκαλούνται, μεταξύ άλλων, από ερεθίσματα που περιβάλλοντος και προκαλούν συμπεριφορά⁴⁸⁴.

Η αποδοχή των θεωριών ταυτότητας υπαγορεύεται από έναν έντονο αντιδυσισμό που διαπνέει τη σύγχρονη φιλοσοφία του νου: τα νοητικά φαινόμενα δεν μπορεί παρά να είναι φυσικά φαινόμενα. Το πρόβλημα με αυτή την αντίληψη ξεκινά από το τι ακριβώς λογίζεται κάθε φορά ως φυσικό. Συγκεκριμένα, παρατηρείται μια περιστολή του φυσικού σε ό,τι μπορεί να μελετηθεί από την επιστήμη της φυσικής. Η οντολογία του φυσικού (natural) περιστέλλεται, δηλαδή, σε μια φυσικαλιστική (physicalistic) οντολογία. Ακόμα και η θεωρία ταυτότητας εκτύπου, που έχει ρητά αναπτυχθεί κατά του φυσικαλιστικού αναγωγισμού, δεν κατορθώνει τελικά να αποφύγει την προαναφερθείσα περιστολή του φυσικού. Ας δούμε, πιο συγκεκριμένα γιατί συμβαίνει αυτό. Αν τα έκτυπα όντων που χρησιμοποιεί η ψυχολογία και η νευροφυσιολογία ταυτίζονται, τότε δεν μπορούν παρά να είναι, κάθε φορά, της ίδιας οντολογικής κατηγορίας. Ο Davidson (1980), για παράδειγμα, μιλά για νοητικά συμβάντα που ταυτίζονται με φυσικά συμβάντα. Επίσης, όταν γίνεται λόγος για νοητικές διαδικασίες αυτές θεωρείται ότι ταυτίζονται με φυσικές διαδικασίες. Τέλος, τις περισσότερες φορές, τα νοητικά φαινόμενα κατανοούνται ως νοητικές καταστάσεις και, σε αυτή την περίπτωση, θεωρείται ότι ταυτίζονται με φυσικές καταστάσεις. Το κοινό χαρακτηριστικό σε όλες αυτές τις ταυτίσεις είναι ότι η οντολογική κατηγορία στην οποία ανήκει το υπό εξέταση έκτυπο νοητικού φαινομένου οφείλει να περιλαμβάνεται στο οντολογικό ρεπερτόριο της φυσικαλιστικής οντολογίας, διότι διαφορετικά δεν θα ήταν δυνατόν το εκάστοτε νοητικό φαινόμενο να ταυτίζεται με κάποιο φυσικό φαινόμενο.

Η αρχή της ταυτότητας εκτύπου, να μην απορρίπτεται το ότι οι διάφορες επιστήμες (εν προκειμένω η νευροφυσιολογία και η ψυχολογία) αναφέρονται στα ίδια είδη ή τύπους όντων, αλλά εξακολουθεί να είναι αρκετά περιοριστική γιατί επιβάλλει ότι οι επιστήμες μιλούν για τα ίδια έκτυπα, τα ίδια άτομα (individuals)⁴⁸⁵ τα οποία δεν μπορεί να είναι άλλα από τα έκτυπα της φυσικής. Πόσο αυτονόητη, όμως, είναι αυτή η αρχή; Αν, προς στιγμήν, δεχτούμε ότι τα έκτυπα της ψυχολογίας είναι, μεταξύ άλλων, οι πεποιθήσεις, γιατί πρέπει να δεχτούμε ότι κάποια από τα έκτυπα που μελετά η νευροφυσιολογία (πόσο μάλλον η φυσική) ταυτίζονται με έκτυπα πεποιθήσεων; Τα έκτυπα της νευροφυσιολογίας είναι, κυρίως, οι νευρώνες και οι διεγέρσεις τους. Κανένας σοβαρός νευροφυσιολόγος δεν ελπίζει ότι μια πεποίθηση μπορεί να ταυτιστεί με τη διέγερση κάποιου μεμονωμένου νευρώνα. Αυτό σημαίνει ότι δεν ισχύει η αρχή της

⁴⁸⁴ Όσοι υποστηρίζουν τον φυσικοκρατικό χαρακτήρα των νοητικών φαινομένων, δηλαδή, τη θέση ότι τα νοητικά φαινόμενα έχουν αιτιακή δράση, ονομάζονται και αποβλεπτικοί ρεαλιστές (intentional realists).

⁴⁸⁵ Βλ. Haugeland 1982/1998.

ταυτότητας εκτύπου: το έκτυπο πεποιθήσης δεν ταυτίζεται με το έκτυπο κάποιου νευροφυσιολογικού τύπου. Ωστόσο, αυτή η αντίρρηση φαίνεται να είναι αφελής: ένα έκτυπο πεποιθήσης μπορεί να μην είναι δυνατόν να ταυτιστεί με τη διέγερση κάποιου μεμονωμένου νευρώνα, αλλά τότε δεν αποκλείει να ταυτιστεί με τη διέγερση ενός συνόλου νευρώνων.

Είναι, όμως, όντως τόσο αφελής η συγκεκριμένη αντίρρηση; Θεωρούμε πως όχι, διότι δείχνει ότι δεν έχουμε ακριβώς να κάνουμε με σχέση ταυτότητας εκτύπων, αλλά με σχέση σύνθεσης (composition) του νοητικού από το φυσικό: το νοητικό έκτυπο συντίθεται από ένα σύνολο φυσικών εκτύπων⁴⁸⁶. Η σχέση σύνθεσης είναι μια σχέση μέρους - όλου: το νοητικό όλον συντίθεται από φυσικά μέρη. Αυτά τα μέρη μπορεί να είναι χωρικά ή χρονικά. Ένα πρόβλημα που δημιουργείται άμεσα με την αρχή της σύνθεσης είναι ότι δεν αποτελεί κάθε ένωση συμβάντων ένα νέο συμβάν, όπως δεν αποτελεί η ένωση δύο οποιωνδήποτε αντικειμένων ένα νέο αντικείμενο⁴⁸⁷. Τα συμβάντα, όπως και τα αντικείμενα, χαρακτηρίζονται από μια ενότητα και ακεραιότητα η οποία δεν μπορεί να επιτευχθεί από την απλή συμπαράθεση οποιονδήποτε μερών. Οι συνθήκες που πρέπει να πληρούνται προκειμένου η σύνδεση των μερών να συνιστά όντως ένα έκτυπο (ενιαίου) συμβάντος / κατάστασης / διαδικασίας / αντικείμενου αποτελεί ένα ανοιχτό πρόβλημα στην οντολογία⁴⁸⁸.

Όμως, ακόμα και αν παραβλέψουμε το γεγονός ότι η σχέση μεταξύ των εκτύπων της ψυχολογίας και της νευροφυσιολογίας (πολύ περισσότερο δε της φυσικής) δεν μπορεί να είναι σχέση ταυτότητας αλλά, στην καλύτερη περίπτωση, σχέση σύνθεσης, πάλι η θέση του Davidson αντιμετωπίζει ανυπέρβλητες δυσκολίες. Όπως έχουν επισημάνει αρκετοί ερευνητές, η έννοια της ταυτότητας στην οποία ανατρέχει εδώ ο Davidson δεν έχει σαφές νόημα. Για παράδειγμα, μια συνθήκη για την κατανόηση σχέσεων ταυτότητας είναι να γνωρίζουμε τι είναι τα δύο ταυτιζόμενα όντα. Όπως επισημαίνει η Stewart (1997, σελ. 237), η πρόταση "ο Έσπερος είναι ο Φώσφορος" γίνεται άμεσα κατανοητή διότι γνωρίζουμε ότι τόσο ο Έσπερος όσο και Φώσφορος είναι πλανήτες. Επειδή, ακριβώς, είναι πράγματα του ίδιου είδους, δηλαδή υπάγονται στην ίδια ταξινομητική έννοια (sortal concept), γίνεται κατανοητό το ότι ταυτίζονται. "Όμως στην περίπτωση των καταστάσεων τις οποίες ο υλοστηρικτής της ταυτότητας εκτύπου επαθμεί να δει ως ταυτιζόμενες, δεν υπάρχει κάποια ταξινομητική έννοια που θα μπορούσε να παίξει αυτόν τον μεσολαβητικό ρόλο. Δεν υπάρχει κάποιο επίπεδο περιγραφής πιο συγκεκριμένο από την κατηγοριακή έννοια 'κατάσταση' (ή 'έκτυπο κατάστασης')" (Stewart 1997, σελ. 237).

⁴⁸⁶ Βλ. Hornsby 1985.

⁴⁸⁷ Το ίδιο ισχύει και στην περίπτωση των καταστάσεων και των διαδικασιών.

⁴⁸⁸ Βλ. Simons 1987.

Ένα δεύτερο, σημαντικότερο, πρόβλημα που καθιστά ακαθόριστη την έννοια της ταυτότητας όπως χρησιμοποιείται από τον Davidson έχει να κάνει με το κριτήριο που προτείνει ο ίδιος για το πότε δύο συμβάντα ταυτίζονται. Ο Davidson (1980, σελ. 179) υποστηρίζει ότι δύο συμβάντα ταυτίζονται αν έχουν τις ίδιες αιτίες και τα ίδια αποτελέσματα. Ωστόσο, αυτό το κριτήριο, όπως έδειξε ο Quine (1985, σελ. 161-171), οδηγεί σε φαύλο κύκλο. Ο Putnam (1994a, σελ. 482-483) δίνει μια ιδιαίτερα παραστατική εξήγηση τού γιατί συμβαίνει κάτι τέτοιο:

"Για να δείτε ότι ο Quine έχει δίκιο, φανταστείτε ότι θέλουμε να αποφασίσουμε εάν η πυροδότηση μιας μικρής ομάδας νευρώνων -ενός από τους 'αυτόνομους υπομηχανισμούς' του Fodor - είναι ή όχι 'ταυτόσημη ως έκτυπο' με μια 'εμπειρία του μπλέ'. Η πυροδότηση μιας ομάδας νευρώνων θα έχει ένα πλήθος αποτελεσμάτων. Τα οποία συνήθως δεν θα θεωρούσαμε ή ονομάζαμε αποτελέσματα της εμπειρίας μου του μπλε' για παράδειγμα, την διέγερση άλλων νευρώνων. Εάν η εμπειρία του μπλε είναι ταυτόσημη με την πυροδότηση μιας ομάδας νευρώνων, τότε αυτές οι άλλες διεγέρσεις είναι 'αποτελέσματα της εμπειρίας του μπλε' εάν ωστόσο, η εμπειρία του μπλε είναι ταυτόσημη με την δραστηριότητα ενός μεγαλύτερου τμήματος του εγκεφάλου, συμπεριλαμβανομένων των άλλων διεγερόμενων νευρώνων, τότε αυτά τα άλλα συμβάντα διεγερσης θα αποτελούν *τμήματα* του συμβάντος που είναι 'η εμπειρία του μπλε' και όχι *αποτελέσματα* αυτού. Δεν υπάρχει τρόπος να αποφασίσει κανείς ποια ομάδα συμβάντων διεγερσης είναι ταυτόσημη με 'την εμπειρία του μπλε' εφαρμόζοντας το κριτήριο του Davidson. Έτσι, μένουμε με ένα είδος ταυτότητας που είναι *sui generis* και χωρίς κριτήρια."

Με άλλα λόγια, για να εξετάσουμε αν ένα έκτυπο ενός νευροφυσιολογικού συμβάντος έχει τις ίδιες αιτίες και τα ίδια αποτελέσματα με ένα έκτυπο ενός νοητικού συμβάντος πρέπει ήδη να γνωρίζουμε αν τα δύο αυτά συμβάντα ταυτίζονται.

Ο ίδιος ο Davidson αποδεχόμενος την κριτική του Quine δηλώνει: "[Ο Quine] λέει ότι το κριτήριο που προτείνει για την εξατομίκευση των συμβάντων είναι ριζικά μη ικανοποιητικό και συμφωνώ. Το αποδέχτηκα μόνο διστακτικά, αλλά τόνισα ότι θεωρούσα ότι ήταν περίπου τόσο καλό κριτήριο όσο εκείνο που έχουμε για τα φυσικά αντικείμενα. Ο Quine μου κατέστησε σαφέστερο το που έσφαλλα στην αρχική μου πρόταση και έτσι εδώ και τώρα την εγκαταλείπω" (Davidson 1985, σελ. 175)⁴⁸⁹.

⁴⁸⁹ Ο Davidson στην συνέχεια υποθετεί το κριτήριο που προτείνει ο Quine: "τα συμβάντα, όπως τα φυσικά αντικείμενα, ταυτίζονται εάν κατέχουν τις ίδιες θέσεις στους ίδιους χρόνους" (Davidson 1985, σελ. 175). Ωστόσο και αυτό το κριτήριο

Ένα πρόσθετο επιχείρημα κατά του κριτηρίου που πρότεινε ο Davidson δίδεται από τον Haugeland (1982/1998, σελ. 101-102). Ο Haugeland επινοεί μια περίπτωση όπου σε μακροσκοπικό επίπεδο υπάρχουν δύο διακριτά, αν και χωροχρονικά ταυτιζόμενα, συμβάντα, ενώ σε μικροσκοπικό επίπεδο υπάρχει ένα σύνολο βιωματικών συμβάντων, δηλαδή συμβάντων που είναι περιγράψιμα στη γλώσσα της (μαθηματικής) φυσικής, με τα οποία δεν ταυτίζεται κανένα από τα δύο μακροσκοπικά συμβάντα. Ας δούμε, όμως, πιο αναλυτικά το επιχείρημα.

Έστω ότι δύο κύματα ξεκινούν από τα δύο άκρα μιας πισίνας και ότι καθώς διαδίδονται κινούν πάνω-κάτω διάφορους φελλούς που βρίσκονται κατά μήκος της διαδρομής τους. Στο μακροσκοπικό επίπεδο, ακόμα και όταν τα δύο κύματα χτυπήσουν τον ίδιο φελλό την ίδια χρονική στιγμή, παραμένουν δύο εντελώς διακριτά –αν και χωροχρονικά ταυτιζόμενα– συμβάντα διότι έχουν διαφορετικές αιτίες και διαφορετικά αποτελέσματα. Συνεπώς, αν ισχυρε η ταυτότητα εκτύπου, τότε θα έπρεπε να υπάρχουν δύο διακριτά μαθηματικά συμβάντα με τα οποία θα ταυτιζόνταν τα δύο κύματα. Κάθε ένα από αυτά τα δύο διακριτά μαθηματικά συμβάντα θα αποτελούταν από ένα πάρα πολύ μεγάλο αριθμό στιγμιαίων θέσεων και ταχύτητων μορίων του νερού στον αντίστοιχο χώρο και χρόνο. Το ερώτημα που τίθεται είναι ποιες από αυτές τις θέσεις και ταχύτητες αποτελούν το ένα μαθηματικό συμβάν και ποιες το άλλο;

Όπως επισημαίνει ο Haugeland (1982/1998, σελ. 102): "Όταν στρέψουμε το μικροσκόπιό μας, ωστόσο, και κοπάζουμε τις θέσεις και τις ταχύτητες των μορίων του νερού, δεν υπάρχει κανένα ίχνος από κάποιο από τα δύο κύματα για να βρεθεί. Τα μεμονωμένα μόρια κινούνται με λιγγώδη ταχύτητα προς όλες τις κατευθύνσεις, συγκρουόμενα και περνώντας το ένα δίπλα στο άλλο με ταχύτητες πολύ μεγαλύτερες από εκείνες της ροής του νερού ή της διάδοσης του κύματος... Καμιά από τις μεμονωμένες θέσεις και ταχύτητες στο μοριακό επίπεδο δεν αποτελεί με καθορισμένο τρόπο συστατικό του ενός κύματος (ή του άλλου ή κανενός) – παρότι καθένα έχει μια απειροελάχιστη συνεισφορά στο συνολικό αποτέλεσμα. Με άλλα λόγια, δεν υπάρχει κανένα 'σύμπλεγμα' τέτοιων μικροσκοπικών μαθηματικών συμβάντων το οποίο κυριολεκτικά να ταυτίζεται με οποιοδήποτε συμπαγές κύμα". Εφόσον, λοιπόν, η ταυτότητα εκτύπου δεν ισχύει για τα κυματικά φαινόμενα θα ήταν εντελώς περιεργό να ισχυριστούμε ότι ισχύει για τα νοητικά και τα φυσικά φαινόμενα.

Τα παραπάνω επιχειρήματα κατά της ταυτότητας εκτύπου δείχνουν ότι, από μόνη της, αυτή η αρχή αντιμετωπίζει πολλά προβλήματα. Αν τώρα συνυπολογίσουμε τα πορίσματα που προέκυψαν από την εξέταση της προηγούμενης παραδοχής, η αρχή της ταυτότητας μεταξύ νοητικών

εκτύπων και εγκεφαλικών εκτύπων καθίσταται εντελώς ανυπόστατη. Όπως δείξαμε, κάθε νοητικό φαινόμενο ανήκει πάντα σε κάποιον. Ωστόσο, όσο και αν προσπαθήσει κανείς δεν μπορεί να ανιχνεύσει στα εγκεφαλικά φαινόμενα τον υποκειμενικό "πόλο" των νοητικών φαινομένων. Κατά συνέπεια, τα έκτυπα της ψυχολογίας δεν μπορούν να ταυτιστούν με (ή, έστω, να συντεθούν από) τα έκτυπα της νευροφυσιολογίας ή της φυσικής.

Το ερώτημα που τίθεται είναι αν απομένει παρόλα αυτά κάτι από τη θεωρία ταυτότητας που μπορεί να διασωθεί. Πιο συγκεκριμένα, το ερώτημα είναι αν μπορούμε, τουλάχιστον, να υποστηρίξουμε την ταυτότητα του αντικειμενικού "πόλου" των νοητικών φαινομένων με τα φυσικά-εγκεφαλικά φαινόμενα. Αυτή φαίνεται μια εύλογη θέση στην οποία μπορεί να υποχωρήσει η γνωστική επιστήμη. Μάλιστα, ο ίδιος ο τρόπος που εκφραζόμαστε μας ωθεί, συχνά, στο να διακρίνουμε τους δύο "πόλους". Για παράδειγμα, πολλές φορές αντί να λέμε "εγώ πιστεύω ότι ρ", λέμε "εγώ έχω την πεποίθηση ότι ρ". Αυτή η ονοματοποίηση των προτασιακών στάσεων οδηγεί άμεσα σε μια εξάλειψη του υποκειμενικού "πόλου". Έτσι, απομένει μόνο ο αντικειμενικός "πόλος" και τα νοητικά φαινόμενα κατανοούνται ως υποστάσεις ή ιδιότητες υποστάσεων που υποκαθιστούν κάποιο αντικείμενο ή ιδιότητα αντικειμένων, δηλαδή κατανοούνται ως νοητικές αναπαραστάσεις⁴⁹⁰. Αν γίνει αυτό το βήμα, τότε αρχίζει να έχει νόημα και το ερώτημα του ιντερναλισμού και του εξτερναλισμού για το πού βρίσκονται τα νοητικά φαινόμενα, αν, δηλαδή, βρίσκονται μέσα ή έξω από το κεφάλι. Η γνωστική επιστήμη ξεκινά όντως από αυτό το επίπεδο, θεματοποιεί αποκλειστικά τον αντικειμενικό "πόλο" των νοητικών φαινομένων και εξετάζει όλα τα γνωστικά φαινόμενα σε ένα υποπροσωπικό επίπεδο.

Είναι δικαιολογημένη αυτή η στρατηγική; Θεωρούμε πως όχι. Καταρχάς, όπως δείξαμε, ο υποκειμενικός και ο αντικειμενικός "πόλος" των νοητικών φαινομένων συνιστούν μια αδιάσπαστη ενότητα που μόνο για λόγους εννοιολογικής ανάλυσης έχει νόημα να διασπούμε. Γι' αυτό είναι λάθος να υποστασιοποιούμε τον αντικειμενικό "πόλο". Είδαμε, επίσης, ότι η συναίσθηση της πατρότητας των νοητικών φαινομένων δεν αποτελεί κάποιο απλό επιφαινόμενο, αλλά έχει συγκροτητικό ρόλο στην οργάνωση του νοητικού βίου. Άρα, λοιπόν, δεν είναι δυνατόν να εξαλείψουμε το συγκεκριμένο "πόλο". Κάθε φορά που αναφερόμαστε σε νοητικά φαινόμενα, αναφερόμαστε στο σύνολο του οργανισμού και όχι στον εγκέφαλο ή κάποιο τμήμα του. Με άλλα λόγια, τα νοητικά φαινόμενα πρέπει να κατανοηθούν ως κατηγορήματα του υποκειμένου στο σύνολό του. Μια πεποίθηση είναι πάντα μια πεποίθηση κάποιου. Δεν υπάρχουν πεποιθήσεις που δεν ανήκουν σε κάποιον και που, απλώς, τυγχάνει να

⁴⁹⁰ Συγκεκριμένα, συνιστούν αναπαραστάσεις σύμφωνα με το κριτήριο της υποκατάστασης το οποίο όμως έχουμε απορρίψει στο τρίτο κεφάλαιο, ενότητα 3.3.1.2.

παρευρίσκονται σε κάποιο κεφάλι, παρά μόνο σε παθολογικές περιπτώσεις. Οι προτασιακές στάσεις - το να πιστεύω, το να επιθυμώ, το να φοβάμαι - αποτελούν ικανότητες του ανθρώπου που, ως τέτοιες, μπορούν να υπάρχουν δυνάμει ή ενεργεία. Όταν ενεργοποιείται κάποια από αυτές τις ικανότητες, τότε ο κόσμος εμφανίζεται κατά ένα συγκεκριμένο τρόπο στον άνθρωπο.

Επίσης, αν λάβουμε υπόψη την κριτική που ασκήσαμε στο έβδομο κεφάλαιο, προκύπτει ότι αυτό που στο φαινομενολογικό επίπεδο ονομάζουμε "αντικειμενικό πόλο" των νοητικών φαινομένων ή απλούστερα "περιεχόμενο" των νοητικών φαινομένων, δεν είναι δυνατόν να εντοπιστεί στο λειτουργικό ή το νευροφυσιολογικό επίπεδο, διότι όλες οι προσπάθειες φυσικοποίησης του νοητικού περιεχομένου αποδείχθηκαν ανεπαρκείς.

Τέλος, όπως δείξαμε στην ενότητα 8.2.1, ακόμα και αυτό που, για λόγους εννοιολογικής ανάλυσης, ονομάζουμε "αντικειμενικό πόλο" των νοητικών φαινομένων δεν μπορεί να κατανοηθεί με όρους νοητικών αναπαραστάσεων. Συγκεκριμένα, είδαμε ότι ένα βασικό χαρακτηριστικό των αντιληπτικών φαινομένων -το ότι μας εμφανίζουν άμεσα τον κόσμο- όσο κι αν προσπαθήσουμε δεν μπορούμε να το εντοπίσουμε στις λειτουργικές ή τις νευροφυσιολογικές καταστάσεις του εγκεφάλου. Η απουσία, ακριβώς, αυτού του "χαρακτηριστικού" οδήγησε στην ριζική κριτική ότι η έννοια της νοητικής αναπαραστάσης στο λειτουργικό ή το νευροφυσιολογικό επίπεδο είναι ασυνάρτητη. Κατά συνέπεια, ο βασικός υποψήφιος με τον οποίον η ΑΥΘΝ επιχειρεί να ταυτίσει, ή τουλάχιστον να εξηγήσει, τα νοητικά φαινόμενα εκλείπει.

Με βάση τα παραπάνω επιχειρήματα μπορούμε να συμπεράνουμε ότι η ιδέα μιας ταύτισης των νοητικών φαινομένων με τα φυσικά φαινόμενα είτε σε επίπεδο τύπων είτε σε επίπεδο εκτύπων πρέπει να εγκαταλειφθεί. Οι θεωρίες ταυτότητας, όντας εγκλωβισμένες σε μια φυσικαλιστική οντολογία είτε στο επίπεδο των εκτύπων είτε και στο επίπεδο των τύπων, καθιστούν εκ προοιμίου αδύνατη κάθε απόπειρα κατανόησης των νοητικών φαινομένων όπως αυτά εμφανίζονται από την οπτική του πρώτου προσώπου.

Αν, όμως, εγκαταλείψουμε τις θεωρίες ταυτότητας, τι άλλο απομένει από το να ξαναγυρίσουμε στον δυϊσμό νου - σώματος; Αν τα νοητικά φαινόμενα δεν ταυτίζονται με εγκεφαλικά φαινόμενα, τότε πώς αποφεύγεται ο καρτεσιανός δυϊσμός;

Θεωρούμε ότι η απόρριψη των θεωριών ταυτότητας δεν συνεπάγεται αναγκαία την υιοθέτηση του δυϊσμού. Μπορούμε να αποδεχτούμε μια χαλαρότερη σχέση μεταξύ του φυσικού και του νοητικού από τις σχέσεις ταυτότητας, χωρίς να οδηγηθούμε στο άκρο του δυϊσμού. Αυτό το είδος σχέσης διαφάνηκε στο παράδειγμα του Haugeland που εξετάσαμε παραπάνω. Εκεί είδαμε ότι παρότι ένα κύμα δεν μπορεί να ταυτιστεί με το σύνολο των στιγμιαίων θέσεων και ταχυτήτων των μορίων του νερού (που κατέχουν την ίδια περιοχή με το κύμα), η ύπαρξη και η συμπεριφορά του κύματος δεν είναι ανεξάρτητες από την ύπαρξη και τη συμπεριφορά των

μορίων του νερού. Κάθε μόριο νερού έχει μια απειροελάχιστη συνεισφορά στο κύμα. Αν δεν συμπεριφέρονταν κατ' αυτόν τον τρόπο τα μόρια του νερού τότε σε μακροσκοπικό επίπεδο δεν θα παρουσιάζονταν κύματα με τις συγκεκριμένες ιδιότητες. Υπ' αυτή την έννοια, η ύπαρξη και η συγκεκριμένη συμπεριφορά των μορίων του νερού αποτελούν αναγκαία προϋπόθεση για την ύπαρξη κυμάτων.

Καθ' όμοιο τρόπο μπορούμε να υποστηρίξουμε ότι η ύπαρξη και η συγκεκριμένη λειτουργία του εγκεφάλου και γενικότερα του σώματος μαζί με την ύπαρξη του περιβάλλοντος εντός του οποίου ζει ο άνθρωπος αποτελούν αναγκαίες προϋποθέσεις για την ύπαρξη νοητικών φαινομένων. Ο Haugeland (1982/1998) ονόμασε αυτή την ειδική σχέση μεταξύ νοητικών και φυσικών φαινομένων "ασθενή επιγένεση"⁴⁹¹, ενώ τα τελευταία χρόνια έχει επικρατήσει να ονομάζεται "συνολική επιγένεση" (global supervenience)⁴⁹². Ο Haugeland (1982/1998, σελ. 96), παραφράζοντας τον Davidson, διατυπώνει την ασθενή επιγένεση ως εξής: "Δεν θα ήταν δυνατόν ο κόσμος να ήταν ακριβώς όπως είναι από όλες τις φυσικές απόψεις, αλλά διαφορετικός από κάποια νοητική άποψη" ή, ο κόσμος δεν μπορεί να αλλάξει από κάποια νοητική άποψη χωρίς να αλλάξει από κάποια φυσική άποψη"⁴⁹³.

Η ασθενής επιγένεση αποφεύγει κάθε αναφορά στη σχέση μεταξύ τύπων ή εκτύπων νοητικών και φυσικών φαινομένων. Το μόνο που λέει είναι ότι το φυσικό έχει μια προτεραιότητα επί του νοητικού. Λέγοντας, "φυσικό" εδώ δεν σημαίνει, κατ' ανάγκη, ότι περιοριζόμαστε στο τι συμβαίνει αποκλειστικά εντός του οργανισμού σε νευροφυσιολογικό επίπεδο. Η διατύπωση του Haugeland θα πρέπει να νοηθεί όσο το δυνατόν πιο γενικά και να συμπεριλάβει και τις φυσικές συνθήκες εκτός του οργανισμού.

Η αρχή της ασθενούς επιγένεσης απελευθερώνει πλήρως την οντολογία της ψυχολογίας από την οντολογία της φυσικής. Όπως, δηλώνει ο Haugeland (1982/1998, σελ. 102-103): "Το συμπέρασμά μου ... είναι ότι τα άτομα ή έκτυπα για τα οποία οι προτάσεις μας είναι αληθείς είναι τόσο 'σχετικά' προς το επίπεδο περιγραφής όσο είναι και τα είδη ή οι τύποι στους οποίους αυτές οι γλωσσικές προτάσεις τα ταξινομούν. Ο κόσμος δεν έρχεται μεταφυσικά εξατομικευμένος, όπως δεν έρχεται μεταφυσικά κατηγοριοποιημένος, πριν και ανεξάρτητα από κάποιους συγκεκριμένους περιγραφικούς τρόπους - το οποίο δεν σημαίνει ότι αυτά τα άτομα (και

⁴⁹¹ Είναι ασθενής κατά το ότι δεν προϋποθέτει ούτε την ταυτότητα τύπου ούτε την ταυτότητα εκτύπου.

⁴⁹² Βλ. Haugeland 1998, σελ. 4 και Wilson 1995, σελ. 145.

⁴⁹³ Πρβλ. υποσημείωση 332 όπου αναφέρεται ο ορισμός που δίνει ο Davidson.

κατηγορίες) δεν είναι απολύτως γνήσια και αντικειμενικά, αφ' ης στιγμής η γλώσσα καθορίζεται"⁴⁹⁴.

Ωστόσο, θα μπορούσε κανείς να αντιτείνει ότι η αρχή της ασθενούς επιγένεσης δεν εξηγεί το πώς ακριβώς σχετίζονται τα νοητικά φαινόμενα με τα φυσικά φαινόμενα και ότι αποτελεί έναν όρο αμυχανίας που απλώς ονομάζει παρά επλύει το πρόβλημα. Αυτό είναι ως ένα βαθμό αληθές. Πρέπει, όμως, να αναρωτηθούμε τι θα αποτελούσε μια *ικανοποιητική* εξήγηση της σχέσης νοητικού και φυσικού; Αν με τον όρο "*ικανοποιητική* εξήγηση" εννοούμε κάποιον ψυχοφυσικό νόμο τότε, σύμφωνα με όσα είπαμε σε αυτό το κεφάλαιο, κάτι τέτοιο δεν είναι εφικτό. Η ασθενής επιγένεση θα πρέπει να νοηθεί απλώς ως μια αρχή που συνθέτει την οντολογική προτεραιότητα του φυσικού έναντι του νοητικού. Ταυτόχρονα, αυτή η αρχή απελευθερώνει την οντολογία του νοητικού από την οντολογία της φυσικής και μας επιτρέπει να αναζητήσουμε έναν καταλληλότερο τρόπο προσδιορισμού των νοητικών φαινομένων⁴⁹⁵.

8.4 *Επιλογος*

Με το κεφάλαιο αυτό ολοκληρώσαμε την κριτική κατά της υπόθεσης περί ύπαρξης νοητικών αναπαραστάσεων. Καταρχάς δείξαμε ότι ένα ουσιαστικό χαρακτηριστικό κάθε νοητικού φαινομένου -η συνείδηση της κανονότητας του πράττειν- δεν είναι αναπαραστάσιμο. Κατά συνέπεια, ακόμα κι αν δεχόμαστε, προσωρινά, ότι τα "συστατικά" των νοητικών φαινομένων που αφορούν τον αντικειμενικό τους "πόλο" είναι αναπαραστάσιμα, σίγουρα η έννοια της νοητικής αναπαράστασης δεν θα επαρκούσε για να συλλάβει τα νοητικά φαινόμενα στην ολότητά τους. Η δεύτερη ριζικότερη κριτική που προέκυψε είναι ότι η έννοια της νοητικής αναπαράστασης είναι ασυνάρτητη. Ειδικότερα, δείξαμε ότι, με δεδομένο ότι η έννοια της αναπαράστασης προϋποθέτει α) την έννοια του χρήστη της αναπαράστασης (βλ. κεφάλαιο 7) και β) τη δυνατότητα άμεσης πρόσβασης του χρήστη είτε στον αναπαραστασιακό φορέα είτε στο αναπαριστώμενο αντικείμενο, οι εσωτερικές καταστάσεις του εγκεφάλου δεν είναι δυνατόν να συνιστούν αναπαραστάσεις.

Παράλληλα με την κριτική αρχίσαμε να αναπτύσσουμε μια μη αναπαραστασιακή περιγραφή των νοητικών φαινομένων. Αναλυτικότερα δείξαμε ότι αν η οπτική του πρώτου προσώπου κατανοηθεί καρτεσιανά ως ένα είδος ενδοσκοπικής, τότε, όπως και στην περίπτωση της επιστημονικής οπτικής του τρίτου προσώπου, διαφεύγουν δύο χαρακτηριστικά γνωρίσματα των νοητικών φαινομένων: η συνείδηση της πατρότητας και η άμεση πρόσβαση στον κόσμο. Όσον αφορά τη

⁴⁹⁴ Βλέπε, επίσης, Stoutland (1985, σελ. 59) για ένα παρόμοιο συμπέρασμα: "τα άτομα της ψυχολογίας είναι εξίσου μη αναγώγιμα στα άτομα της φυσικής, όσο μη αναγώγιμα είναι τα είδη της ψυχολογίας στα είδη της φυσικής".

⁴⁹⁵ Πρβλ. Putnam 1994, σελ. 290.

συναίσθηση της πατρότητας δείξαμε ότι δεν "βρίσκεται" κάπου στο εσωτερικό της υποκειμενικής σφαίρας για να εντοπιστεί με κάποιου είδους εσωτερική αντίληψη. Η συναίσθηση της πατρότητας δεν πρέπει να περιγραφεί καρτεσιανά, ως μια συναίσθηση της εσωτερικότητας, αλλά ως μια συναίσθηση της ικανότητας του πράττειν. Αυτή η θέση ενισχύεται από το γεγονός ότι στην περίπτωση ορισμένων παραισθήσεων ένα υποκειμενο μπορεί να αισθάνεται ότι έχει σκέψεις "μέσα" στο νού του χωρίς ταυτόχρονα να τις αντιλαμβάνεται ως δικές του. Ο υποκειμενικός και ο αντικειμενικός "πόλος" των νοητικών φαινομένων αποτελούν μια πρωταρχική ενότητα: Η ίδια η επιτέλεση ενός νοητικού ενεργήματος επιφέρει και τη συναίσθηση της πατρότητάς του. Το ενέργημα και αυτός που το επιτελεί δεν είναι δύο διακριτά φαινόμενα που συνδέονται εξωτερικά, αλλά αποτελούν ένα ενιαίο φαινόμενο – μια οργανική ενότητα. Επίσης, δείξαμε ότι η συναίσθηση της ικανότητας του πράττειν δεν είναι απλό επικαιρόνιο γιατί η απουσία της από ένα νοητικό φαινόμενο (όπως, για παράδειγμα, συμβαίνει στην περίπτωση ορισμένων παραισθήσεων) προκαλεί ριζικές αλλαγές στο νοητικό βίο και την ανθρώπινη συμπεριφορά. Εκτός από την αμεσότητα με την οποία έχουμε επίγνωση ότι ο κόσμος εμφανίζεται σε εμάς, τα νοητικά φαινόμενα χαρακτηρίζονται και από μια δεύτερη αμεσότητα - διαφορετικού τύπου - που αφορά τον αντικειμενικό τους "πόλο": με τα νοητικά φαινόμενα μάς εμφανίζεται άμεσα ο κόσμος, χωρίς διάμεσους. Δείξαμε ότι στο φαινομενολογικό επίπεδο δεν εμφανίζονται νοητικές αναπαραστάσεις ούτε εσωτερικά αντικείμενα, αλλά ο ίδιος ο κόσμος. Κατά συνέπεια, εφόσον στο φαινομενολογικό επίπεδο ούτε ο αντικειμενικός ούτε ο υποκειμενικός "πόλος" των νοητικών φαινομένων δίδονται υπό μορφή αναπαραστάσεων ή εσωτερικών αντικειμένων, η πρόσβαση από την οπτική του πρώτου προσώπου δεν πρέπει να κατανοηθεί καρτεσιανά ως κάποιου είδους ενδοσκοπηση: τα νοητικά φαινόμενα δεν βρίσκονται κάπου μέσα στο υποκείμενο αλλά αποτελούν τους τρόπους εμφάνισης του κόσμου στον άνθρωπο. Τέλος, δείξαμε ότι σημαντικό ρόλο στη διαφοροποίηση μεταξύ ειδών νοητικών φαινομένων παίζει η αντίστοιχη τροποποίηση της συναίσθησης της ικανότητας του πράττειν: κάθε είδος νοητικού φαινομένου (αντίληψη, ανάμνηση, φαντασία κ.λπ.) χαρακτηρίζεται από ένα ιδιαίτερο σύνολο προδιαθέσεων για συμπεριφορά και δεσμεύσεων εκ μέρους του υποκειμένου. Γι' αυτό άλλωστε, ενώ κατά την παραίσθηση και τη φαντασία μπορεί να έχουμε πρόσβαση στο ίδιο περιεχόμενο, αυτό συνοδεύεται στις δύο περιπτώσεις από ένα πολύ διαφορετικό σύνολο δεσμεύσεων και προδιαθέσεων για συμπεριφορά: όταν φανταζόμαστε ένα λιοντάρι, σε αντίθεση με την περίπτωση που έχουμε την παραίσθηση ενός λιονταριού, δεν το βάζουμε στα πόδια.

Τα νοητικά φαινόμενα, εφόσον ουσιαστικά ανήκουν πάντα σε κάποιον, δεν μπορούν να αναλυθούν σε υποπροσωπικό επίπεδο χωρίς να χάσουν την ταυτότητά τους. Αν παραμείνουμε εγκλωβισμένοι σε μια έννοια ταυτότητας (τύπου (type) ή εκτύπου (token)) μεταξύ νοητικών και εγκεφαλικών φαινομένων, τότε αναγκάζομαστε να υποστασιοποιήσουμε

τα νοητικά φαινόμενα και, εξαλείφοντας τον υποκειμενικό τους "πόλο", μιλούμε απλά για νοητικές καταστάσεις, όπως πεποιθήσεις, επιθυμίες κ.λπ. Έτσι, δημιουργείται το "ανόητο" ερώτημα "που βρίσκονται οι νοητικές καταστάσεις;" το οποίο καθοδηγεί την έρευνα στο υποπροσωπικό επίπεδο. Η έρευνα της γνωσιακής επιστήμης δεν πρέπει να θεωρείται ότι αφορά τα ίδια τα νοητικά φαινόμενα αλλά ορισμένες αναγκαίες προϋποθέσεις για την εμφάνισή τους. Η κριτική της δεύτερης παραδοχής, δηλαδή της αποδοχής κάποιου είδους ταυτότητας μεταξύ νοητικών καταστάσεων και νευροφυσιολογικών ή λειτουργικών καταστάσεων, μας οδήγησε στην υιοθέτηση μιας χαλαρότερης σχέσης μεταξύ φυσικών και νοητικών φαινομένων, τη συνολική επιγένεση (global supervenience).

ΕΠΙΛΟΓΟΣ

Με το προηγούμενο κεφάλαιο ολοκληρώσαμε την κριτική της έννοιας της νοητικής αναπαράστασης και, παράλληλα, δώσαμε μια διαφορετική, προφανώς όχι πλήρη, περιγραφή των νοητικών φαινομένων η οποία δεν βασίζεται σε αυτή την έννοια. Τα δύο κύρια γνωρίσματα των νοητικών φαινομένων που εντοπίσαμε ότι διαφεύγουν από την επιστημονική οπτική του τρίτου προσώπου (ήτοι, η συναίσθηση της ικανότητας του πράττειν και η άμεση πρόσβαση στον κόσμο) δείξαμε ότι διαφεύγουν εξίσου από την οπτική του πρώτου προσώπου αν αυτή κατανοηθεί καρτεσιανά ως κάποιο είδος εσωτερικής αντίληψης. Έτσι, ασκήσαμε κριτική όχι μόνο στην έννοια της νοητικής αναπαράστασης του λειτουργικού/νευροφυσιολογικού επιπέδου, αλλά και σε εκείνη του φαινομενολογικού επιπέδου. Η διτλή αμεσότητα που χαρακτηρίζει τα νοητικά φαινόμενα – άμεση πρόσβαση στον κόσμο και άμεση επίγνωση της πατρότητας των νοητικών φαινομένων – δεν μπορεί να αναλυθεί με όρους αναπαραστάσεων.

Τα παραπάνω επιτρέπουν έναν πρώτο επαναπροσδιορισμό του πεδίου έρευνας της γνωσιακής επιστήμης. Το ερώτημα που εγείρεται, εν προκειμένω, είναι πώς πρέπει να κατανοηθεί η έρευνα που διεξάγεται από τη γνωσιακή επιστήμη, αν δεχθούμε ότι δεν είναι δυνατόν να υπάρχουν νοητικές αναπαραστάσεις.

Στα κεφάλαια 3 έως 8 ασκήσαμε πολλαπλή κριτική στην έννοια της νοητικής αναπαράστασης και φτάσαμε στο συμπέρασμα ότι νοητικές αναπαραστάσεις δεν είναι δυνατόν να υπάρχουν. Από την άλλη μεριά, όπως είδαμε στο δεύτερο κεφάλαιο, η κλασική γνωσιακή επιστήμη βασίζεται καθοριστικά στην έννοια της νοητικής αναπαράστασης. Κάθε θεωρία που αναπτύσσεται στο πλαίσιο αυτής της επιστήμης προϋποθέτει την έννοια της νοητικής αναπαράστασης και της αποδίδει εξηγητικό ρόλο. Τίθεται, λοιπόν, το ερώτημα πώς αξιολογεί κανείς την έρευνα που διεξάγει η γνωσιακή επιστήμη αν δεχθεί την κριτική που προηγήθηκε;

Θα μπορούσαμε να υποστηρίξουμε ότι πρόκειται για μια επιστήμη που έχει οικοδομηθεί σε σαθρά θεμέλια. Ωστόσο, η γνωσιακή επιστήμη αποτελεί μολαταύτα μια ανθούσα επιστήμη, με μεγάλη διάδοση και πλήθος ερευνητικών προγραμμάτων. Πώς είναι δυνατόν να παρυσιάζει μια τόσο μεγάλη ανάπτυξη, αν η βασικότερη θεωρητική οντότητα που προϋποθέτουν οι θεωρίες της δεν είναι δυνατόν να υπάρχει;

Η θέση που θα υποστηρίξουμε είναι ότι η έρευνα που διεξάγεται στο χώρο της γνωσιακής επιστήμης είναι καθ' όλα χρήσιμη και πληροφοριακή υπό τον όρο ότι οι επονομαζόμενες "νοητικές αναπαραστάσεις" δεν κατανοούνται ως αναπαραστάσεις αλλά ως

εγγραφές⁴⁹⁶. Η έννοια της εγγραφής εισήχθη στο τρίτο κεφάλαιο (ενότητα 3.3.1.4) ως μια γενίκευση της έννοιας της προσομοίωσης. Ξεκινώντας από την έννοια του μοντέλου, επισημάναμε ότι η ομοιότητα μεταξύ ενός μοντέλου A και ενός μοντελοποιούμενου B δεν αποτελεί ούτε επαρκή ούτε αναγκαία συνθήκη για να έχει το A αναφορική σχέση προς το B, ήτοι για να καταστεί το A αναπαράσταση (συγκεκριμένα, απεικόνιση) του B⁴⁹⁷. Το ίδιο είδαμε ότι ισχύει και για τη σχέση αιτιότητας: το γεγονός ότι το A συνιστά αιτία του B δεν είναι ούτε επαρκής ούτε αναγκαία συνθήκη για να αποτελεί το B αναπαράσταση (συγκεκριμένα, ενδείκτη) του A⁴⁹⁸. Η ταινία που προκύπτει από τη μαγνητοσκόπηση ενός πλοίου δεν συνιστά αναπαράσταση του πλοίου, αλλά απλά μια εγγραφή του πλοίου. Το ίδιο συμβαίνει και με τις αναλογικές μετρήσεις: η ένδειξη ενός θερμομέτρου δεν αποτελεί από μόνη της μια αναπαράσταση της θερμοκρασίας αλλά απλά μια εγγραφή της θερμοκρασίας. Η εγγραφή από μόνη της, δηλαδή ανεξάρτητα από τον τρόπο που χρησιμοποιείται από τον χρήστη, δεν έχει αναφορική λειτουργία - δεν ενδεικνύει τη θερμοκρασία.

Επίσης, στο τρίτο κεφάλαιο είδαμε ότι ένας τρόπος να αναδειχθεί η διαφορά μεταξύ εγγραφών και αναπαραστάσεων είναι να συγκρίνουμε τη διαδικασία μετάβασης από ένα είδος εγγραφής σε ένα άλλο (μετεγγραφή) με τη διαδικασία μετάβασης από ένα είδος αναπαράστασης σε ένα άλλο (μετάφραση): Η μετεγγραφή, σε αντίθεση με τη μετάφραση, μπορεί να εκτελεστεί μηχανικά. Αυτό συμβαίνει γιατί η διαδικασία της μετεγγραφής στηρίζεται αποκλειστικά στη μορφή των εγγραφών, ενώ η μετάφραση στηρίζεται στο περιεχόμενο των αναπαραστάσεων. Ως χαρακτηριστικό παράδειγμα αναφέραμε το πόσο πιο δύσκολη είναι η περιγραφή με λόγια του περιεχομένου μιας φωτογραφίας από την ψηφιοποίηση μιας αναλογικής εγγραφής: η πρώτη σε αντίθεση με τη δεύτερη δεν είναι μηχανική διαδικασία.

Η διάκριση μετεγγραφής και μετάφρασης φανερώνει ότι δεν αρκεί να επικαλούμαστε, γενικώς και αορίστως, μια έννοια χρήσης για να δικαιολογήσουμε πότε ένα πράγμα λειτουργεί ως αναπαράσταση. Υπάρχουν χρήσεις των εγγραφών που αποτελούν απλώς μετεγγραφές, ενώ άλλες που καθιστούν τις εγγραφές αναπαραστάσεις⁴⁹⁹. Στο παραπάνω παράδειγμα, η ψηφιοποίηση μιας φωτογραφίας από ένα μηχάνημα είναι μετεγγραφή διότι διεξάγεται ερήμην του όποιου περιεχομένου έχει αυτή

⁴⁹⁶ Ο όρος "εγγραφή", έχει μια αμφισημία: αναφέρεται τόσο στη διαδικασία όσο και στο αποτέλεσμα. Εδώ χρησιμοποιούμε τον όρο "εγγραφή" με τη δεύτερη σημασία. Για την πρώτη σημασία θα χρησιμοποιούμε τους όρους "διαδικασία εγγραφής" ή "τρόπος εγγραφής".

⁴⁹⁷ Βλ., επίσης, το παράρτημα 1.

⁴⁹⁸ Αυτό το ζήτημα εξετάστηκε διεξοδικά στο έβδομο κεφάλαιο, ενότητα 7.2.4.1.

⁴⁹⁹ Στο τρίτο κεφάλαιο (ενότητα 3.3.1.2) εντοπίσαμε ως βασική συνθήκη της χρήσης αναπαραστάσεων την ικανότητα του χρήστη να συμπεριφέρεται έναντι του αναφερόμενου ως να είναι μη σωματικά παρόν.

στα μάτια κάποιου που την χρησιμοποιεί ως αναπαράσταση. Στο έβδομο κεφάλαιο (ενότητα 7.2.5.1) εξετάσαμε αυτή την ιδέα σε μια πιο γενική μορφή όταν ασχοληθήκαμε με την σημασιολογία λειτουργικού ρόλου. Εκεί είδαμε ότι οι μορφικές ή οι λειτουργικές σχέσεις μιας κατάστασης με άλλες καταστάσεις δεν αρκούν για να αποκτήσει αυτή η κατάσταση αναπαραστασιακό περιεχόμενο.

Συμπερασματικά, λοιπόν, μια εγγραφή δεν αποτελεί αναπαράσταση για δύο λόγους:

1) η σχέση αιτιότητας ή η σχέση ομοιότητας μεταξύ του εγγραφόμενου και της εγγραφής δεν είναι ούτε αναγκαία ούτε επαρκής συνθήκη για να καταστεί η εγγραφή *αναπαράσταση* του εγγραφόμενου, δηλαδή για να *αναφέρεται* η εγγραφή στο εγγραφόμενο, και

2) η επεξεργασία και ο μετασηματισμός της εγγραφής με βάση αποκλειστικά τη μορφή της δεν αποτελεί *αναπαραστασιακή* χρήση της εγγραφής (δηλαδή δεν καθιστά την εγγραφή αναπαράσταση) αλλά συνιστά απλώς μια μετεγγραφή της.

Από τη διερεύνηση που κάναμε στα προηγούμενα κεφάλαια προκύπτει ότι οι δύο παραπάνω λόγοι ισχύουν τόσο για τα νοητικά σύμβολα όσο και για τα νοητικά είδωλα. Κατά συνέπεια, τα δύο αυτά είδη των υποτιθέμενων αναπαραστάσεων είναι στη πραγματικότητα διαφορετικά είδη εγγραφών. Με δεδομένο αυτό το αποτέλεσμα, όλη η διερεύνηση που κάναμε στα κεφάλαια 4 και 5 όσον αφορά τα νοητικά σύμβολα και τα νοητικά είδωλα θα πρέπει να ξαναδιαβαστεί ως μια διερεύνηση διαφορετικών ειδών εγγραφών. Για παράδειγμα, η έννοια του κατοπτρισμού (*mirroring*) του αναπαραστασιακού περιεχομένου στην αναπαραστασιακή μορφή πρέπει να κατανοηθεί ως κατοπτρισμός του εγγραφόμενου στην εγγραφή. Ειδικότερα, το ίδιο ισχύει και για ό,τι στο πρώτο κεφάλαιο⁵⁰⁰ ονομάστηκε "κατοπτρισμός των σημασιολογικών σχέσεων στις συντακτικές σχέσεις". Εδώ πάλι, σύμφωνα με την ερμηνεία που προτείνουμε, πρόκειται για τον κατοπτρισμό σχέσεων μεταξύ των εγγραφόμενων σε σχέσεις μεταξύ των εγγραφών. Εκτός από την έννοια του κατοπτρισμού, άλλες έννοιες με συναφές περιεχόμενο είναι η αντιστοίχιση (*correspondence*), ο ισομορφισμός, η χαρτογράφηση (*mapping*) και η ιχνηλάτηση (*tracking*)⁵⁰¹. Όλες αυτές οι έννοιες θεωρούμε ότι προσδιορίζουν τρόπους εγγραφής. Επίσης, τρόπους εγγραφής προσδιορίζουν οι έννοιες της κωδικοποίησης (*coding*) και της αποκωδικοποίησης (*decoding*) στις οποίες αναφερθήκαμε στο έκτο κεφάλαιο (ενότητα 6.7.5) αναφορικά με τις υποτιθέμενες αναπαραστάσεις εισόδου και εξόδου στα νευρωνικά δίκτυα.

Όλες οι παραπάνω έννοιες, λοιπόν, δεν σχετίζονται με την έννοια της αναπαράστασης αλλά αποτελούν εξειδικεύσεις της έννοιας της

⁵⁰⁰ Βλ. ενότητα 1.3.4. Βλ., επίσης, ενότητες 2.4.2 και 7.1.

⁵⁰¹ Βλ. ενότητα 7.2.4.3.2.

εγγραφής. Οι εγγραφές, εφόσον δεν είναι αναπαραστάσεις δεν έχουν αναφορική λειτουργία και, κατά συνέπεια, δεν μπορούν να εξηγήσουν την αποβλεπτικότητα των νοητικών φαινομένων. Οι εγγραφές αποτελούν, στην καλύτερη περίπτωση, προσομοιώσεις διαφορετικών ειδών περιεχομένου. Κατά συνέπεια, η γνωσιακή επιστήμη, στο βαθμό που οι νοητικές αναπαραστάσεις για τις οποίες κάνει λόγο είναι στην πραγματικότητα εγγραφές, μελετά τα νοητικά φαινόμενα προσομοιώνοντας όψεις του περιεχομένου τους. Αν, μάλιστα δεχτούμε ότι ο εγκέφαλος επεξεργάζεται εγγραφές – μια εμπειρική υπόθεση για την οποία θα αποφανθεί η επιστήμη- τότε πρέπει να δεχτούμε ότι η γνωσιακή επιστήμη μελετά τους μηχανισμούς του εγκεφάλου, δηλαδή ορισμένες αναγκαίες συνθήκες για την εμφάνιση νοητικών φαινομένων. Εδώ κάνουμε λόγο απλώς για αναγκαίες συνθήκες βασιζόμενοι στην κριτική που ασκήσαμε στις έννοιες της ταυτότητας τύπου ή έκτυπου μεταξύ νοητικών και εγκεφαλικών φαινομένων: τα νοητικά φαινόμενα δεν είναι ταυτόσημα ούτε ως τύποι ούτε ως έκτυπα με τα εγκεφαλικά φαινόμενα, αλλά η ύπαρξη των εγκεφαλικών φαινομένων αποτελεί αναγκασία προϋπόθεση για την ύπαρξη των νοητικών φαινομένων.

Τα τελευταία χρόνια έχουν αρχίσει να αναπτύσσονται ορισμένες μη-αναπαραστασιακές προσεγγίσεις μέσα στο πλαίσιο της γνωσιακής επιστήμης. Πρόκειται για επιστημονικές προσεγγίσεις που δεν βασίζονται στην έννοια της νοητικής αναπαράστασης. Η κυριότερη από αυτές είναι η προσέγγιση των γνωσιακών διαδικασιών μέσω δυναμικών συστημάτων (Dynamic Systems Approach to Cognition)⁵⁰² την οποία παρουσιάζουμε εν συντομία στο δεύτερο παράρτημα της διατριβής. Επίσης, ορισμένοι ερευνητές θεωρούν ότι και τα νευρωνικά δίκτυα αποτελούν μια μη – αναπαραστασιακή προσέγγιση. Στο έκτο κεφάλαιο συσταχτήκαμε με αυτή τη θέση και προτείναμε συγκεκριμένα επιχειρήματα για την υποστήριξη της⁵⁰³. Με δεδομένη τη θέση μας κατά των νοητικών αναπαραστάσεων, το ερώτημα που εγείρεται σε σχέση με τις προαναφερθείσες μη-αναπαραστασιακές προσεγγίσεις είναι αν αυτές βρίσκονται σε πλεονεκτικότερη θέση όσον αφορά την εξήγηση των νοητικών φαινομένων. Θα ξεκινήσουμε με τα νευρωνικά δίκτυα.

Όπως είδαμε στο έκτο κεφάλαιο (ενότητα 6.7.4), τα νευρωνικά δίκτυα δεν φέρουν αναπαραστάσεις, όχι επειδή οι νοητικές αναπαραστάσεις που προϋποθέτει η γνωσιακή επιστήμη δεν μπορούν παρά να είναι απλώς εγγραφές, αλλά διότι το περιεχόμενο που κατοπτρίζεται στα νευρωνικά δίκτυα δεν είναι προτασιακό αλλά κάποιο είδος πρακτικής

⁵⁰² Βλ. van Gelder 1995 και van Gelder & Port (1995).

⁵⁰³ Άλλες μη αναπαραστασιακές προσεγγίσεις είναι η καταστασιακή ρομποτική (Situated Robotics) (Brooks 1991a, Brooks 1991b), η τεχνητή ζωή (Boden 1996), καθώς και μια καλύτερη προσέγγιση, η οικολογική ψυχολογία (Gibson 1986, Michaels & Carello 1981).

γνώσης δεξιοτήτων⁵⁰⁴. Δεδομένου δε ότι η πρακτική γνώση δεξιοτήτων, όπως δείξαμε στο τέταρτο κεφάλαιο, δεν έχει αναφορική διάσταση, η μορφή στην οποία κατοπτρίζεται το συγκεκριμένο περιεχόμενο δεν συνιστά αναπαράσταση για κάποιον επαρκή χρήστη. Εφόσον, όπως υποστηρίξαμε παραπάνω, η έννοια του κατοπτρισμού προσδιορίζει την έννοια της εγγραφής και όχι την έννοια της αναπαράστασης, θα πρέπει να δεχθούμε ότι ο τρόπος με τον οποίο κατοπτρίζεται το περιεχόμενο της πρακτικής γνώσης δεξιοτήτων στα νευρωνικά δίκτυα αποτελεί μια μορφή εγγραφής. Έτσι, η έννοια της ισχυρής υπέρθεσης που προτείναμε στο έκτο κεφάλαιο (ενότητα 6.7.2) προκειμένου να εξηγήσουμε τη συμπεριφορά των νευρωνικών δικτύων πρέπει να κατανοηθεί ως μια πολύ σύνθετη μορφή εγγραφής.

Ένα ερώτημα που δημιουργείται μετά από αυτή την τοποθέτηση είναι το εξής: επιτυγχάνουν τα νευρωνικά δίκτυα κάτι παραπάνω από την απλή εγγραφή / προσομοίωση της πρακτικής γνώσης δεξιοτήτων; Με άλλα λόγια, θα μπορούσαν τα νευρωνικά δίκτυα να θεωρηθούν ότι δεν προσομοιώνουν απλώς αλλά *πραγματώνουν* (realize) κάποια είδη δεξιοτήτων;

Ας εξετάσουμε λίγο πιο συγκεκριμένα το ερώτημα. Στο τέταρτο κεφάλαιο είδαμε ότι η πρακτική γνώση δεξιοτήτων είναι ένα είδος γνώσης που δεν μπορεί να διακριθεί από τον χρήστη της και να περιγραφεί μέσω αναπαραστάσεων που θα ίσχυαν ανεξάρτητα από αυτόν. Η πρακτική γνώση δεξιοτήτων είναι μια μορφή ενσώματης γνώσης. Το ερώτημα, λοιπόν, με τα νευρωνικά δίκτυα είναι αν υπολείπονται σε κάτι από τον τρόπο που ένας ειδικός ασκεί τη δεξιότητά του. Υπάρχει, με άλλα λόγια, κάτι που να εμποδίζει να δεχθούμε ότι αυτό που κάνουν τα νευρωνικά δίκτυα συνιστά *πραγμάτωση* δεξιοτήτων;

Θεωρούμε ότι η δραστηριότητα των νευρωνικών δικτύων δεν συνιστά πραγμάτωση δεξιοτήτων αλλά απλά εγγραφή δεξιοτήτων διότι η συγκεκριμένη δραστηριότητα δεν χαρακτηρίζεται από κάποια συναίσθηση της ικανότητας του πράττειν. Όπως είδαμε στο τέταρτο κεφάλαιο, η πρακτική γνώση δεξιοτήτων ανήκει πάντα σε κάποιον, δηλαδή χαρακτηρίζεται ουσιαστικά από μια συναίσθηση της ικανότητας του πράττειν. Γι' αυτό, άλλωστε, η άσκηση μιας δεξιότητας, παρά την πλήρη αυτοματοποίησή της, διακρίνεται πάντα από ένα εξαρτημένο αντανακλαστικό: Η άσκηση μιας δεξιότητας, σε αντίθεση με το εξαρτημένο αντανακλαστικό, κατανοείται από το υποκείμενο ως δική του πράξη. Χάρη στη συναίσθηση της ικανότητας του πράττειν ο κάτοχος της πρακτικής γνώσης δεξιοτήτων είναι σε θέση να διακρίνει αν η εκάστοτε δραστηριότητά του είναι επιδέξια ή όχι, δηλαδή είναι σε θέση να κάνει κανονιστικές διακρίσεις. Στα νευρωνικά δίκτυα δεν υπάρχει αυτή η διάσταση και γι' αυτό θεωρούμε ότι αποτελούν απλώς *προσομοιωτές* της

⁵⁰⁴ Βλέπε ενότητα 6.7.4.

πρακτικής γνώσης δεξιοτήτων. Εφόσον, δηλαδή τα νευρωνικά δίκτυα δεν μπορούν να κάνουν κανονιστικές διακρίσεις δεν φέρουν όντως κάποια μορφή γνωσιακού περιεχομένου αλλά απλά το προσομοιώνουν. Αν υπάρχει κάτι που τα νευρωνικά δίκτυα *πραγματώνουν* αυτό είναι, ενδεχομένως, τα εξαρτημένα αντανακλαστικά, δηλαδή μια μορφή δραστηριότητας απαλλαγμένη από κάθε είδος γνωσιακού περιεχομένου.

Ας έρθουμε τώρα στην προσέγγιση των γνωσιακών διαδικασιών μέσω δυναμικών συστημάτων. Η προσέγγιση αυτή έχει αναδυθεί εντελώς πρόσφατα και βρίσκεται ακόμα υπό διαμόρφωση. Γι' αυτόν κυρίως το λόγο την παρουσιάζουμε σε παράρτημα. Αφετηριακή παραδοχή της δυναμικιστικής προσέγγισης είναι ότι "[τ]α φυσικά γνωσιακά συστήματα αποτελούν συγκεκριμένα είδη δυναμικών συστημάτων και είναι καλύτερα να κατανοηθούν από την άποψη της δυναμικής" (van Gelder & Port 1995a, σελ. 5). Στο δεύτερο παράρτημα είδαμε, σε μια πρώτη προσέγγιση, τι είναι ένα δυναμικό σύστημα και σε τι διαφέρει από τα υπολογιστικά συστήματα. Υπάρχουν πολλές αντιρρήσεις για το αν και πώς μπορεί να γίνει αυστηρά αυτή η διάκριση. Δεν θα ασχοληθούμε περισσότερο με αυτό το ζήτημα, ωστόσο, γιατί δεν αφορά άμεσα το πλαίσιο της παρούσας διατριβής. Η δυναμικιστική προσέγγιση έχει δεχθεί κριτική και σε άλλα σημεία της⁵⁰⁵. Ένα είδος κριτικής, ίσως το πιο γενικό, είναι ότι η συγκεκριμένη προσέγγιση δίνει υπερβολική έμφαση στη ζεύξη του οργανισμού με το περιβάλλον. Η εν λόγω ζεύξη επιτυγχάνει, ως ένα βαθμό, να εξηγήσει την *προσαρμοστικότητα* του οργανισμού στο περιβάλλον, αλλά δεν μπορεί να εξηγήσει την (γνωσιακή) *αυτονομία* του.

Ακόμα, όμως, και αν παραμείνουμε σε γνωσιακά έργα που απαιτούν πολύ στενή αλληλεπίδραση του ανθρώπου με το άμεσο περιβάλλον του, όπως σωματικές δεξιότητες που βασίζονται σε άριστο συντονισμό αντίληψης και πράξης, πάλι δεν ισχύει ότι τα δυναμικά συστήματα επιτυγχάνουν κάτι παραπάνω από την προσομοίωση κάποιου είδους μη προτασιακού περιεχομένου. Όπως και στην περίπτωση των νευρωνικών δικτύων, τα δυναμικά συστήματα δεν μπορούν να εξηγήσουν τη συναισθηση της ικανότητας του πράττειν που χαρακτηρίζει την άσκηση μιας δεξιότητας. Αυτή η έλλειψη αναδεικνύει το γεγονός ότι η δυναμικιστική προσέγγιση τελικά, παρά τη ρητορεία της περί του αντιθέτου, δεν αφορά το προσωπικό επίπεδο. Οι δυναμικιστικές εξηγήσεις είναι υποπροσωπικές εξηγήσεις. Με άλλα λόγια, δεν αρκεί απλά η μεταβολή της μονάδας ανάλυσης από τον απομονωμένο εγκέφαλο στον ενσώματο και ενταγμένο στο περιβάλλον εγκέφαλο για να μεταβούμε στο προσωπικό επίπεδο ανάλυσης. Αυτή η μετάβαση επιτρέπει, ενδεχομένως, να συλλάβουμε τον τρόπο που ένας έμβιος οργανισμός λειτουργεί μέσα στο περιβάλλον του, αλλά σίγουρα δεν συλλαμβάνει το προσωπικό

⁵⁰⁵ Βλ. ενδεικτικά Grush 1997, Eliasmith 1997, Bechtel 1996.

επίπεδο στο οποίο ο κόσμος εμφανίζεται στον οργανισμό και ταυτόχρονα ο οργανισμός έχει μια συναίσθηση της ικανότητας του πράττειν.

Ο δυναμικισμός αντιδρώντας στον καρτεσιανισμό της γνωστικοκρατίας που συνοψίζεται στη θέση ότι ο νους είναι μέσα στο κεφάλι, φτάνει στο άλλο, εξίσου απαράδεκτο, άκρο ότι, δηλαδή, ο νους είναι (και) έξω από το κεφάλι. Αυτός ο ριζικός εξτερναλισμός παρότι εμφανίζεται να υπερβαίνει τον φυσικοποιημένο καρτεσιανισμό της γνωστικοκρατίας, δηλαδή τον τοπολογικό χωρισμό νου και κόσμου, εξακολουθεί να παραμένει εγκλωβισμένος σε μια καρτεσιανή εικόνα γιατί υποστασιοποιεί τα νοητικά φαινόμενα. Αποτέλεσμα της υποστασιοποίησης είναι να εξαφανίζεται ο υποκειμενικός "πόλος" των νοητικών φαινομένων και να δημιουργείται το ("ανόητο") ερώτημα "πού βρίσκονται τα νοητικά φαινόμενα;"⁵⁰⁶. Έτσι, νομιμοποιούνται ένα πλήθος από πιο εξειδικευμένα ερωτήματα τα οποία προβάλλουν επιτακτικά: "πού βρίσκονται οι πεποιθήσεις, οι επιθυμίες, οι αναμνήσεις κ.λπ."

Αν ακολουθήσουμε όσα υποστηρίζαμε στο όγδοο κεφάλαιο, τα νοητικά φαινόμενα δεν βρίσκονται ούτε "μέσα" ούτε "έξω" από το κεφάλι γιατί δεν είναι οντότητες για να βρίσκονται κάπου. Τα νοητικά φαινόμενα ως τρόποι φανέρωσης του κόσμου στον άνθρωπο αποτελούν τη συνθήκη δυνατότητας για την κατανόηση της διάκρισης μεταξύ εντός και εκτός. Γενικότερα, ο νους δεν είναι ένα απομονώσιμο όργανο αλλά ένα σύνολο ικανοτήτων (capacities) του ανθρώπου⁵⁰⁷ που, όταν ενεργοποιούνται, επιτρέπουν την φανέρωση του κόσμου με διαφορετικούς τρόπους. Κατά συνέπεια, μόνο στο προσωπικό επίπεδο ανάλυσης και περιγραφής μπορούν να προσεγγιστούν τα ίδια τα νοητικά φαινόμενα και όχι κάποιες συνθήκες που είναι απλώς αναγκαίες για την εμφάνιση των νοητικών φαινομένων.

Το προσωπικό επίπεδο ανάλυσης παραδοσιακά προσεγγίζεται από την καθημερινή οπτική του τρίτου προσώπου⁵⁰⁸ και αποσυνδέεται από το φαινομενολογικό επίπεδο. Για παράδειγμα, ο Wittgenstein επιχειρώντας να περιγράψει το τι σημαίνει "κατανοώ μια έννοια" επισημαίνει ότι δύο υποκείμενα ενδέχεται να έχουν τα ίδια νοητικά είδωλα αλλά να εννοούν κάτι διαφορετικό: "Και είναι ουσιαστικό να δούμε ότι στο νου μας μπορεί να παρουσιαστεί το ίδιο πράγμα όταν ακούμε τη λέξη, και όμως η εφαρμογή της να είναι διαφορετική. Έχει λοιπόν την ίδια σημασία και τις δύο φορές; Πιστεύω πως θα απαντήσουμε αρνητικά." (Wittgenstein

⁵⁰⁶ Βλ. όγδοο κεφάλαιο, ενότητα 8.3.

⁵⁰⁷ Βλ. Putnam 1994, σελ. 453, υποσημ. 18.

⁵⁰⁸ Υπενθυμίζουμε ότι στο όγδοο κεφάλαιο (ενότητα 8.2) διακρίναμε την καθημερινή οπτική του τρίτου προσώπου από την επιστημονική οπτική του τρίτου προσώπου: ενώ η πρώτη αποδίδει νοητικά φαινόμενα σε ολόκληρο το υποκείμενο, η δεύτερη τα εντοπίζει κάπου στο εσωτερικό του και ειδικότερα στον εγκέφαλο. Γι' αυτό οι εξηγήσεις που παρέχει η επιστημονική οπτική του τρίτου προσώπου δεν αφορούν το προσωπικό επίπεδο περιγραφής.

1952/1977, §140, 152). Επίσης, ο Putnam (1981a, σελ. 19-20) βασιζόμενος στην υπόδειξη αυτή του Wittgenstein επισημαίνει: "Είδαμε ότι η κατοχή μιας έννοιας δεν είναι ζήτημα κατοχής νοητικών ειδώλων (ας πούμε, δένδρων - ή ακόμα 'οπτικών' ή 'ακουστικών' νοητικών ειδώλων κάποιων γλωσσικών προτάσεων ή ολόκληρων κειμένων (discourses)) εφόσον κάποιος μπορεί να κατέχει όποιο σύστημα νοητικών ειδώλων θέλετε χωρίς να έχει την *ικανότητα* (ability) να χρησιμοποιεί γλωσσικές προτάσεις με τρόπο κατάλληλο ανάλογα με την εκάστοτε κατάσταση (θεωρώντας ότι η 'καταστασιακή καταλληλότητα' καθορίζεται τόσο από γλωσσικούς παράγοντες -τι έχει επωθεί πριν- όσο και από μη γλωσσικούς παράγοντες). Ένας άνθρωπος μπορεί να έχει όσα νοητικά είδηλα θέλετε και παρόλα αυτά να είναι εντελώς χαμένος όταν κάποιος του πει 'δείξε ένα δένδρο', ακόμα κι' αν υπάρχουν μπροστά του πολλά δένδρα... Ανεξάρτητα από τι είδους εσωτερικά φαινόμενα δεχθούμε ως δυνατές *εκφράσεις* της σκέψης, ... δεν είναι τα φαινόμενα τα ίδια αυτά που συνιστούν την κατανόηση, αλλά μάλλον η *ικανότητα* του στοχαστή να *εφαρμόζει* αυτά τα φαινόμενα, να παράγει τα σωστά φαινόμενα στις σωστές περιστάσεις". Το απόσπασμα αυτό του Putnam βρίσκεται, ως ένα βαθμό, σε συμφωνία με όσα υποστηρίξαμε στο τέταρτο κεφάλαιο περί του ακολουθείν *επιδέξια* έναν κανόνα. Εκεί δείξαμε ότι όταν ακολουθούμε *επιδέξια* έναν κανόνα δεν "έχουμε στο νου μας" κάποιο προτασιακό κανόνα ανεξάρτητο από το πλαίσιο χρήση τον οποίο εφαρμόζουμε στις εκάστοτε συνθήκες. Αυτό θα συνιστούσε μια *ερμηνεία* των κανόνων και θα οδηγούσε σε επ' άπειρον αναδρομή. Ούτε, πολύ λιγότερο, "έχουμε στο νου μας" κάποιο νοητικό είδωλο που εξίσου θα απαιτούσε *ερμηνεία*. Όταν ακολουθούμε *επιδέξια* έναν κανόνα, λειτουργούμε ως *χρήστες* και όχι ως *ερμηνευτές* του κανόνα γιατί δεν υπάρχει απόσταση μεταξύ του κανόνα και της εφαρμογής του. Όταν ακολουθούμε *επιδέξια* έναν κανόνα, δεν έχουμε μια προτασιακή γνώση που στη συνέχεια την εφαρμόζουμε στο τρέχον πλαίσιο, αλλά διαθέτουμε μια πρακτική γνώση δεξιοτήτων που μας επιτρέπει να ανταποκρινόμαστε κατάλληλα στις εκάστοτε συνθήκες. Αν δεχθούμε την *καντιανή* κατά βάση άποψη ότι οι έννοιες είναι κανόνες, τότε η κατανόηση μιας έννοιας δεν είναι τίποτε άλλο από την *επιδέξια* χρήση της έννοιας. Σε συμφωνία με αυτή τη γραμμή σκέψης, ο Putnam ακολουθώντας τον Wittgenstein, ισχυρίζεται ότι η κατανόηση μιας έννοιας δεν καθορίζεται από το τι και πώς εμφανίζεται στο υποκείμενο, αλλά από τον τρόπο χρήσης της⁵⁰⁹. Με άλλα λόγια, θα μπορούσαμε να πούμε ότι η γνώση μιας έννοιας δεν καθορίζεται από την στιγμιακή συνείδηση, δηλαδή δεν μπορεί να εξατομικευτεί εστιάζοντας στα περιεχόμενα της στιγμιακής συνείδησης.

Για τους παραπάνω λόγους, ο Putnam (1981a, σελ. 20) οδηγείται στο συμπέρασμα ότι η φαινομενολογική έρευνα δεν βοηθάει στην εξήγηση της κατανόησης και της σκέψης: "...η απόπειρα να κατανοηθεί η σκέψη

⁵⁰⁹ Βλ., επίσης, McGinn 1989, σελ. 32 n45 και McCulloch 1995, σελ. 118.

μέσω αυτού που αποκαλείται 'φαινομενολογική' έρευνα είναι θεμελιωκά παραπλανητική· διότι εκείνο που οι φαινομενολόγοι δεν καταφέρνουν να δουν είναι ότι αυτό που περιγράφουν είναι η εσωτερική έκφραση της σκέψης, αλλά η κατανόηση αυτής της έκφρασης – η κατανόηση που έχει κανείς των ιδίων των δικών του σκέψεων – δεν είναι ένα συμβάν (occurrence) αλλά μια ικανότητα (ability)". Αν ακολουθήσει κανείς τον Putnam, τότε συνάγεται ότι στο προσωπικό επίπεδο η ανάλυση των νοητικών φαινομένων δεν μπορεί να βασιστεί στη "φαινομενολογική έρευνα" των νοητικών φαινομένων, αλλά μόνο στην καθημερινή οπτική του τρίτου προσώπου όπου βασική εξηγητική έννοια είναι η έννοια της χρήσης.

Αν και η κριτική του Putnam στην "φαινομενολογική έρευνα" - όπως την αντιλαμβάνεται - είναι ορθή, το παραπάνω συμπέρασμα δεν προκύπτει. Ο Putnam φαίνεται να παραμένει εδώ εγκλωβισμένος σε μια καρτεσιανή εικόνα και αντιλαμβάνεται την φαινομενολογική έρευνα ως ένα είδος ενδοσκόπησης που μας αποκαλύπτει την "εσωτερική έκφραση της σκέψης", δηλαδή κάποια "εσωτερικά φαινόμενα"⁵¹⁰. Στο όγδοο κεφάλαιο ασκήσαμε κριτική σε αυτή την εικόνα και επιχειρήσαμε να αποκαθάρουμε, κατά το δυνατόν, την οπτική του πρώτου προσώπου από καρτεσιανές προκείμενες. Πιο συγκεκριμένα, δείξαμε ότι στο φαινομενολογικό επίπεδο δεν μας εμφανίζονται νοητικές αναπαραστάσεις ή εσωτερικά αντικείμενα, αλλά ο ίδιος ο κόσμος. Όταν κατανοώ κάτι τότε αυτό το ίδιο μου φανερώνεται με ένα συγκεκριμένο τρόπο. Κατά συνέπεια, η φανέρωση του κόσμου αποτελεί συστατικό "χαρακτηριστικό" της κατανόησης του κόσμου. Ταυτόχρονα, κάθε φορά που κατανοώ κάτι έχω πάντα μια συναίσθηση ότι εγώ είμαι αυτός που το κατανοεί. Έτσι, ένα δεύτερο συστατικό "χαρακτηριστικό" της κατανόησης είναι ότι πάντα "ανήκει" σε κάποιον.

Ο Putnam προσπαθώντας να αποφύγει τον καρτεσιανισμό φτάνει στο άλλο άκρο και περιγράφει το τι σημαίνει κατανόηση με όρους χρήσης, δηλαδή αποκλειστικά από την καθημερινή οπτική του τρίτου προσώπου. Κατ' αυτόν τον τρόπο ωστόσο, φαίνεται να οδηγείται σε δύο τουλάχιστον προβλήματα.

Καταρχάς, δεν είναι σαφές πώς μπορεί να αποφύγει το συμπεριφορισμό, διότι δεν διαθέτει το εννοιολογικό οπλοστάσιο που θα του επέτρεπε να διακρίνει την επιδέξια χρήση (που προσιδιάζει, μεταξύ άλλων, στην κατανόηση μιας έννοιας) από την αντανάκλαστική "χρήση" που τυγχάνει να είναι ορθή αλλά δεν εκλαμβάνεται από το ίδιο το υποκείμενο που την επιτελεί ως δικιά του πράξη. Εδώ χρειάζεται η έννοια της συναίσθησης της ικανότητας του πράττειν που αναδεικνύεται από τη φαινομενολογική διερεύνηση.

⁵¹⁰ Βλ. το προαναφερθέν παράθεμα από το Putnam 1981a, σελ. 19-20.

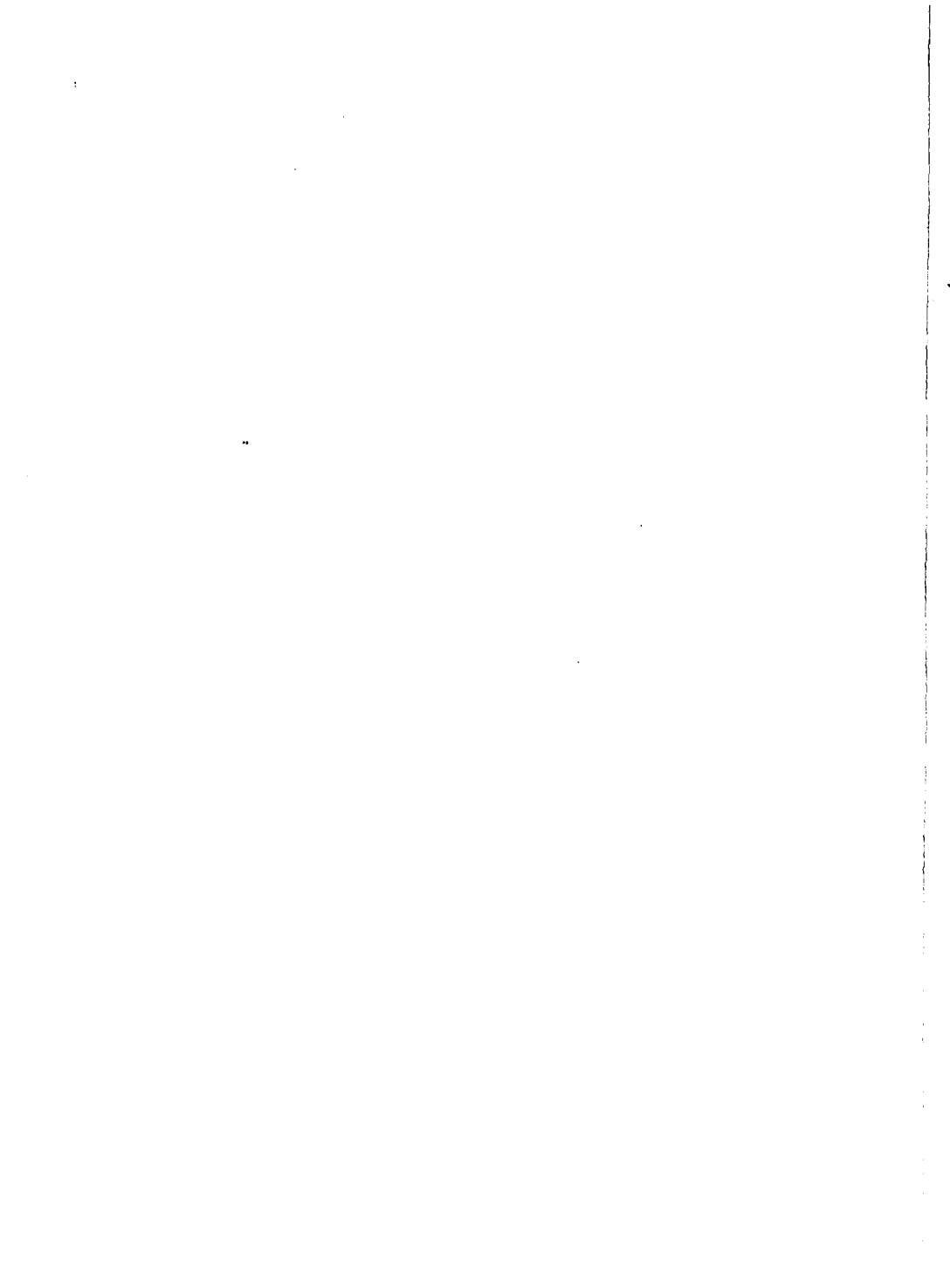
Το δεύτερο πρόβλημα έχει να κάνει με την κατανόηση (εξωτερικών) αναπαραστάσεων: αν ο μόνος τρόπος που μπορεί να αναλυθεί είναι με όρους χρήσης, τίθεται το ερώτημα πώς διακρίνεται η χρήση αναπαραστάσεων από τη χρήση άλλου είδους εργαλείων. Όταν χρησιμοποιώ ένα πράγμα ως εικόνα κάνω κάτι πολύ διαφορετικό απ' ό,τι όταν το χρησιμοποιώ ως σφυρί. Στο τρίτο κεφάλαιο υποστηρίξαμε ότι αν κάποιος γνωρίζει τη χρήση αναπαραστάσεων, τότε, στην παρουσία του αναπαραστασιακού φορέα, έχει την ικανότητα να συμπεριφέρεται ως εάν να είναι αντιμέτωπος με το αναπαριστώμενο αντικείμενο με τη διαφορά ότι δεν το αντιλαμβάνεται ως σωματικά παρόν. Ακριβέστερα, όταν κάποιος γνωρίζει τη χρήση αναπαραστάσεων τότε του φανερώνονται αντικείμενα ενεργεία και συμπεριφέρεται έναντι αυτών ανάλογα⁵¹¹. Αν δεν συμπεριληφθεί η έννοια της φανέρωσης στην περιγραφή με όρους χρήσης που γίνεται στην έννοια της αναπαράστασης, τότε δεν είναι σαφές πώς μπορεί να διακριθεί η χρήση αναπαραστάσεων από τη χρήση άλλων εργαλείων. Το ίδιο πρόβλημα φαίνεται να προκύπτει και για την αντίληψη, τουλάχιστον τη *θεματική* αντίληψη. Όταν βλέπω απέναντί μου κάτι ως σφυρί, δηλαδή *κατανοώ* τι βλέπω, τότε το σφυρί μου φανερώνεται ενεργεία, δηλαδή μου παρουσιάζεται. Αυτός ο τρόπος σχετισμού με το σφυρί είναι πολύ διαφορετικός από εκείνον που έχω όταν το χρησιμοποιώ "τυφλά" για να επιτελέσω ένα έργο. Προφανώς και ο δεύτερος τρόπος σχετισμού αποτελεί κάποιο είδος κατανόησης το οποίο, ωστόσο, φαίνεται να είναι πολύ διαφορετικού τύπου από το πρώτο είδος κατανόησης. Αν προσπαθήσουμε να περιγράψουμε τη διαφορά αποκλειστικά με όρους χρήσης, τότε ενδεχομένως θα οδηγηθούμε στη διαπίστωση ότι στην πρώτη περίπτωση έχουμε μια "παθητική" σχέση με το αντικείμενο, ενώ στη δεύτερη μια "ενεργητική". Αυτή η διάκριση, ωστόσο, διαστρέφει το χαρακτήρα της αντίληψης η οποία, όπως υποστηρίξαμε στο όγδοο κεφάλαιο (ενότητα 8.2.2.2.2), είναι μια δυναμική δραστηριότητα του ανθρώπου και όχι ένα στιγμιακό και στατικό συμβάν. Ακόμα, όμως, κι αν δεχθούμε αυτή την ορθή, κατά τη γνώμη μας, θέση για την αντίληψη, ανακύπτει εκ νέου το πρόβλημα: πώς μπορεί να διακριθεί *αποκλειστικά* με όρους χρήσης η θεματική αντίληψη ενός σφυριού από την "τυφλή" χρήση του σφυριού;

Τα δύο προαναφερθέντα προβλήματα αναδεικνύουν το πόσο δύσκολο είναι να περιγραφούν τα νοητικά φαινόμενα αποκλειστικά από την καθημερινή οπτική του τρίτου προσώπου, δηλαδή αποκλειστικά με όρους χρήσης. Στην παρούσα διατριβή δεν επιχειρήσαμε και ούτε είχαμε ως στόχο να αποδείξουμε ότι ένα τέτοιο εγχείρημα είναι αδύνατο. Εκείνο που επιδιώξαμε, μιλώντας εντελώς γενικά, είναι να συνδυάσουμε την ανάλυση με όρους χρήσης με μια μη καρτεσιανή φαινομενολογική περιγραφή της πράξης και της αντίληψης με στόχο την ανάπτυξη μιας μη

⁵¹¹ Βλ. κεφάλαιο 8, ενότητα 8.2.2.2.3.

αναπαραστασιακής περιγραφής των νοητικών φαινομένων. Ειδικότερα, στο τέταρτο κεφάλαιο ασχοληθήκαμε με το ζήτημα της επιδέξιας πράξης ενώ στο όγδοο κεφάλαιο εξετάσαμε το ζήτημα της (θεματικής) αντίληψης. Θεωρούμε ότι τα αποτελέσματα αυτών των κεφαλαίων δείχνουν ότι η φαινομενολογική περιγραφή μπορεί να λειτουργήσει συμπληρωματικά με την ανάλυση με όρους χρήσης βοηθώντας στην επίλυση πρόβλημάτων όπως τα δύο που προαναφέρθηκαν. Πολλά ζητήματα παραμένουν φυσικά ανοικτά όσον αφορά την καθημερινή οπτική του τρίτου προσώπου και την μη καρτεσιανή οπτική του πρώτου προσώπου, καθώς και τη μεταξύ τους σχέση. Για παράδειγμα, ένα βασικό ερώτημα που εγείρεται είναι πώς μπορεί να αναλυθεί περαιτέρω η συναίσθηση της ικανότητας του πράττειν. Υπάρχουν αρκετές έννοιες σε αυτή τη θεματική περιοχή που χρειάζεται να διερευνηθούν και να αναδειχθούν οι μεταξύ τους σχέσεις, όπως για παράδειγμα, η κιναισθηση, το συναίσθημα, το αίσθημα, το έμβιο σώμα, η διάθεση, η δέσμευση (commitment), η γνώση υποβάθρου, η δεξιότητα, κ.α.⁵¹² Πολλά άλλα ζητήματα εγείρονται όσον αφορά τις δύο οπτικές και τη μεταξύ τους σχέση τα οποία θα εξετασθούν σε επόμενες εργασίες.

⁵¹² Για μια πρώτη διερεύνηση ορισμένων από αυτές τις έννοιες βλέπε, ενδεικτικά, Bermudez, Marcel & Eilan 1995.



ΠΑΡΑΡΤΗΜΑ 1: Η θεωρία σημειογραφίας του Nelson Goodman

Εισαγωγή

Ο Goodman στο βιβλίο του *Languages of Art* αναπτύσσει μια αρκετά σύνθετη θεωρία σημειογραφίας (notation) με βάση την οποία ταξινομεί τα διάφορα είδη αναπαραστάσεων. Παρότι στόχος του σε αυτό το βιβλίο είναι, όπως διαφαίνεται και από τον τίτλο του, να εξετάσει τις διαφορετικές μορφές αναπαράστασης που χρησιμοποιούνται στην τέχνη, η θεωρία σημειογραφίας καθ' εαυτή βρίσκει ευρύτερη εφαρμογή. Στο πλαίσιο της παρούσας εργασίας η συγκεκριμένη θεωρία θα μας φανεί χρήσιμη για τρεις κυρίως λόγους:

1. παρέχει σαφή κριτήρια για τη διάκριση των γλωσσικών από τις εικονιστικές αναπαραστάσεις.
2. εισάγει ένα μεγάλο αριθμό όρων με βάση τους οποίους μπορούν να γίνουν σημαντικές διακρίσεις τόσο μεταξύ των γλωσσικών όσο και μεταξύ των εικονιστικών αναπαραστάσεων και, τέλος,
3. ορίζει τι είναι αναλογικός και τι ψηφιακός υπολογιστής

Με βάση όλες αυτές τις διακρίσεις, η θεωρία του Goodman προσδιορίζει πολύ συγκεκριμένα κριτήρια που πρέπει να πληρούνται προκειμένου ένα αναπαραστασιακό σύστημα- όπως το υποτιθέμενο αναπαραστασιακό σύστημα του νου- να είναι γλωσσικής ή, αντίθετα, εικονιστικής μορφής⁵¹³. Παράλληλα με την παρουσίαση της θεωρίας του Goodman, θα επιχειρήσουμε και την εφαρμογή της στον "χώρο" των νοητικών αναπαραστάσεων.

Η επίθεση κατά της έννοιας της ομοιότητας

Καμία από τις συνθήκες που θέτει ο Goodman για να διακρίνει τις εικονιστικές από τις γλωσσικές αναπαραστάσεις δεν βασίζεται στην έννοια της ομοιότητας. Η ύπαρξη ομοιότητας ανάμεσα σε δύο αντικείμενα δεν αποτελεί ούτε επαρκή ούτε αναγκαία συνθήκη προκειμένου το ένα να συνιστά απεικόνιση του άλλου. Ο Goodman δικαιολογεί αυτή τη θέση του ως εξής: η σχέση ομοιότητας σε αντίθεση με τη σχέση απεικόνισης είναι συμμετρική και ανακλαστική. Συγκεκριμένα, αν το α ομοιάζει προς το β,

⁵¹³ Το γεγονός ότι ο Goodman είναι συμβατιστής δεν απαγορεύει την εφαρμογή των πορισμάτων του στην υποτιθέμενη αντισυμβατιστική γλώσσα της σκέψης. Το ερώτημα σε τι βασίζεται η σχέση αναπαράστασης -αναπαραστώμενου είναι ανεξάρτητο από το ερώτημα για τα ποιοτικά χαρακτηριστικά των διαφόρων μορφών αναπαράστασης.

τότε και το β ομοιάζει προς το α, ενώ αν το α απεικονίζει το β τότε το β δεν απεικονίζει το α. Επίσης, κάθε αντικείμενο ομοιάζει προς τον εαυτό του, δεν ισχύει όμως ότι και κάθε αντικείμενο απεικονίζει τον εαυτό του. Άρα, λοιπόν, δεν είναι κάθε σχέση ομοιότητας σχέση απεικόνισης. Μήπως όμως ισχύει το αντίθετο, δηλαδή, κάθε σχέση απεικόνισης είναι σχέση ομοιότητας; Ούτε και αυτό μπορεί να υποστηριχθεί γιατί ορισμένες εικόνες αναφέρονται σε φανταστικά αντικείμενα και, κατά συνέπεια, δεν υπάρχει κάτι προς το οποίο να ομοιάζουν. Επίσης, το γεγονός ότι μια καρικατούρα μπορεί να διαφέρει εντελώς από το πρόσωπο που απεικονίζει δεν μας εμποδίζει να την θεωρούμε απεικόνιση του προσώπου.

Ο Goodman δεν αρκείται, απλώς, στο να δείξει ότι η ομοιότητα δεν αποτελεί ούτε αναγκαία ούτε ικανή συνθήκη για την απεικόνιση, αλλά επιτίθεται επίσης στην ίδια την έννοια της ομοιότητας.

Συνηθίζουμε να λέμε ότι μια απεικόνιση είναι τόσο περισσότερο ρεαλιστική όσο περισσότερο μοιάζει με αυτό που απεικονίζει. Όσο περισσότερες ιδιότητες του αντικειμένου συλλαμβάνει η απεικόνιση, τόσο πιο πιστή και επιτυχής είναι. Ο Goodman απορρίπτει ρητά μια τέτοια προσέγγιση, διότι αυτή βασίζεται στην παραδοχή ότι τα αντικείμενα υπάρχουν κατά ένα μόνο τρόπο και κατά συνέπεια δεν εξαρτώνται από τον τρόπο με τον οποίο τα βλέπουμε. Μόνο αν ισχύει αυτή η παραδοχή έχει νόημα να ισχυρισθεί κανείς ότι μια απεικόνιση μοιάζει περισσότερο ή λιγότερο με το απεικονιζόμενο. Σύμφωνα με τον Goodman, το πώς είναι τα πράγματα εξαρτάται ουσιαστικά από τις ανάγκες και τις προκαταλήψεις μας⁵⁴. Η ιδέα του αθώου βλέμματος και η σύστοιχη προς αυτή ιδέα του γυμνού δεδομένου δεν είναι παρά μύθοι. Παραφράζοντας τον Καντ λέει ότι "η αθά ματιά είναι τυφλή και ο παρθένος νους άδειος" (σελ. 8).

Ο Goodman φυσικά δεν απορρίπτει εντελώς την έννοια της ομοιότητας. Αυτό που απορρίπτει είναι τη θέση ότι πρόκειται περί μιας αντικειμενικής ιδιότητας, μιας ιδιότητας δηλαδή που δεν εξαρτάται από την ερμηνεία του υποκειμένου. Η διαπίστωση ότι δύο αντικείμενα μοιάζουν βασίζεται, σύμφωνα με τον Goodman, σε μια πρακτική που υπαγορεύει ότι ορισμένες μόνο ιδιότητες⁵⁵ των δύο αντικειμένων είναι σημαντικές, ενώ πολλές άλλες δεν λαμβάνονται υπόψη.

⁵⁴ Πρόκειται για ρητή αναφορά στο περίφημο άρθρο του Bruner (1957) και τη νέα προσέγγιση στη ψυχολογία που ονομάστηκε "Νέα ματιά στην αντίληψη" (New Look in Perception). Βλ. Κεφ. 1, ενότητα 1.1.2.

⁵⁵ Ένα ερώτημα που ανακύπτει εδώ αλλά δεν θα εξετάσουμε περαιτέρω είναι αν αυτές οι ιδιότητες προ-υπάρχουν αντικειμενικά; Ο ίδιος ο Goodman λέει ότι "οποιαδήποτε δύο πράγματα έχουν τον ίδιο αριθμό κοινών ιδιοτήτων. Η ομοιότητα ποικίλει με την σπουδαιότητα που συγκριτικά έχουν οι κοινές ιδιότητες και έτσι ποικίλει με το ενδιαφέρον, το πλαίσιο και τη συνήθεια." (Goodman 1988, σελ. 122).

Η Θεωρία Σημειογραφίας

Ο Goodman έχοντας απορρίψει τη θέση ότι η ομοιότητα μπορεί να αποτελέσει είτε ικανή είτε αναγκαία συνθήκη για την απεικόνιση, θέτει ως στόχο του να προσδιορίσει κάποιο εναλλακτικό κριτήριο για τη διάκριση εικονιστικών και γλωσσικών αναπαραστάσεων. Έτσι, λοιπόν, ξεκινά από τη διαπίστωση ορισμένων διαφορών τους που δεν προϋποθέτουν την έννοια της ομοιότητας. Μία τέτοια διαφορά είναι ότι οι γλωσσικές αναπαραστάσεις, σε αντίθεση με τις εικονιστικές, διαφυλάσσουν την ταυτότητα ενός έργου τέχνης από εκτέλεση σε εκτέλεση. Το αντίγραφο μιας παρτιτούρας αναφέρεται στο ίδιο ακριβώς μουσικό έργο με εκείνο στο οποίο αναφέρεται και η πρωτότυπη. Κάτι τέτοιο προφανώς δεν ισχύει για κάποια άλλα είδη τέχνης. Κάθε ζωγραφικός πίνακας θεωρείται μοναδικός και δεν αντιγράφεται, διότι πέρα από τα ανυπέβλητα προβλήματα που παρουσιάζει η ακριβής αντιγραφή του, το αντίγραφο δεν θεωρείται ότι αποτελεί το ίδιο το έργο τέχνης, αλλά απλώς μια απομίμηση. Αντίθετα μια παρτιτούρα (όπως και ένα ποίημα) μπορεί πολύ εύκολα να αντιγραφεί και κάθε αντίγραφο της θεωρείται εξίσου γνήσιο με το πρωτότυπο⁵¹⁶. Μάλιστα, δεν χρειάζεται καν η παρτιτούρα να είναι από χαρτί: οι νότες θα μπορούσαν εξίσου καλά να σκαλιστούν σε πέτρα, να διαβαστούν και να ηχογραφηθούν σε κασέτα. Υπάρχει, λοιπόν, μια ανεξαρτησία από το μέσο. Αντίθετα ένας ζωγραφικός πίνακας είναι αναπόσπαστος από τον υλικό φορέα του. Μία παρτιτούρα (ή ένα ποίημα), σε αντίθεση με έναν ζωγραφικό πίνακα, δεν εξαρτάται από την ιστορία παραγωγής της. Επίσης, κάθε ζωγραφικός πίνακας αλλοιώνεται με το πέρασμα του χρόνου και αυτό έχει ως συνέπεια την αλλοίωση της ταυτότητας του έργου. Οποιαδήποτε μεταβολή του πίνακα έχει σημασία. Αντίθετα οι αλλοιώσεις που υφίσταται μια παρτιτούρα δεν επηρεάζουν την ταυτότητα του μουσικού έργου⁵¹⁷.

Σε τι έγκεινται αυτές οι διαφορές; Η θέση του Goodman είναι ότι η παρτιτούρα σε αντίθεση με ένα ζωγραφικό πίνακα αποτελεί σημειογραφικό σύστημα και αυτό είναι που διασφαλίζει την ταυτότητα του μουσικού έργου από εκτέλεση σε εκτέλεση με όλες τις επακόλουθες ιδιότητες που αναφέραμε. Μόνο τα σημειογραφικά συστήματα καθιστούν δυνατή την ταυτοποίηση ενός έργου ανεξάρτητα από την ιστορία παραγωγής του και έτσι καθιστούν γνήσια οποιαδήποτε αντιγραφή του. Πότε όμως ένα σύστημα είναι σημειογραφικό; Ο Goodman καταλήγει ότι πρέπει να πληρούνται πέντε συνθήκες: ένα σύστημα είναι σημειογραφικό όταν είναι συντακτικά αποσυναρθρωμένο (disjoint) και διαφοροποιημένο (differentiated), σημασιολογικά αποσυναρθρωμένο και διαφοροποιημένο

⁵¹⁶ Γι' αυτό ίσως δεν έχει πολύ νόημα να υπογράφει κανείς το ποίημά του, όπως συνηθίζουν οι ζωγράφοι με τα έργα τους.

⁵¹⁷ Εκτός φυσικά από την περίπτωση που οι αλλοιώσεις είναι τόσο μεγάλες ώστε τμήματα της παρτιτούρας έχουν καταστραφεί ολοκληρωτικά.

καθώς και μη-αμφίσημο. Σε ό,τι ακολουθεί θα εξετάσουμε αυτές τις συνθήκες αναλυτικά.

Συντακτικές συνθήκες

Ο Goodman εξετάζει αρχικά τις συντακτικές ιδιότητες των σημειογραφικών συστημάτων. Ένα σημειογραφικό σύστημα ιδωμένο από συντακτικής απόψεως ονομάζεται σημειογραφικό σχήμα. Η διάκριση γίνεται διότι μπορούμε να θεωρήσουμε σημειογραφικά σχήματα χωρίς σημασιολογικές-αναπαραστασιακές⁵¹⁸ ιδιότητες. Μόνο όταν ένα σημειογραφικό σχήμα αποκτά αναπαραστασιακές ιδιότητες καθίσταται σημειογραφικό σύστημα. Τα σημειογραφικά σχήματα είναι μια ειδική κατηγορία συμβολικών σχημάτων. Όπως κάθε συμβολικό σχήμα, αποτελούνται από χαρακτήρες και τρόπους συνδυασμού αυτών για τον σχηματισμό σύνθετων χαρακτήρων. Κάθε χαρακτήρας είναι μια τάξη σημάδιων (marks)⁵¹⁹. Κάθε σημάδι ενός χαρακτήρα μπορεί να αντικατασταθεί με οποιοδήποτε άλλο σημάδι του ίδιου χαρακτήρα χωρίς αυτό να επιφέρει την παραμικρή διαφορά από συντακτική άποψη. Με άλλα λόγια, όλα τα σημάδια ενός χαρακτήρα είναι *συντακτικώς ισοδύναμα*. Αυτό δεν σημαίνει ωστόσο ότι είναι, κατ' ανάγκη, *μορφικά* όμοια: για να ανήκουν δύο σημάδια στον ίδιο χαρακτήρα δεν απαιτείται να είναι το ένα ακριβές αντίγραφο (exact duplicate) του άλλου. Για παράδειγμα, ένα κεφαλαίο σίγμα ("Σ") και ένα μικρό σίγμα ("σ") παρότι πολύ διαφορετικά μεταξύ τους ανήκουν στον ίδιο χαρακτήρα, ενώ αντίθετα ένα μικρό σίγμα ("σ") και ένα μικρό όμικρον ("ο") παρά την ομοιότητά τους ανήκουν σε διαφορετικούς χαρακτήρες⁵²⁰. Ο Goodman, για να διακρίνει τα σημάδια του ίδιου χαρακτήρα από τα ακριβή αντίγραφα τα ονομάζει "ομοιώτυπα" (replicas). "Κανένας βαθμός ομοιότητας δεν είναι αναγκαίος ή ικανός για την ύπαρξη ομοιώτυπας" (σελ. 131).

Ο Goodman ονομάζει τα ομοιώτυπα "σημάδια αδιάφορα ως προς τον χαρακτήρα" (character-indifferent marks): δύο σημάδια είναι αδιάφορα ως προς τον χαρακτήρα εάν κάθε ένα ανήκει σε κάποιο χαρακτήρα και κανένα δεν ανήκει σε χαρακτήρα που το άλλο δεν ανήκει. Με βάση αυτή την έννοια, διατυπώνει την 1^η συνθήκη που πρέπει να πληρούν τα σημειογραφικά συστήματα, τη συνθήκη της συντακτικής

⁵¹⁸ Ο Goodman στο *Languages of Art* με τον όρο "αναπαραστασιακός" αναφέρεται αποκλειστικά στις εικονιστικές αναπαραστάσεις. Για λόγους συνέπειας με τα προηγούμενα κεφάλαια δεν θα ακολουθήσουμε αυτή τη σύμβαση, αλλά με τον όρο αυτό θα αναφερόμαστε γενικότερα σε κάθε είδος σχέσης όπου ένα πράγμα παραπέμπει σε κάποιο άλλο.

⁵¹⁹ Με την διάκριση ανάμεσα σε χαρακτήρα και σημάδι ο Goodman θέλει ουσιαστικά να συλλάβει τη γνωστή διάκριση τύπου/έκτυπου (type/token). (Βλ. Peirce 1967)

⁵²⁰ Επίσης, κάθε άνθρωπος έχει διαφορετικό γραφικό χαρακτήρα και μάλιστα ορισμένοι έχουν εξαιρετικά "κακό" γραφικό χαρακτήρα.

αποσυνάρθρωσης (disjointness): οι χαρακτήρες ενός σημειογραφικού συστήματος πρέπει να είναι συντακτικώς αποσυναρθρωμένοι, δηλαδή σε κάθε έναν να ανήκουν σημάδια αδιάφορα ως προς τον χαρακτήρα. Συνέπεια αυτής της συνθήκης είναι ότι κανένα σημάδι δεν μπορεί να ανήκει σε περισσότερους του ενός χαρακτήρες. Για παράδειγμα, ακριβώς επειδή τα γράμματα του αλφάβητου αποτελούν συντακτικώς αποσυναρθρωμένους χαρακτήρες, όταν δούμε ένα σημάδι γράμματος που μοιάζει τόσο με τον χαρακτήρα "ζ" όσο και με τον χαρακτήρα "ς" δεν προχωρούμε παρά μόνο αν διαπιστώσουμε σε ποιόν από τους δύο ανήκει - είναι αδιανόητο το να ανήκει και στους δύο.

Σε αυτό το σημείο θα πρέπει να επισημάνουμε ότι ο Goodman αποσυνδέει το αίτημα της αποσυνάρθρωσης από την αντιληπτική ικανότητα που έχει ο χρήστης να την διαπιστώσει. Το αίτημα αυτό αφορά απλώς το πώς είναι δομημένο το σημειογραφικό σχήμα. Για παράδειγμα, ένα σημάδι γράμματος γνωρίζουμε ότι ανήκει σε ένα μόνο χαρακτήρα ακόμα και αν δεν καταφέρουμε ποτέ να ανακαλύψουμε ποιος είναι αυτός.

Ωστόσο, υπάρχουν συμβολικά συστήματα που είναι κατά τέτοιο τρόπο ορισμένα ώστε είναι φυσικώς αδύνατο να προσδιορισθεί αν ένα σημάδι ανήκει σε ένα συγκεκριμένο χαρακτήρα ή σε κάποιον άλλο. Για παράδειγμα, έστω ότι ορίζουμε ένα συμβολικό σχήμα που αποτελείται από ευθύγραμμο τμήματα διαφορετικού μήκους. Κάθε ευθύγραμμο τμήμα μήκους x αποτελεί ένα σημάδι του χαρακτήρα x , όπου x ανήκει στους ρητούς αριθμούς. Είναι φανερό ότι ένα τέτοιο σχήμα θα έχει άπειρους χαρακτήρες. Το πρόβλημα με αυτό το σχήμα έγκειται στο ότι είναι φυσικώς αδύνατο να διαπιστωθεί το εάν κάποιον σημάδι ανήκει στον χαρακτήρα x ή σε κάποιον κοντινό του.⁵²¹ Αυτό συμβαίνει επειδή οι χαρακτήρες του συγκεκριμένου σχήματος δεν είναι πεπερασμένα διαφοροποιημένοι (finitely differentiated).

Με βάση την έννοια της πεπερασμένης διαφοροποίησης ο Goodman ορίζει την δεύτερη συνθήκη που πρέπει να πληρούν τα σημειογραφικά συστήματα: "Για κάθε δύο χαρακτήρες K και K' και κάθε σημάδι m που στην πραγματικότητα δεν ανήκει και στους δύο, ο καθορισμός του εάν το m δεν ανήκει στο K , ή του εάν το m δεν ανήκει στο K' είναι θεωρητικά δυνατός." (σελ. 136). Με τον όρο "θεωρητικά δυνατός" εννοεί όλες τις περιπτώσεις που δεν είναι λογικά ή μαθηματικά αδύνατες (π.χ. άπειρη διαφοροποίηση).

Το παράδειγμα που χρησιμοποιήθηκε για να εισαχθεί η έννοια της πεπερασμένης διαφοροποίησης, υποβάλλει, ενδεχομένως, την ιδέα ότι αυτή

⁵²¹ Το βαθύτερο πρόβλημα είναι ότι το σχήμα επιβάλλει συνθήκες που αφορούν την μορφή των δειγμάτων: δύο σημάδια του ίδιου χαρακτήρα δεν μπορούν παρά να είναι ακριβή αντίγραφα. Κάθε διαφοροποίησή τους τα καθιστά σημάδια διαφορετικού χαρακτήρα. Δεν υπάρχει περιθώριο διαφοροποίησης μέσα στο πλαίσιο του ίδιου χαρακτήρα.

συνδέεται άμεσα με το αν το συμβολικό σχήμα διαθέτει ένα πεπερασμένο αριθμό χαρακτήρων. Ωστόσο, κάτι τέτοιο δεν ισχύει. Αν, αντί για ευθύγραμμο τμήματα επιλέξουμε ως χαρακτήρες την ανηγμένη κλασματική μορφή των ρητών στην αραβική αρίθμηση, τότε παίρνουμε ένα συμβολικό σχήμα ξανά με άπειρους χαρακτήρες το οποίο ωστόσο είναι πεπερασμένο διαφοροποιημένο: για κάθε σημάδι ρητού αριθμού μπορεί χωρίς δυσκολία να καθορισθεί το σε ποιον χαρακτήρα αυτό ανήκει.

Σε τι έγκειται όμως αυτή η διαφορά, δεδομένου ότι και τα δύο συμβολικά σχήματα έχουν άπειρο αριθμό χαρακτήρων; Μια προφανής απάντηση είναι ότι η αραβική αρίθμηση βασίζεται σε 10 ατομικούς χαρακτήρες και όλοι οι υπόλοιποι κατασκευάζονται ως συνδυασμοί, ενώ στο συμβολικό σχήμα των ευθυγράμμων τμημάτων δεν υπάρχουν σύνθετοι χαρακτήρες. Ο Goodman ωστόσο δεν αναφέρει μια τέτοιου τύπου εξήγηση, ούτε χρησιμοποιεί καθόλου τη διάκριση ανάμεσα σε ατομικούς και σύνθετους χαρακτήρες προκειμένου να εξηγήσει την πεπερασμένη διαφοροποίηση. Ο λόγος, ενδεχομένως, είναι ότι η διαπίστωση ότι ένα σημάδι ανήκει σε ένα ατομικό και όχι σε έναν σύνθετο χαρακτήρα δεν συνεπάγεται τη διαπίστωση της ταυτότητας του συγκεκριμένου χαρακτήρα⁵²². Αυτό μπορεί να φανεί καλύτερα με το ακόλουθο παράδειγμα: έστω ένα συμβολικό σχήμα που αποτελείται από δύο μόνο ατομικούς χαρακτήρες με βάση τους οποίους στη συνέχεια κατασκευάζεται μια απειρία σύνθετων χαρακτήρων. Οι δύο ατομικοί χαρακτήρες ορίζονται ως εξής:

ο πρώτος χαρακτήρας είναι η κλάση όλων των ευθυγράμμων τμημάτων μήκους x , όπου $1 > x \geq 2$,

ο δεύτερος χαρακτήρας είναι η κλάση όλων των ευθυγράμμων τμημάτων μήκους x , όπου $2 > x \geq 3$,

Το πρόβλημα με αυτό το συμβολικό σχήμα είναι ότι δεν είναι δυνατό για κάθε σημάδι του να προσδιορισθεί σε ποιόν χαρακτήρα ανήκει: για ένα σημάδι μήκους πολύ κοντά στο 2 δεν είναι θεωρητικά δυνατό (με την έννοια που προαναφέρθηκε) να προσδιορισθεί σε ποιον από τους δύο ατομικούς χαρακτήρες αυτό ανήκει. Βλέπουμε λοιπόν ότι η πεπερασμένη διαφοροποίηση δεν είναι απόρροια της ύπαρξης ενός πεπερασμένου αριθμού ατομικών χαρακτήρων. Κατά συνέπεια, η ιδιότητα της πεπερασμένης διαφοροποίησης δεν μπορεί να θεωρηθεί ότι οφείλεται ούτε στην ύπαρξη ενός πεπερασμένου αριθμού χαρακτήρων. Ανακεφαλαιώνοντας, λοιπόν, θα πρέπει να πούμε ότι η ύπαρξη πεπερασμένης διαφοροποίησης είναι ανεξάρτητη από το εάν το συμβολικό

⁵²² Πρέπει να παραδεχτούμε, ωστόσο, ότι την καθιστά ευκολότερη. Το γεγονός ότι ο Goodman δεν λαμβάνει υπόψη του αυτό το κριτήριο, δεν πρέπει να του χρεωθεί ως παράληψη, διότι στόχος του είναι να προσδιορίσει τις λογικές και όχι τις πραγματιστικές συνθήκες που πρέπει να πληρούνται για να είναι ένα σύστημα σημειογραφικό.

σχήμα έχει άπειρο ή πεπερασμένο αριθμό χαρακτήρων. Η ιδιότητα αυτή έχει να κάνει με το πώς σχετίζονται μεταξύ τους σημάδια διαφορετικών χαρακτήρων: αν ορισμένα από τα σημάδια δύο ή περισσότερων χαρακτήρων βρίσκονται *απείρως κοντά*, τότε δεν υπάρχει θεωρητικά δυνατότητα να εντοπιστεί σε ποιον από αυτούς τους χαρακτήρες ανήκουν⁵²³.

Θα πρέπει να επισημάνουμε σε αυτό το σημείο ότι η έννοια της *θεωρητικής* δυνατότητας στην οποία επιμένει ο Goodman, πέρα από το ότι δεν είναι αρκετά σαφής, δεν φαίνεται να είναι ούτε και δικαιολογημένη. Τα σημειογραφικά συστήματα μπορούμε να πούμε ότι καθίστανται τέτοια στο βαθμό που χρησιμοποιούνται. Χρήση θα πει ανάγνωση και εγγραφή. Η πρώτη δραστηριότητα, σχηματικά μιλώντας, έγκειται στην αναγνώριση ότι ένα σημάδι ανήκει σε κάποιον χαρακτήρα, ενώ η δεύτερη στην παραγωγή ενός σημάδιού με βάση τον χαρακτήρα του. Ωστόσο, και οι δύο αυτές δραστηριότητες μπορούν να πραγματοποιηθούν μόνο στο βαθμό που είναι *πρακτικά* δυνατή η διαπίστωση της διαφοροποίησης των χαρακτήρων (κατά τρόπο που προσιδιάζει στο εκάστοτε σύστημα⁵²⁴). Ένα σημειογραφικό σύστημα που έχει κατασκευαστεί έτσι ώστε οι χαρακτήρες του να είναι πεπερασμένα διαφοροποιημένοι είναι εντελώς άχρηστο αν δεν είναι ταυτόχρονα *πρακτικά* δυνατό να διαπιστωθεί αυτή η διαφοροποίηση. Ο ίδιος ο Goodman σε ένα μεταγενέστερο κείμενό του επισημαίνει: "Στην πράξη αυτό που έχει σημασία δεν είναι απλά το εάν η διαφοροποίηση είναι θεωρητικά δυνατή αλλά το εάν μπορεί να επιτευχθεί με μέσα διαθέσιμα και κατάλληλα για τη δεδομένη χρήση του δεδομένου σχήματος"⁵²⁵. Για παράδειγμα, οι συνήθεις υπολογιστές κάνουν χρήση δύο χαρακτήρων του "0" και του "1". Τα σημάδια του "1" είναι όλες εκείνες οι καταστάσεις κατά τις οποίες σε κάποιο συγκεκριμένο σημείο του υπολογιστή περνά ρεύμα (περίπου) 5 Volt ενώ τα σημάδια του "0" εκείνες κατά τις οποίες στο σημείο αυτό δεν περνά ρεύμα. Η διαφορά μεταξύ 0 και 5 Volt είναι επίτηδες μεγάλη για να αποφευχθεί οποιαδήποτε "σύγχυση" εξαιτίας κάποιας τυχαίας φόρτισης. Με άλλα λόγια, το πόσο "μάκριά" πρέπει να βρίσκονται οι χαρακτήρες ενός σημειογραφικού συστήματος εξαρτάται

⁵²³ Ένα ανοικτό ερώτημα εδώ είναι: από ποια άποψη τα σημάδια βρίσκονται απείρως κοντά; Από μορφική άποψη; Σε αυτή την περίπτωση, όμως, δεν επανεισάγεται ένα κριτήριο με βάση την ομοιότητα –όχι μεταξύ σημάδιού και αναπαριστώμενου αλλά μεταξύ σημαδιών; Εξάλλου, όπως είπαμε το σημάδι ενός γράμματος μπορεί να είναι απείρως κοντά από μορφική άποψη στα σημάδια κάποιου άλλου γράμματος, χωρίς αυτό να δημιουργεί κανένα πρόβλημα.

⁵²⁴ Αυτό μπορεί να επιτευχθεί με πολλούς διαφορετικούς τρόπους; με μέτρηση ή γενικότερα τη χρήση οργάνων, με σύγκριση, με την βοήθεια των συμφραζομένων κ.λπ.

⁵²⁵ Goodman (1988), σελ. viii

άμεσα από τις διακριτικές ικανότητες⁵²⁶ του χρήστη και όχι από το αν το σύστημα είναι απλώς θεωρητικά διαφοροποιημένο. Σε ό,τι ακολουθεί, θα υιοθετήσουμε, λοιπόν, τη συνθήκη περί πεπερασμένης διαφοροποίησης με την αίρεση ότι πρέπει να είναι *πρακτικώς* δυνατή η διαπίστωση της ισχύος της από τον χρήστη του συστήματος. Η έννοια του "πρακτικώς" διευκρινίζεται περαιτέρω με τις δύο παρακάτω επισημάνσεις:

1. Η χρήση ενός συμβολικού σχήματος, όπως ήδη επισημάναμε, δεν έγκειται μόνο στις διαδικασίες ανάγνωσης αλλά και στις διαδικασίες εγγραφής. Ο τρόπος με τον οποίο ο Goodman διατυπώνει τη συνθήκη της πεπερασμένης διαφοροποίησης περιορίζει τη μεσολάβηση του χρήστη στο πρώτο είδος διαδικασιών, διότι κάνει λόγο για "[θεωρητική δυνατότητα] καθορισμού του εάν το σημάδι π δεν ανήκει στον χαρακτήρα Κ ή δεν ανήκει στον χαρακτήρα Κ'". Είναι προφανές ότι η συνθήκη θα πρέπει να συμπληρωθεί ώστε να προβλέπει και την *πρακτική δυνατότητα εγγραφής* ενός σημαδιού με βάση κάποιον χαρακτήρα.
2. Όπως δείχνει το παράδειγμα με τον υπολογιστή, ο χρήστης ενός συμβολικού σχήματος δεν είναι αναγκαίο να είναι άνθρωπος, ούτε είναι αναγκαίο τα σημάδια του σχήματος να είναι "εξωτερικά" ως προς τον χρήστη. Τίποτα μέσα στο ορισμό που δίνει ο Goodman για την πεπερασμένη διαφοροποίηση δεν αποκλείει αυτή την επέκταση του τρόπου χρήσης.


Οι δύο παραπάνω επισημάνσεις δείχνουν ότι η συνθήκη του Goodman θα μπορούσε ενδεχομένως να χρησιμοποιηθεί ακόμα και στον "χώρο" της γλώσσας της σκέψης. Δηλαδή, σε ένα υποτιθέμενο συμβολικό σύστημα που είναι αξεδιάλυτο από τον χρήστη του και το οποίο, προφανώς δεν υπόκειται μόνο σε διαδικασίες ανάγνωσης αλλά και σε διαδικασίες εγγραφής.

Κατά την διερεύνηση της έννοιας της πεπερασμένης διαφοροποίησης είδαμε ότι υπάρχουν μη-πεπερασμένα διαφοροποιημένα σχήματα τόσο με πεπερασμένο όσο και με μη πεπερασμένο αριθμό χαρακτήρων. Ο Goodman για να διακρίνει αυτά τα δύο είδη σχημάτων ονομάζει τα τελευταία "συντακτικώς πυκνά": ένα συμβολικό σχήμα είναι συντακτικώς πυκνό "αν παρέχει άπειρα πολλούς χαρακτήρες διατεταγμένους κατά τέτοιο τρόπο ώστε ανάμεσα σε κάθε δύο από αυτούς να υπάρχει ένας τρίτος" (σελ. 136). Η έννοια της συντακτικής πυκνότητας θα μας φανεί πολύ χρήσιμη παρακάτω διότι, όπως θα δούμε, μέσω αυτής,

⁵²⁶ Οι οποίες, ωστόσο, δεν θα πρέπει να θεωρηθεί ότι βασίζονται μόνο στις αντιληπτικές δυνατότητες του χρήστη, αλλά και στις πραγματολογικές του γνώσεις (π.χ. αναγνώριση της ταυτότητας ενός γράμματος από τα συμφραζόμενα).

σε συνδυασμό με την έννοια της σημασιολογικής πυκνότητας, ορίζεται το τι είναι αναλογικό σύστημα.

Αν ένα σύστημα δεν είναι συντακτικώς πυκνό, τότε ονομάζεται συντακτικώς ασυνεχές (discontinuous). Τα ασυνεχή συστήματα δεν είναι κατ' ανάγκη διαφοροποιημένα. Όπως είδαμε υπάρχουν μη διαφοροποιημένα σχήματα με πεπερασμένο αριθμό χαρακτήρων (βλέπε περιοχή 2 στο παρακάτω σχεδιάγραμμα⁵²⁷)

Πεπερασμένα διαφοροποιημένα Σχήματα		Μη-πεπερασμένα διαφοροποιημένα Σχήματα
Α σ υ ν ε χ ή	Ε χ ή μ α τ α	Πυκνά Σχήματα
1	2	

Σχήμα 1

Σημασιολογικές συνθήκες

Ο Goodman συμπληρώνει τη θεωρία του περί σημειογραφικών συστημάτων με τρεις επιπλέον συνθήκες που τις αντλεί από τον χώρο της σημασιολογίας. Κατ' αυτόν, ένα συμβολικό σχήμα καθίσταται συμβολικό σύστημα όταν συσχετίζεται με ένα πεδίο αναφοράς. Προτού προχωρήσουμε, πρέπει να εισαγάγουμε κάποια ορολογία. Κάθε αντικείμενο του πεδίου αναφοράς παραπέμπεται (is denoted) από κάποιο χαρακτήρα ή, ισοδύναμα, συναρμόζει (complies) με κάποιο χαρακτήρα. Όλα τα αντικείμενα του πεδίου αναφοράς που συναρμόζουν με κάποιον χαρακτήρα αποτελούν την τάξη συναρμογής (compliance class) ή την έκταση (extension) του χαρακτήρα. Κάθε αντικείμενο της έκτασης ενός χαρακτήρα θα ονομάζεται και συναρμόζον (compliant). Ο Goodman διευκρινίζει ότι η συναρμογή του αντικειμένου δεν είναι βασικά με τον χαρακτήρα αλλά με κάποιο σημάδι του χαρακτήρα.

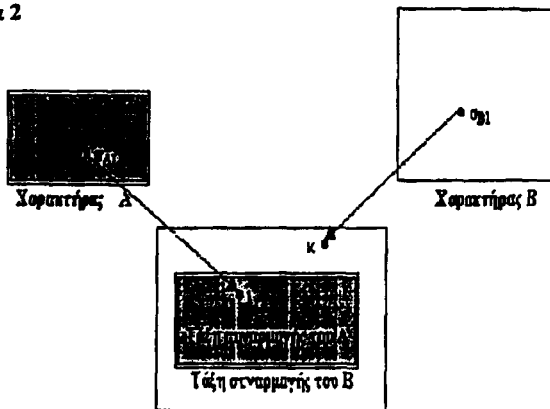
Η πρώτη σημασιολογική συνθήκη που πρέπει να πληρούν τα σημειογραφικά συστήματα είναι η μη-αμφισημία. Ένα σύστημα είναι μη-αμφίσημο αν κάθε χαρακτήρας του είναι μη-αμφίσημος. Για να είναι ένας χαρακτήρας μη-αμφίσημος θα πρέπει τα σημάδια του να πληρούν δύο περιορισμούς. Ας δούμε ποιοι είναι αυτοί. Κατ' αρχήν, κάθε σημάδι ενός χαρακτήρα μπορεί να παραπέμπει σε περισσότερα του ενός πράγματα. Αν η τάξη των πραγμάτων στην οποία παραπέμπει ένα σημάδι, δηλαδή η τάξη συναρμογής του, είναι η ίδια σε κάθε χρονική στιγμή και σε κάθε πλαίσιο

⁵²⁷ Το σχήμα επίσης προϋποθέτει ότι τα ασυνεχή και πυκνά σχήματα δεν εξαντλούν το σύνολο των δυνατών σχημάτων. Μπορούμε να φανταστούμε σχήματα που δεν προϋποθέτουν την έννοια της διάταξης.

τότε το σημάδι είναι μη-αμφίσημο (unambiguous). Προκειμένου ένας χαρακτήρας να είναι μη-αμφίσημος δεν αρκεί κάθε σημάδι του να είναι μη αμφίσημο, θα πρέπει επίσης όλα τα σημάδια του να έχουν την ίδια τάξη συναρμογής. Διαφορετικά, ορισμένα σημάδια αυτού του χαρακτήρα, παρότι συντακτικώς ισοδύναμα, δεν θα είναι και σημασιολογικώς ισοδύναμα. Κατά συνέπεια, η αντικατάσταση του ενός από το άλλο δεν θα διατηρεί την ταυτότητα του έργου.

Ωστόσο, ακόμα κι αν όλα τα σημάδια ενός χαρακτήρα έχουν την ίδια τάξη συναρμογής, δηλαδή παραπέμπουν στα ίδια αντικείμενα, τίποτα μέχρι στιγμής δεν εξασφαλίζει ότι ορισμένα από αυτά τα αντικείμενα δεν ανήκουν και στην τάξη συναρμογής σημαδιών άλλων χαρακτήρων. Σε μια τέτοια περίπτωση, μετά από μια σειρά μεταβάσεων από ένα αντικείμενο στο σημάδι με το οποίο αυτό συναρμόζει και από 'κει σε κάποιο άλλο αντικείμενο της τάξης συναρμογής του σημαδιού κ.ο.κ., είναι ενδεχόμενο να μην καταλήξουμε σε σημάδι του ίδιου χαρακτήρα με εκείνον στον οποίο ανήκει το αρχικό σημάδι. Για παράδειγμα, στο σχήμα 2 μπορεί κανείς να περάσει από το αντικείμενο κ στο σημάδι σ_{B1} του χαρακτήρα Β και από εκεί στο αντικείμενο ι, το οποίο όμως ανήκει και στην τάξη συναρμογής του σημαδιού σ_{A1} του χαρακτήρα Α.

Σχήμα 2



Έτσι λοιπόν, μια δεύτερη σημασιολογική συνθήκη που πρέπει να πληρούν τα σημειογραφικά συστήματα είναι ότι οι τάξεις συναρμογής κάθε χαρακτήρα δεν πρέπει να έχουν κοινά στοιχεία μεταξύ τους. Ο Goodman ονομάζει αυτή τη συνθήκη, συνθήκη της σημασιολογικής αποσυνάρθρωσης.

Οι περισσότερες από τις συνήθειες γλώσσες δεν πληρούν τη συγκεκριμένη συνθήκη. Για παράδειγμα, για οποιουδήποτε δύο όρους που έχουν σχέση είδους-γένους οι αντίστοιχες τάξεις συναρμογής

αποτελούν η μία γνήσιο υποσύνολο της άλλης: η τάξη συναρμογής της λέξης "θηλαστικό" περιλαμβάνει, μεταξύ άλλων, και όλα εκείνα τα αντικείμενα που συναρμόζουν με την λέξη "άνθρωπος".

Τέλος, η τρίτη σημασιολογική συνθήκη που πρέπει να πληρούν τα σημειογραφικά συστήματα είναι η σημασιολογική πεπερασμένη διαφοροποίηση:

"για κάθε δύο χαρακτήρες K και K' τέτοιους ώστε οι τάξεις συναρμογής τους δεν είναι ταυτόσημες, και κάθε αντικείμενο h που δεν συναρμόζει και με τους δύο, ο καθορισμός είτε ότι το h δεν συναρμόζει με το K , είτε ότι το h δεν συναρμόζει με το K' είναι θεωρητικά δυνατό." (σελ. 152)

Ας δούμε ένα παράδειγμα συμβολικού συστήματος που δεν ικανοποιεί αυτή την συνθήκη. Έστω, το συμβολικό σχήμα που αποτελείται απ' όλους τους χαρακτήρες της αραβικής αριθμησης που αντιστοιχούν στους ρητούς και βρίσκονται στην ανηγμένη κλασματική μορφή. Όπως έχουμε δει, το σχήμα αυτό ικανοποιεί τις συντακτικές συνθήκες της αποσυνάρθρωσης και της πεπερασμένης διαφοροποίησης. Αν τώρα θεωρήσουμε ως τάξη συναρμογής όλα τα φυσικά αντικείμενα που καθένα έχει βάρος ίσο με κάποιο ρητό αριθμό κιλών, τότε το συμβολικό σύστημα που ορίζεται δεν πληροί την συνθήκη της πεπερασμένης διαφοροποίησης: για οποιοδήποτε αντικείμενο της τάξης συναρμογής είναι αδύνατο να προσδιορισθεί το εάν αυτό συναρμόζει με κάποιο συγκεκριμένο χαρακτήρα ή με κάποιο άλλον. Καμία μέτρηση, οσοδήποτε ακριβής, δεν μπορεί να πιστοποιήσει κάτι τέτοιο⁵²⁸. Παράλληλα, ωστόσο, το σύστημα πληροί τις συνθήκες της μη-αμφισημίας και της σημασιολογικής αποσυνάρθρωσης. Η πρώτη πληρείται διότι για οποιοδήποτε χαρακτήρα όλα τα σημάδια του έχουν την ίδια τάξη συναρμογής και η δεύτερη συνθήκη πληρείται διότι κάθε αντικείμενο της τάξης συναρμογής εφόσον έχει κάποιο συγκεκριμένο βάρος συναρμόζει με έναν και μόνο χαρακτήρα.

Ο Goodman όπως και στην περίπτωση της συντακτικής πεπερασμένης διαφοροποίησης έτσι και στην σημασιολογική της εκδοχή κάνει λόγο για *θεωρητική* δυνατότητα διαπίστωσής της. Ωστόσο, για τους ίδιους λόγους που αναφέραμε και στην πρώτη περίπτωση, θα υιοθετήσουμε και πάλι την συνθήκη με την τροποποίηση ότι θα πρέπει να είναι *πρακτικά* δυνατή η διαπίστωση της ισχύος της.

Κατ' αντιστοιχία προς την έννοια της συντακτικής πυκνότητας μπορούμε τώρα να ορίσουμε την έννοια της σημασιολογικής πυκνότητας: ένα συμβολικό σύστημα είναι σημασιολογικά πυκνό αν διαθέτει "έναν άπειρο αριθμό χαρακτήρων με τάξεις συναρμογής κατά τέτοιο τρόπο διατεταγμένες ώστε ανάμεσα σε οποιοδήποτε δύο από αυτές να υπάρχει μία τρίτη" (σελ. 153). Θα πρέπει να επισημανθεί ότι δεν είναι μόνο τα

⁵²⁸ Βλέπε, Goodman (1968), σελ. 153.

πυκνά συστήματα αυτά στα οποία δεν ισχύει η συνθήκη της πεπερασμένης διαφοροποίησης. Ακόμα και ένα σημασιολογικά ασυνεχές σύστημα είναι δυνατόν να μην πληροί την συνθήκη (βλέπε σχήμα 3, περιοχή 2). Για παράδειγμα, αν ορίσουμε ένα συμβολικό σύστημα με δύο χαρακτήρες κατά τρόπο ώστε στον πρώτο να ανήκουν όλα τα αντικείμενα που ζυγίζουν ένα κιλό ή λιγότερο και στον δεύτερο όλα τα αντικείμενα που ζυγίζουν περισσότερο από ένα κιλό, τότε το σύστημα αυτό, αν και σημασιολογικά ασυνεχές (μια και αποτελείται από δύο μόνο τάξεις συναρμογής), δεν είναι παρόλα αυτά πεπερασμένα διαφοροποιημένο (διότι για κάποιο αντικείμενο πολύ κοντά στο ένα κιλό δεν είναι δυνατόν να προσδιορισθεί το εάν αυτό συναρμόζει με τον πρώτο ή με τον δεύτερο χαρακτήρα). Το παράδειγμα αυτό δείχνει ότι η ικανοποίηση της συνθήκης της σημασιολογικής αποσυνάρθρωσης δεν συνεπάγεται την συνθήκη της σημασιολογικής πεπερασμένης διαφοροποίησης και όπως δείχνει ο Goodman δεν ισχύει ούτε το αντίθετο. Κατά συνέπεια, οι δύο συνθήκες είναι λογικά ανεξάρτητες.

Πεπερασμένα διαφοροποιημένα Συστήματα	Μη-πεπερασμένα διαφοροποιημένα Συστήματα			
<table border="1"> <tr> <td>Ασυνεχέ 1</td> <td>Συστήματα 2</td> <td>Πυκνά Συστήματα</td> </tr> </table>	Ασυνεχέ 1	Συστήματα 2	Πυκνά Συστήματα	
Ασυνεχέ 1	Συστήματα 2	Πυκνά Συστήματα		

Σχήμα 3

Με βάση τις πέντε συνθήκες της συντακτικής αποσυνάρθρωσης και πεπερασμένης διαφοροποίησης, της μη-αμφισημίας και της σημασιολογικής αποσυνάρθρωσης και πεπερασμένης διαφοροποίησης μπορούν πλέον να ορισθούν τα σημειογραφικά συστήματα ως όλα εκείνα τα συμβολικά συστήματα που πληρούν αυτές τις συνθήκες.

Σε τι διαφέρουν οι εικονιστικές από τις γλωσσικές αναπαραστάσεις;

Με βάση τις παραπάνω διακρίσεις μπορούμε τώρα να προσδιορίσουμε σε τι διαφέρουν οι εικονιστικές από τις γλωσσικές αναπαραστάσεις. Η διαφορά αυτή, σύμφωνα με τον Goodman, δεν θα πρέπει να αναζητηθεί στην εσωτερική δομή μιας αναπαράστασης μεμονωμένα, γιατί η ίδια αναπαράσταση "που περιγράφει σε κάποια συστήματα, μπορεί να απεικονίζει σε κάποια άλλα" (σελ. 231). Το εάν μια αναπαράσταση είναι εικονιστική ή γλωσσική καθορίζεται από το συμβολικό σύστημα στο οποίο ανήκει.

Ας ξεκινήσουμε από τις εικονιστικές αναπαραστάσεις. Μια αναπαράσταση είναι εικονιστική, σύμφωνα με τον Goodman, αν το σύστημα στο οποίο ανήκει είναι συντακτικά και σημασιολογικά πυκνό. Ωστόσο, θα μπορούσε κανείς να αντιτείνει ότι δεν είναι όλες οι εικονιστικές αναπαραστάσεις του ίδιου είδους: υπάρχουν νατουραλιστικοί πίνακες, εξπρεσιονιστικοί πίνακες, σκίτσα, κάρικατούρες, ασπρόμαυρες φωτογραφίες, έγχρωμες φωτογραφίες κ.λπ. Και ακόμα γενικότερα,

υπάρχουν μοντέλα, διαγράμματα, χάρτες κ.λπ. Πώς διαφοροποιούνται όλα αυτά τα διαφορετικά είδη μεταξύ τους; Ο Goodman για να εξηγήσει αυτές τις διαφορές εισάγει ένα πρόσθετο κριτήριο με βάση την έννοια της πληρότητας (repleteness). Η πληρότητα έχει να κάνει με την ποσότητα των χαρακτηριστικών μιας αναπαράστασης που παίζουν συμβολικό ρόλο μέσα στο πλαίσιο ενός αναπαραστασιακού συστήματος. Ένα σύμβολο είναι τόσο περισσότερο πλήρες όσο "αναλογικά περισσότερα από τα χαρακτηριστικά του λειτουργούν συμβολικά"⁵²⁹. Για παράδειγμα, σε ένα απλό διάγραμμα, συνήθως το πάχος και το χρώμα των γραμμών δεν έχει κανένα αναπαραστασιακό ρόλο. Το μόνο που έχει σημασία είναι το σχήμα των γραμμών. Αντίθετα σε μια εικόνα όλα είναι σημαντικά: το χρώμα, το σχήμα, το πάχος των γραμμών, η υφή της επιφάνειας κ.λπ. Όσο περισσότερες διαστάσεις μεταβολής έχουν συμβολική λειτουργία εντός ενός συστήματος τόσο περισσότερο εικονιστικές είναι οι αναπαραστάσεις εντός αυτού. Οι απεικονίσεις έχουν υψηλό βαθμό πληρότητας⁵³⁰. Υπ' αυτή την έννοια, ένα μη-βαθμονομημένο θερμόμετρο, δεδομένου ότι έχει μία μόνο διάσταση μεταβολής με συμβολική λειτουργία δεν μπορεί να θεωρηθεί ότι απεικονίζει την θερμοκρασία, διότι έχει πολύ χαμηλή πληρότητα⁵³¹.

Ανακεφαλιώνοντας, λοιπόν, τα εικονιστικά συστήματα χαρακτηρίζονται από συντακτική και σημασιολογική πυκνότητα και υψηλή πληρότητα. Ας εξετάσουμε τώρα ποια χαρακτηριστικά ικανοποιούν οι γλωσσικές αναπαραστάσεις. Τα γλωσσικά συστήματα περιλαμβάνουν, μεταξύ άλλων, τα σημειογραφικά συστήματα, δηλαδή συστήματα που ικανοποιούν τις πέντε συνθήκες που έχουμε ήδη αναφέρει. Ωστόσο, υπάρχουν και γλωσσικά συστήματα που δεν είναι σημειογραφικά. Όπως είδαμε, η συνήθης (διαλογική) γλώσσα είναι αμφίσημη και σημασιολογικά μη-αποσυναρθρωμένη. Επίσης, ούτε η συνθήκη της σημασιολογικής πεπερασμένης διαφοροποίησης ικανοποιείται πάντα⁵³². Όλες όμως οι γλώσσες πληρούν τις συνθήκες της συντακτικής αποσυναρθρωσης και της συντακτικής πεπερασμένης διαφοροποίησης⁵³³. Κατά συνέπεια, η ικανοποίηση των δύο αυτών συνθηκών αποτελεί αναγκαίο (όχι όμως και ικανό) όρο προκειμένου ένα συμβολικό σύστημα να είναι γλωσσικό. Αν η υποτιθέμενη γλώσσα της σκέψης είναι πράγματι

⁵²⁹ Βλέπε, Goodman (1988), σελ. 123.

⁵³⁰ Goodman 1988, σελ. 123.

⁵³¹ Ένα εύλογο ερώτημα εδώ είναι ποια θα ήταν η ελάχιστη πληρότητα που θα καθιστούσε μια αναπαράσταση εικονιστική; Επίσης, ποια θα ήταν η μέγιστη πληρότητα πάνω από την οποία μια αναπαράσταση θα σταματούσε να είναι εικονιστική (και θα γίνονταν για παράδειγμα μοντέλο, δείγμα κ.λπ.); Πιστεύω ότι οποιαδήποτε επιβολή κάποιων τέτοιων ορίων θα ήταν αυθαίρετη.

⁵³² Βλ. σχετικό παράδειγμα στο Goodman (1968), σελ. 163.

⁵³³ Βλ. Goodman (1968), σελ. 163, υποσημ. 17.

γλώσσα θα πρέπει και αυτή να ικανοποιεί τις συνθήκες της συντακτικής αποσυνάρθρωσης και της συντακτικής πεπερασμένης διαφοροποίησης.

Αναλογικοί και Ψηφιακοί Υπολογιστές

Κατά την παρουσίαση της θεωρίας του Fodor είδαμε ότι μια διαφορά ανάμεσα στους μορφομετατροπείς και τους αυτόνομους υπομηχανισμούς είναι ότι οι πρώτοι αποτελούν αναλογικούς υπολογιστές ενώ οι δεύτεροι ψηφιακούς υπολογιστές. Ωστόσο, η διάκριση ανάμεσα σε αναλογικούς και ψηφιακούς υπολογιστές παραμένει σε ένα διαισθητικό επίπεδο μέσα στο έργο του Fodor αλλά και γενικότερα μέσα στην γνωσιακή επιστήμη, γεγονός που καθιστά προβληματική κάθε θεωρία που έχει ως αφηρησία της αυτές τις δύο έννοιες⁵³⁴.

Ο Goodman μπορεί να αναπληρώσει αυτό το κενό γιατί με βάση τις διακρίσεις που αναπτύσσει για τη θεωρία σημειογραφίας δίνει έναν σαφή ορισμό των αναλογικών και των ψηφιακών συστημάτων. Συγκεκριμένα, ορίζει τα αναλογικά συστήματα ως όλα εκείνα τα συμβολικά συστήματα που είναι συντακτικά και σημασιολογικά πυκνά, ενώ τα ψηφιακά ως εκείνα τα συμβολικά συστήματα που είναι συντακτικά και σημασιολογικά πεπερασμένα διαφοροποιημένα.

Ας δούμε ορισμένα παραδείγματα. Μια ζυγαριά με δείκτη που περιστρέφεται γύρω από το κέντρο ενός μη βαθμονομημένου δισκου είναι ένας αναλογικός υπολογιστής γιατί συνιστά ένα συντακτικά και σημασιολογικά πυκνό σύστημα. Συγκεκριμένα, αν θεωρηθεί ότι κάθε θέση του δείκτη προσδιορίζει διαφορετικό χαρακτήρα τότε ανάμεσα σε κάθε δύο χαρακτήρες υπάρχει ένας τρίτος. Παρόμοια στο σημασιολογικό επίπεδο, ανάμεσα σε κάθε δύο αντικείμενα διαφορετικού βάρους υπάρχει πάντα κάποιο τρίτο με βάρος μεταξύ των δύο πρώτων. Αντίθετα, ένας κερματοδέκτης με ψηφιακή ένδειξη αποτελεί, σύμφωνα με τον ορισμό του Goodman, ψηφιακό υπολογιστή: διότι, στον βαθμό που η ψηφιακή ένδειξη χρησιμοποιεί την αραβική αρίθμηση, το σύστημα είναι συντακτικά πεπερασμένα διαφοροποιημένο. Επίσης, είναι σημασιολογικά πεπερασμένα διαφοροποιημένο διότι για κάθε επιμέρους κέρμα μπορεί να καθορισθεί σε ποιον τύπο κέρματος (δηλαδή, σε ποιον χαρακτήρα) ανήκει.

Τα ψηφιακά και τα αναλογικά συστήματα δεν εξαντλούν όλο το φάσμα των δυνατών συμβολικών συστημάτων: υπάρχουν συστήματα που δεν είναι ούτε ψηφιακά ούτε αναλογικά. Όπως φαίνεται και στο Σχήμα 4 (βλέπε γκριζες περιοχές (2)-(8)), τέτοιου τύπου συστήματα μπορούμε να πάρουμε με πολλούς διαφορετικούς τρόπους.

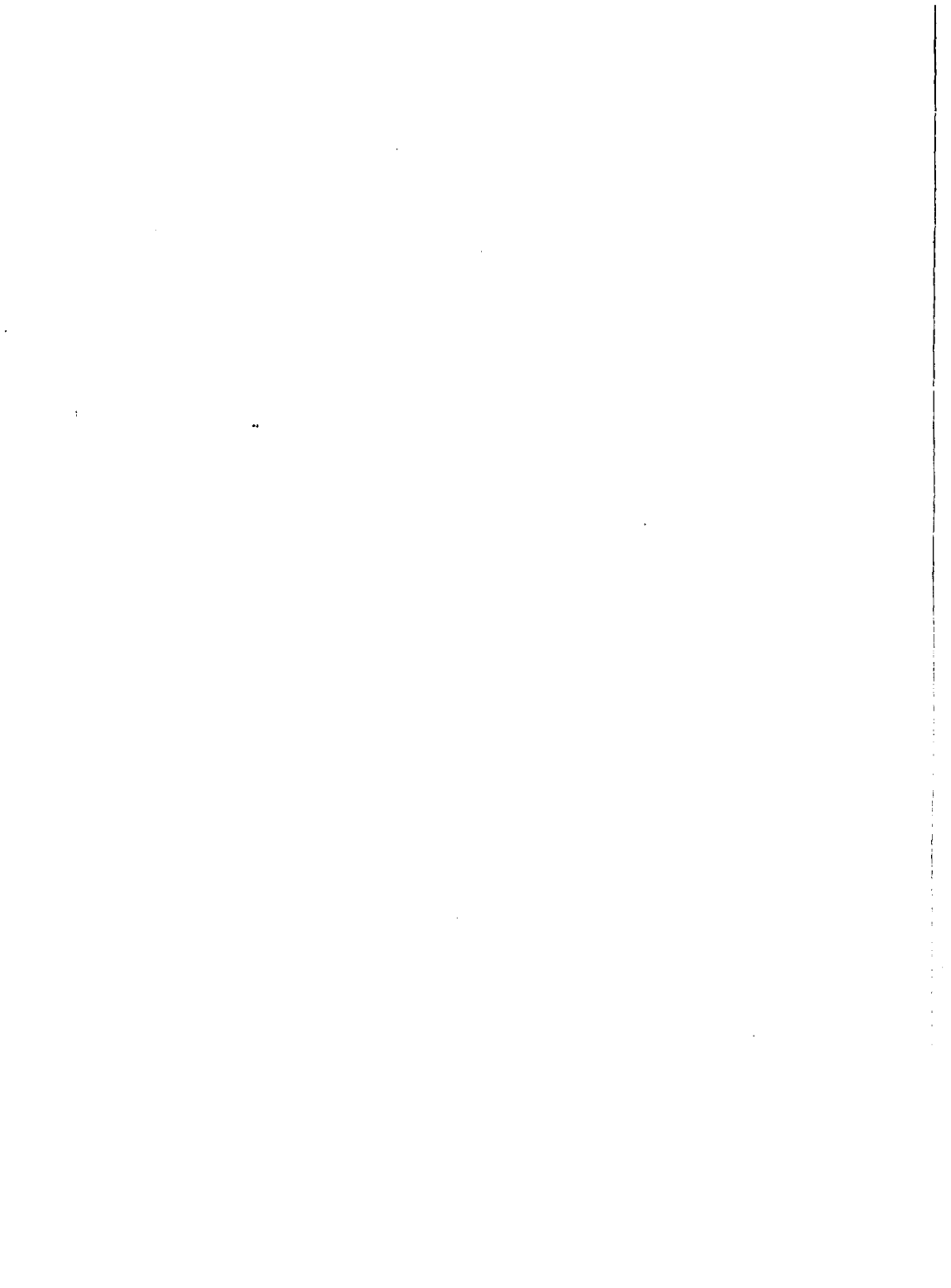
Ας δούμε ένα παράδειγμα συστήματος που δεν είναι ούτε αναλογικό ούτε ψηφιακό από την περιοχή (7), δηλαδή ένα σύστημα που

⁵³⁴ Τα τελευταία χρόνια η κατάσταση αυτή αρχίζει να αλλάζει με την εισαγωγή της θεωρίας δυναμικών συστημάτων στη μελέτη των γνωσιακών διαδικασιών. Βλ. παράρτημα 2 και van Gelder 1999.

είναι συντακτικά πεπερασμένα διαφοροποιημένο και σημασιολογικά πυκνό. Τέτοιο παράδειγμα θα αποτελούσε μια ζυγαριά με δείκτη που περιστρέφεται γύρω από έναν βαθμονομημένο δίσκο. Έστω ότι η βαθμονόμηση χωρίζει την περιφέρεια του δίσκου σε 10 περιοχές έτσι ώστε κάθε μία από αυτές να έχει κάποιο μικρό κενό από τις δύο γειτονικές της. Προκειμένου τώρα η ζυγαριά να είναι συντακτικώς πεπερασμένα διαφοροποιημένη θα πρέπει όλα τα σημεία της εκάστοτε περιοχής να θεωρούνται ότι ανήκουν σε έναν χαρακτήρα. Έτσι, κάθε φορά που ο δείκτης βρίσκεται πάνω σε μία από τις παραπάνω περιοχές θα θεωρηθεί ότι ενδεικνύει έναν από τους 10 χαρακτήρες. Αν, αντίθετα, θεωρηθεί ότι αυτό που έχει σημασία είναι η απόλυτη θέση του δείκτη πάνω στην περιφέρεια του δίσκου και ότι η βαθμονόμηση απλώς υποβοηθεί τον προσδιορισμό της θέσης, τότε το σύστημα εξακολουθεί να είναι συντακτικά πυκνό. Βλέπουμε, λοιπόν, ότι το είδος ενός συμβολικού συστήματος καθορίζεται από τον τρόπο που χρησιμοποιείται (εν προκειμένω, από τον τρόπο ανάγνωσής του).

Σχήμα 4

		Συντακτικές Συνθήκες			
		Συντακτική ασυνέχεια		Συντακτική πυκνότητα	
		Συντακτική πεπερασμένη διαφοροποίηση	Συντακτική μη-πεπερασμένη διαφοροποίηση	Συντακτική μη-πεπερασμένη διαφοροποίηση	
Σημασιολογικές Συνθήκες	Σημασιολογική ασυνέχεια	Σημασιολογική πεπερασμένη διαφοροποίηση	Ψηφιακά συστήματα (1)	Μη-ψηφιακά μη-αναλογικά συστήματα (2)	Μη-ψηφιακά μη-αναλογικά συστήματα (3)
		Σημασιολογική μη-πεπερασμένη διαφοροποίηση	Μη-ψηφιακά μη-αναλογικά συστήματα (4)	Μη-ψηφιακά μη-αναλογικά συστήματα (5)	Μη-ψηφιακά μη-αναλογικά συστήματα (6)
	Σημασιολογική πυκνότητα	Σημασιολογική μη-πεπερασμένη διαφοροποίηση	Μη-ψηφιακά μη-αναλογικά συστήματα (7)	Μη-ψηφιακά μη-αναλογικά συστήματα (8)	Αναλογικά συστήματα (9)



ΠΑΡΑΡΤΗΜΑ 2: Η προσέγγιση των γνωσιακών διαδικασιών μέσω δυναμικών συστημάτων

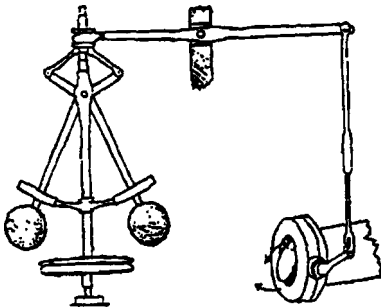
Τα τελευταία χρόνια έχει αναπτυχθεί μια νέα προσέγγιση μέσα στο πλαίσιο της γνωσιακής επιστήμης η οποία αμφισβητεί την πρωτοκαθεδρία του ψηφιακού υπολογιστή ως μοντέλου για την εξήγηση του νου. Πρόκειται για την προσέγγιση των γνωσιακών διαδικασιών μέσω δυναμικών συστημάτων (Dynamic Systems Approach to Cognition). Η δυναμικιστική προσέγγιση έκανε δυναμικά την εμφάνισή της το 1995 με την έκδοση από τους van Gelder & Port του συλλογικού τόμου *Mind as Motion: Explorations in the Dynamics of Cognition*⁵³⁵. Φυσικά επιμέρους εργασίες προς αυτή την κατεύθυνση υπήρχαν ήδη από παλιότερα στον ορίζοντα. Μάλιστα, η βασική ιδέα δεν είναι και καινούργια, αλλά έχει υποστηριχθεί από την εποχή της Κυβερνητικής⁵³⁶. Ωστόσο, η έκδοση του συγκεκριμένου βιβλίου προσέδωσε το χαρακτήρα ενός συνολικού ερευνητικού προγράμματος που αντιπαρατίθεται επί ίσοις όρους στην κλασική γνωσιακή προσέγγιση. Η εξέλιξη αυτή δεν είναι τυχαία. Η αλματώδης πρόοδος της υπολογιστικής τεχνολογίας επέτρεψε κυρίως από τα μέσα της δεκαετίας του 1980 την προσομοίωση της συμπεριφοράς σύνθετων δυναμικών συστημάτων - μια δυνατότητα που δεν υπήρχε στην εποχή της Κυβερνητικής.

Αφετηριακή παραδοχή των δυναμικιστών είναι ότι "[τ]α φυσικά γνωσιακά συστήματα αποτελούν συγκεκριμένα είδη δυναμικών συστημάτων και είναι καλύτερα να κατανοηθούν από την άποψη της

⁵³⁵ Είναι ενδιαφέρον να παρατηρήσει κανείς την προσπάθεια μίμησης της δομής που έχει ο τίτλος της "βιβλίου" του συνδετισμού των McLelland & Rumelhart (1986): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*.

⁵³⁶ Βλ. Ashby 1952 και Wiener 1948/1961. Η Κυβερνητική ασχολούταν με τη μελέτη των μαθηματικών των μηχανών, όχι από την άποψη του τρόπου σύνδεσης των λειτουργικών συστατικών μιας μηχανής αλλά από την άποψη όλων των δυνατών συμπεριφορών που μια μηχανή μπορεί να παράγει. Η μηχανή χαρακτηριζόταν μέσω των εισόδων και των εξόδων της, ενώ η εσωτερική λειτουργία της αντιμετωπιζόταν ως μαύρο κουτί. Τα εργαλεία ανάλυσης ήταν συχνά διαφορικές και ολοκληρωτικές εξισώσεις και αυτά τα εργαλεία περιόρισαν την Κυβερνητική, κατά την περίοδο εμφάνισής της, στη μελέτη καταστάσεων όπου οι οριακές συνθήκες δεν άλλαζαν γρήγορα. Σε αντίθεση με την Τεχνητή Νοημοσύνη, η Κυβερνητική δεν βασιζόταν στην έννοια της συμβολικής αναζήτησης - κάθε έρευνα διεξαγόταν με όρους ελαχιστοποίησης μιας συνάρτησης. Τέλος, η Κυβερνητική πρόσβευε ότι ο οργανισμός θα έπρεπε να μοντελοποιηθεί μαζί με το περιβάλλον προκειμένου να κατανοηθεί η παραγόμενη συμπεριφορά.

δυναμικής" (van Gelder & Port 1995a, σελ. 5). Γι' αυτό οι δυναμικιστές ανατρέχουν σε έννοιες και εργαλεία από την μαθηματική θεωρία δυναμικών συστημάτων για να αναλύσουν τις γνωσιακές διαδικασίες. Έτσι, από μαθηματική σκοπιά, το κύριο μέσο προσομοίωσης των μελετώμενων φαινομένων δεν είναι πλέον η θεωρία υπολογισμού αλλά η θεωρία δυναμικών συστημάτων. Ο οργανισμός και το περιβάλλον μοντελοποιούνται ως δύο δυναμικά συστήματα σε ζεύξη⁵³⁷. Ο εγκέφαλος παύει να θεωρείται χειριστής συμβολικών αναπαραστάσεων και αντιμετωπίζεται ως ένας ρυθμιστής της ενσώματης και ενταγμένης στο περιβάλλον δραστηριότητας του οργανισμού (van Gelder 1995). Σύμφωνα με τους δυναμικιστές, ο εγκέφαλος δεν είναι παρά τμήμα ενός συνολικότερου συστήματος που περιλαμβάνει το σώμα του οργανισμού και το περιβάλλον: "Εφόσον, το νευρικό σύστημα, το σώμα και το περιβάλλον εξελίσσονται όλα συνεχώς και αλληλοεπηρεάζονται ταυτόχρονα, το γνωσιακό σύστημα δεν μπορεί να είναι απλώς ο απομονωμένος εγκέφαλος είναι, μάλλον, ένα μοναδικό ενοποιημένο σύστημα που περιλαμβάνει και τα τρία [προαναφερθέντα υποσυστήματα] ... οι εσωτερικές και οι εξωτερικές διαδικασίες βρίσκονται σε ζεύξη κατά τέτοιο τρόπο ώστε και τα δύο σύνολα διαδικασιών επηρεάζουν συνεχώς το ένα το άλλο"⁵³⁸. Στο δυναμικισμό, λοιπόν, η εξήγηση των γνωσιακών διαδικασιών βασίζεται στο μοντέλο των δυναμικών συστημάτων το οποίο αντικαθιστά το κλασικό μοντέλο του ψηφιακού υπολογιστή.



Ο van Gelder είναι, ίσως, ο βασικότερος υποστηρικτής του δυναμικισμού και αυτός που πρώτος επεχείρησε να αρθρώσει με θεωρητικούς όρους τα κοινά γνωρίσματα διαφόρων, ανεξάρτητων μεταξύ τους, επιστημονικών προσεγγίσεων των γνωσιακών διαδικασιών που αντιτίθενται στη κλασική γνωσιακή επιστήμη⁵³⁹. Στο

σημαντικό άρθρο "What might Cognition Be if not Computation?" ο van Gelder (1995) προτείνει, στη θέση της μηχανής Turing, ένα νέο παράδειγμα (exemplar) ως αφετηρία για την κατανόηση των γνωσιακών διαδικασιών. Πρόκειται για το ρυθμιστή του Watt (βλέπε διπλανό σχήμα).

⁵³⁷ Τα δυναμικά συστήματα περιγράφονται κυρίως μέσω διαφορικών εξισώσεων. Ένα παράδειγμα συζευγμένων διαφορικών εξισώσεων είναι το ακόλουθο: $dx/dt=x+y$, $dy/dt=x^2+y$

⁵³⁸ van Gelder & Port 1995, σελ. 13.

⁵³⁹ Βλ. υποσημείωση 503.

Ο ρυθμιστής του Watt είναι ένα σύστημα αυτόματου ελέγχου που επανοήθηκε από τον Watt για να διατηρεί σταθερή, παρά τις αυξομειώσεις του φορτίου, την ταχύτητα περιστροφής ενός άξονα ο οποίος κινείται από μια ατμομηχανή. Αυτό το επιτυγχάνει με το να ανοίγει ή να κλείνει ανάλογα το στόμιο της βαλβίδας του ατμού όταν μειώνεται ή αυξάνεται αντίστοιχα η ταχύτητα περιστροφής του άξονα. Πιο συγκεκριμένα, όπως φαίνεται και στο παρακάτω σχήμα, όσο μεγαλύτερη είναι η ταχύτητα περιστροφής του άξονα, τόσο πιο μακριά από το άξονα περιστρέφονται – λόγω φυγόκεντρης δύναμης– οι δύο σφαίρες που είναι προσαρμοσμένες πάνω του. Ως συνέπεια αυτού, ανυψώνεται ο μοχλός της βαλβίδας και κλείνει ανάλογα το στόμιο.

Σύμφωνα με τον van Gelder (1995), ο ρυθμιστής του Watt αποτελεί ένα δυναμικό σύστημα γιατί δεν χρησιμοποιεί αναπαραστάσεις⁵⁴⁰ και υπολογισμούς επί των αναπαραστάσεων προκειμένου να επιτελέσει το έργο του. Αυτό γίνεται φανερότερο αν εξετάσουμε πως λειτουργεί ένας αντίστοιχος υπολογιστικός ρυθμιστής. Ένας υπολογιστικός ρυθμιστής, λοιπόν, προκειμένου να εκτελέσει το ίδιο έργο πρέπει να ακολουθεί κάποιον αλγόριθμο. Ένας τέτοιος αλγόριθμος θα ήταν, για παράδειγμα, ο εξής:

1. Μέτρησε την ταχύτητα του άξονα
 2. Σύγκρινε την τρέχουσα ταχύτητα με την επιθυμητή
 3. Αν δεν υπάρχει διαφορά, επέστρεψε στο βήμα 1. Αλλιώς,
 - α. μέτρησε την τρέχουσα πίεση του ατμού
 - β. υπολόγισε την επιθυμητή μεταβολή στην πίεση του ατμού
 - γ. υπολόγισε την αναγκαία προσαρμογή της βαλβίδας
 4. κάνε την προσαρμογή της βαλβίδας
- Επέστρεψε στο βήμα 1⁵⁴¹.

Ο ρυθμιστής του Watt, σε αντίθεση με τον υπολογιστικό ρυθμιστή, δεν λειτουργεί ακολουθιακά (sequentially) και κυκλικά: "Εφόσον, δεν υπάρχουν διακριτά βήματα επεξεργασίας, δεν μπορεί να υπάρχει κάποια ακολουθία με την οποία συμβαίνουν αυτά τα [υποτιθέμενα] βήματα"⁵⁴². Στον ρυθμιστή του Watt, "η είσοδος, η εσωτερικής δραστηριότητα και η έξοδος συμβαίνουν όλα συνεχώς και ταυτόχρονα"⁵⁴³. Γι' αυτό το λόγο, η λειτουργία του ρυθμιστή του Watt δεν περιγράφεται υπολογιστικά, δηλαδή με τη βοήθεια κάποιου αλγορίθμου, αλλά με τη χρήση κάποιας διαφορετικής εξίσωσης⁵⁴⁴.

⁵⁴⁰ Πρόκειται για ό,τι στον Επίλογο θα ερμηνεύσουμε ως (ψηφιακές) εγγραφές. Βλ., επίσης, κεφάλαιο 3, ενότητα 3.3.1.4.

⁵⁴¹ van Gelder 1995, σελ. 348

⁵⁴² ibid, σελ. 354

⁵⁴³ ibid, σελ. 354.

⁵⁴⁴ "Γενικά, μια διαφορετική εξίσωση είναι κάθε εξίσωση που περιλαμβάνει μια συνάρτηση και μία ή περισσότερες από τις παραγώγους της" (van Gelder 1995, σελ. 368).

Η διαφορά μεταξύ του ρυθμιστή του Watt και του αντίστοιχου υπολογιστικού συστήματος μοιάζει εκ πρώτης όψεως σαφής. Ωστόσο, αν κανείς επιχειρήσει να προσδιορίσει αυστηρά το σε τι ακριβώς έγκειται η διαφορά μεταξύ υπολογιστικών και δυναμικών συστημάτων εν γένει, βρίσκεται αντιμέτωπος με σημαντικότερα προβλήματα. Πιγή των προβλημάτων αποτελεί, εν μέρει, το ότι δεν υπάρχει μια γενική θεωρία περί του τι είναι υπολογισμός, ούτε ένας γενικός προσδιορισμός της έννοιας του αλγορίθμου και, κατ' επέκταση, της έννοιας των υπολογισμών συναρτήσεων⁵⁴⁵. Αντίστοιχα, δεν υπάρχει γενική συναίνεση ούτε για το τι αποτελεί δυναμικό σύστημα.

Ο van Gelder, με μια σειρά δημοσιεύσεων οι οποίες προκάλεσαν αρκετή συζήτηση, επιχειρήσε να προσδιορίσει τη διαφορά των υπολογιστικών συστημάτων από τα δυναμικά συστήματα, τουλάχιστον σε ένα πρώτο επίπεδο. Σύμφωνα με αυτόν η σημαντικότερη διαφορά έγκειται στο ότι οι μεταβλητές στα δυναμικά συστήματα είναι αριθμητικές, δηλαδή έχουν ποσοτικές ιδιότητες. Τα δυναμικά συστήματα διαφέρουν από τους ψηφιακούς υπολογιστές κατά το ότι είναι ποσοτικά συστήματα. Η έννοια της ποσοτικότητας μπορεί να αφορά είτε τις διάφορες καταστάσεις του συστήματος είτε την αλληλεξάρτηση καταστάσεων - χρόνου είτε, τέλος, το ρυθμό μεταβολής του συστήματος.

Πώς όμως μπορεί να κατανοηθεί πιο συγκεκριμένα η έννοια της ποσοτικότητας σε αυτό το πλαίσιο; Σίγουρα δεν αρκεί να πούμε απλώς ότι οι μεταβλητές στα δυναμικά συστήματα είναι αριθμητικές γιατί και στους ψηφιακούς υπολογιστές θα μπορούσε κανείς να αντιτείνει ότι οι καταστάσεις τους μπορούν να περιγραφούν αριθμητικά. Όλες οι καταστάσεις ενός ψηφιακού υπολογιστή μπορούν να εκφραστούν ως συνδυασμός 0 και 1. Μια δεύτερη επισήμανση είναι ότι η διαφορά που εντοπίζει ο van Gelder δεν θα πρέπει να αναχθεί ούτε στη διαφορά διακριτού - συνεχούς. Ένας εύκολος τρόπος να δεχθεί αυτό είναι ότι υπάρχουν δυναμικά συστήματα που δεν περιγράφονται από διαφορικές εξισώσεις, ήτοι δεν ορίζονται στους πραγματικούς αριθμούς, αλλά περιγράφονται από τις λεγόμενες εξισώσεις διαφορών οι οποίες ορίζονται στους ακεραίους.

Η διαφορά που εντοπίζει ο van Gelder έχει να κάνει με τον τρόπο που σχετίζονται μεταξύ τους οι διάφορες καταστάσεις του συστήματος. Υπαρξη ποσοτικής σχέσης μεταξύ των καταστάσεων του συστήματος σημαίνει ότι μπορεί να ορισθεί μια απόσταση ή μετρική⁵⁴⁶ που

⁵⁴⁵ Βλ. κεφάλαιο 1, ενότητα 1.3.5.

⁵⁴⁶ Απόσταση ή μετρική επί ενός συνόλου E ονομάζεται κάθε μη αρνητική αριθμητική συνάρτηση d ορισμένη επί του ExE και η οποία έχει τις ακόλουθες ιδιότητες:

1. $d(x,y) \geq 0 \Leftrightarrow x=y$
2. $d(x,y) = d(y,x)$, για κάθε $(x,y) \in ExE$ (συμμετρική ιδιότητα)

προσδιορίζει τη σχέση δύο καταστάσεων κατά ένα τέτοιο τρόπο ώστε η απόσταση να σχετίζεται συστηματικά με τη συμπεριφορά του συστήματος. "Πιο συγκεκριμένα, ένα σύστημα είναι ποσοτικό ως προς τις καταστάσεις του εάν υπάρχει μια μετρική που ορίζεται επί του συνόλου των καταστάσεων τέτοια ώστε η συμπεριφορά [του συστήματος] να σχετίζεται συστηματικά με τις αποστάσεις όπως μετρώνται από την μετρική"⁵⁴⁷. Μια μετρική, που όπως προσθέτει ο van Gelder (1998a) στην απάντηση προς τους κριτές του, θα πρέπει να καθορίζεται ανεξάρτητα από τη συμπεριφορά του συστήματος⁵⁴⁸.

Αν τώρα αντιπαραβάλλουμε προς αυτή τη συνθήκη το τι γίνεται στην περίπτωση των ψηφιακών υπολογιστών, θα διαπιστώσουμε ότι σε μια μηχανή Turing η μετάβαση από τη μια κατάσταση στην άλλη δεν μπορεί να συσχετισθεί με κάποια έννοια απόστασης μεταξύ των δύο καταστάσεων διότι αυτές δεν είναι καν συγκρίσιμες ποσοτικά. Παρότι οι καταστάσεις μιας μηχανής Turing μπορούν να περιγραφούν με αριθμούς, αυτοί οι αριθμοί δεν υπάρχουν σε κάποια μετρική, δεν είναι συγκρίσιμοι μεταξύ τους. Η συμπεριφορά μιας μηχανής Turing εξαρτάται κάθε φορά αποκλειστικά από το σε ποια κατάσταση βρίσκεται και δεν σχετίζεται καθόλου με την απόσταση της εν λόγω κατάστασης με άλλες καταστάσεις της μηχανής. Αυτό που έχει σημασία, με άλλα λόγια, είναι ο τύπος κατάστασης στον οποίο ανήκει το εκάστοτε έκτυπο κατάστασης. Για παράδειγμα, αν είναι τύπος 0 ή τύπος 1. Σε αυτό το πλαίσιο δεν έχει κανένα νόημα η απόσταση μεταξύ του 0 και του 1. Οι υπολογιστικές διαδικασίες βασίζονται αποκλειστικά στη δομή των καταστάσεων του συστήματος και οι υπολογιστικοί κανόνες είναι ευαίσθητοι στη δομή των εγγραφών⁵⁴⁹. Όπως επισημαίνει ο van Gelder (1998, σελ. 16): "Οι μηχανές Turing μετακινούνται μέσα στον χώρο των καταστάσεων τους κατά τρόπους που θα έμοιαζαν τελείως ακανόνιστοι μέχρις ότου κανείς συνειδητοποιήσει ότι η διάταξή τους στηρίζεται σε μορφικές ιδιότητες και όχι σε ποσοτικές ιδιότητες". Γι' αυτό και μέσα στο πλαίσιο της

3. $d(x,z) \leq d(x,y) + d(y,z)$ για οποιαδήποτε x,y,z του E (τριγωνική ιδιότητα). (βλ. Χαήνης 1984, σελ. 12).

⁵⁴⁷ van Gelder 1998.

⁵⁴⁸ Ο van Gelder προσθέτει αυτή τη συνθήκη για να αποφύγει ένα έξυπνο κόλπο με το οποίο οι Chater & Hahn (1998, σελ. 633) επεχείρησαν να δείξουν ότι και στις μηχανές Turing μπορεί να οριστεί μια μετρική. Πιο συγκεκριμένα, πρότειναν το εξής: "Ορίστε την ακόλουθη μετρική: η απόσταση μεταξύ δύο καταστάσεων είναι ο ελάχιστος αριθμός βημάτων μεταξύ αυτών. Η συμπεριφορά της μηχανής Turing συστηματικά σχετίζεται προς αυτή τη μετρική (σε κάθε βήμα, η μηχανή θα βιάσει στη γειτονική κατάσταση αυτής της μετρικής)". Όπως επισημαίνει ο van Gelder (1998a) "η μετρική των Chater & Hahn ... ορίζεται με βάση τη συμπεριφορά του ίδιου του συστήματος. Δεν μπορούμε να ξέρουμε ποιες είναι οι αποστάσεις στον χώρο των καταστάσεων μέχρι να μάθουμε πώς συμπεριφέρεται το σύστημα".

⁵⁴⁹ Βλ. κεφάλαιο 2, ενότητα 2.5.

υπολογιστικής θεωρίας που προσδιορίζει τη λειτουργία μιας μηχανής Turing δεν έχουν θέση ερωτήματα όπως: "Σε ποια κατάσταση ήταν η μηχανή στο χρόνο 1,5; Πόσο χρόνο ήταν η μηχανή στην κατάσταση 1; Πόσο χρόνο πήρε στη μηχανή να μεταβεί από την κατάσταση 1 στην κατάσταση 2;"⁵⁵⁰

Ένας δεύτερος τρόπος με τον οποίο ένα δυναμικό σύστημα μπορεί να είναι ποσοτικό αφορά την αλληλεξάρτηση καταστάσεων και χρόνου. Συγκεκριμένα, "ένα σύστημα είναι ποσοτικό ως προς τον χρόνο όταν ο χρόνος αποτελεί ποσότητα, δηλαδή υπάρχει μια μετρική επί του συνόλου των χρόνων τέτοια ώστε η συμπεριφορά του συστήματος να σχετίζεται συστηματικά με τις αποστάσεις που μετρώνται με αυτή τη μετρική"⁵⁵¹. Έτσι, σε ένα σύστημα που είναι ποσοτικό ως προς τον χρόνο, το ποσό μεταβολής από μια κατάσταση στην άλλη εξαρτάται συστηματικά από το ποσό του διανυόμενου χρόνου. Ξανά μπορούμε εδώ να δούμε τη διαφορά από την περίπτωση των ψηφιακών υπολογιστών. Σε μια μηχανή Turing δεν υπάρχει κανένας περιορισμός ως προς τον χρόνο που θα χρειαστεί η μηχανή για να μεταβεί από την μια κατάσταση στην άλλη⁵⁵². Κατά συνέπεια, δεν υφίσταται κάποια σχέση μεταξύ της χρονικής απόστασης δύο καταστάσεων της μηχανής και των ιδίων των καταστάσεων. Ενώ, δηλαδή, σε ένα δυναμικό σύστημα μπορούμε, γνωρίζοντας την τρέχουσα κατάσταση στην οποία βρίσκεται, να εντοπίσουμε την κατάσταση στην οποία θα βρεθεί μετά ορισμένο χρόνο t , δεν υπάρχει αντίστοιχη δυνατότητα στην περίπτωση ενός ψηφιακού υπολογιστή.

Τέλος ένας τρίτος τρόπος με τον οποίο ένα σύστημα μπορεί να είναι ποσοτικό αφορά τον ρυθμό μεταβολής του. Σε αυτά τα συστήματα οι ρυθμοί μεταβολής τους εξαρτώνται από τον εκάστοτε τρέχοντα ρυθμό μεταβολής. Ο ρυθμός μεταβολής μιας μεταβλητής εκφράζεται από την πρώτη παράγωγο της μεταβλητής ως προς τον χρόνο.

Ο van Gelder (1998), λοιπόν, εκλαμβάνει ένα σύστημα ως δυναμικό αν είναι ποσοτικό με έναν από τους παραπάνω τρεις τρόπους. Η αναφορά που κάναμε εδώ στο κριτήριο της ποσοτικότητας δεν στοχεύει να εξαντλήσει το ζήτημα, αλλά να δώσει μια πρώτη εικόνα για το τι είναι δυναμικό σύστημα και σε τι μπορεί να διαφέρει η λειτουργία του από τη λειτουργία της μηχανής Turing. Η τοποθέτηση του van Gelder δεν είναι φυσικά χωρίς προβλήματα. Υπάρχει ήδη εν εξελίξει μια συζήτηση γύρω από το εάν το κριτήριο της ποσοτικότητας είναι επαρκές για τη διάκριση

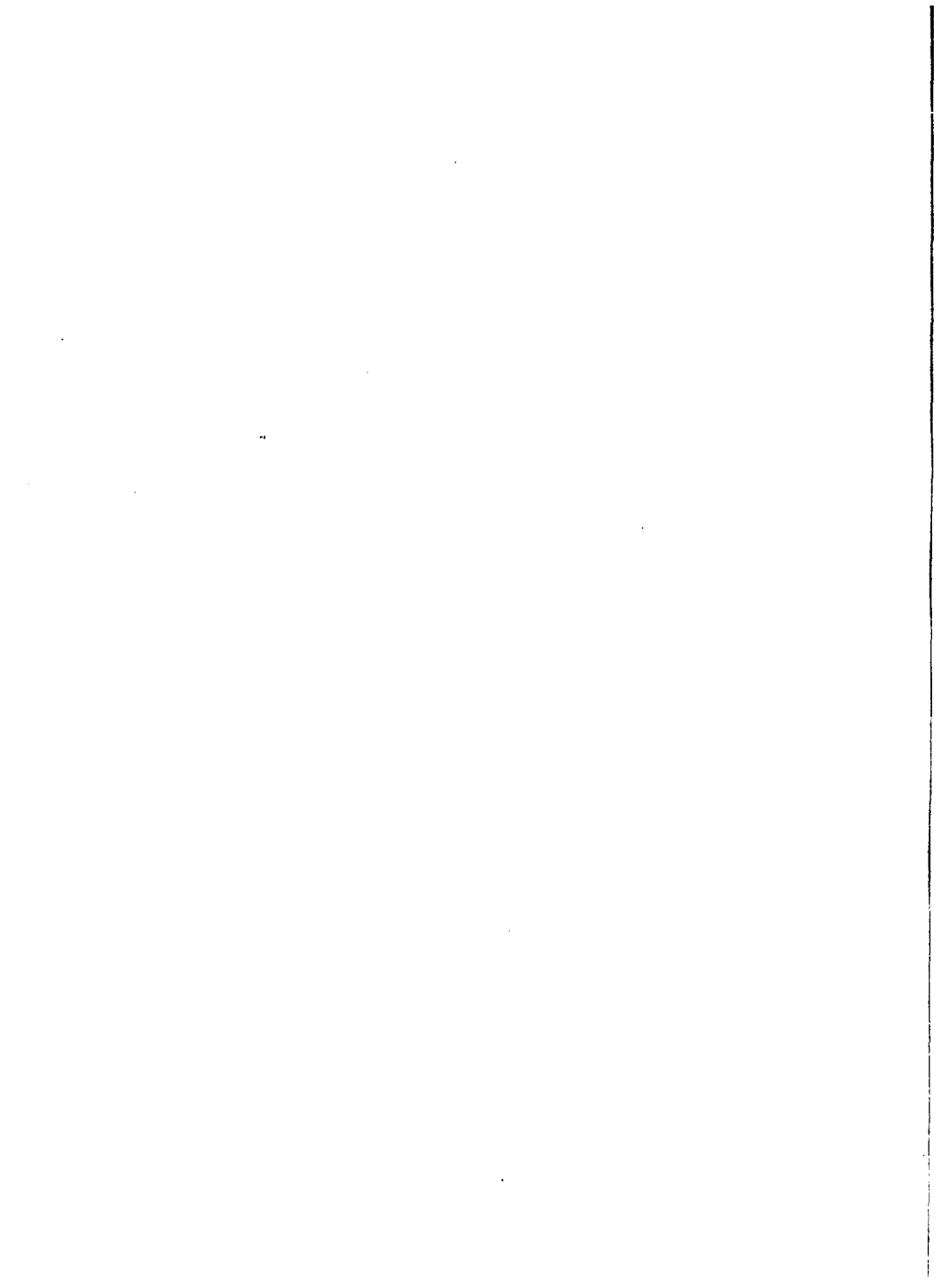
⁵⁵⁰ van Gelder & Port 1995, σελ. 20.

⁵⁵¹ van Gelder 1998.

⁵⁵² Γι' αυτό και στο υπολογιστικό ρυθμιστή που αναφέρθηκε παραπάνω πρέπει εξαγωγώς να περιοριστεί και να ελεγχθεί η χρονική διάρκεια του κύκλου λειτουργίας του.

των δυναμικών από τα υπολογιστικά συστήματα⁵⁵³. Ωστόσο, αυτό το ζήτημα υπερβαίνει το στόχο του παραρτήματος και δεν θα διερευνηθεί περαιτέρω.

⁵⁵³ Βλ., ενδεικτικά, τα σχόλια των κριτών του van Gelder (1998) στο *Behavioral and Brain Sciences*, όπου και δημοσιεύτηκε το άρθρο του van Gelder.



ΒΙΒΛΙΟΓΡΑΦΙΑ

- Anapolitanos, D. (1998): *Leibniz: Representation, Continuity and the Spatio-temporal*. Kluwer Academic Publishers.
- Andler, D. (ed.) (1992): *Introduction aux Science Cognitives*. Gallimard.
- Appiah, A. (1986): "Truth Conditions: A Causal Theory" στο Butterfield 1986.
- Aquinas, T. (1992): "That the Soul never Thinks without an Image" στο Beakley & Ludlow 1992.
- Armstrong, D. (1973): *Belief, Truth and Knowledge*. Cambridge University Press.
- Ashby, W.R. (1952): *Design for a Brain*. Chapman & Hall.
- Atneave, F. (1960): "In Defense of Homunculi" στο Rosenblith W. (ed.) (1960): *Sensory Communication*. The MIT Press.
- Austin, J. (1962): *Sense and Sensibilia*. Oxford University Press.
- Austin, J. (1995): "Επιτελεστικές Εκφορές", *Δευκαλίων*, 13/2-3, σελ. 215-230.
- Ayer, A. (1936): *Language, Truth and Logic*. Penguin Books.
- Ayer, A. (1956): *The problem of Knowledge*. Penguin Books.
- Ayer, A. (1987): "Frege, Russell, and Modern Logic" στο Magee 1987.
- Baker, R. L. (1995): *Explaining Attitudes - A Practical Approach to the Mind*. Cambridge University Press.
- Ballard, D. H. & Hayes, P. J. (1984): "Parallel Logical Inference". *Proceedings of the Sixth Conference of the Cognitive Science Society*.
- Baltas, A. (1987): "Ideological 'Assumptions' in Physics: Social Determination of Internal Structures", στο Fine, A. & Machamer, P. (1987): *PSA 1986 Vol. 2*. Philosophy of Science Association.
- Baumgartner et. al. (eds) (1995): *Phenomenology and Cognitive Science*. Röhl Verlag.
- Baumgartner, P. & Peyr, S. (eds) (1995): *Speaking Minds - Interviews with Twenty Eminent Cognitive Scientists*. Princeton University Press.
- Beakley, B. & Ludlow, P. (eds) (1992): *The Philosophy of Mind: Classical Problems/Contemporary Issues*. A Bradford Book - The MIT Press.
- Bechtel, W. & Abrahamsen, A. (1991): *Connectionism and the Mind: An Introduction to Parallel Processing in Networks*. Blackwell.
- Bechtel, W. (1996): "Yet Another Revolution? Defusing the Dynamical System Theorists' Attack on Mental Representations", Πηγή: Internet
- Bermudez, J.L. (1995): "Syntax, Semantics and Levels of Explanation". *The Philosophical Quarterly*, 45 (180) σελ. 361-367.
- Bermudez, J.L., Marcel, A., & Eilan, N. (eds) (1995): *The Body and the Self*. A Bradford Book - The MIT Press.

- Bilgrami, A. (1992): *Belief & Meaning - The Unity & Locality of Mental Content*. Blackwell.
- Black, M. (1962): *Models and Metaphors*. Cornell University Press.
- Blackburn, S. (1984): *Spreading the Word*. Oxford: Clarendon Press.
- Block, N. (1978/1980): "Troubles with Functionalism" στο Block 1980.
- Block, N. (1983): "Mental Pictures and Cognitive Science" στο Lycan 1990.
- Block, N. (1986): "Advertisement for a Semantics for Psychology" στο Stich & Warfield 1994.
- Block, N. (ed.) (1980): *Readings in Philosophy of Psychology - Volume One*. Harvard University Press.
- Block, N. (ed.) (1981): *Readings in Philosophy of Psychology - Volume Two*. Harvard University Press.
- Block, N. (υπό έκδοση): "Semantics, Conceptual Role" λήμμα στο *The Routledge Encyclopedia of Philosophy*. Πηγή: Internet.
- Bobrow, D. & Norman, D. (1975): "Some Principles of Memory Schemata" στο Bobrow & Collins 1975.
- Bobrow, D. & Winograd, T. (1977): "An Overview of KRL, a knowledge Representation Language", *Cognitive Science*, 1 (1), σελ. 3-46.
- Bobrow, G. D. & Collins, A. (eds) (1975): *Representation & Understanding - Studies in Cognitive Science*. Academic Press, Inc. .
- Boden, M. (1977): *Artificial Intelligence & Natural Man*. Basic Books, Inc., Publishers.
- Boden, M. (1988): *Computer Models of Mind*. Cambridge University Press.
- Boden, M. (ed.) (1990): *The Philosophy of Artificial Intelligence*. Oxford University Press.
- Boden, M. (ed.) (1996): *The Philosophy of Artificial Life*. Oxford Univ. Press.
- Born, R. (ed.) (1987): *Artificial Intelligence: The Case Against*. Groom Helm London & Sydney.
- Braddon-Mitchell, D. & Jackson, F. (1996): *Philosophy of Mind and Cognition*. Blackwell.
- Bradshaw, D.E. (1991): "Connectionism and the Specter of Representationalism" στο Horgan & Tienson 1991.
- Brandom, R. (1994): *Make It Explicit: Reasoning, Representing, and Discursive Commitment*. Harvard University Press.
- Brandom, R. (2000): *Articulating Reasons- An Introduction to Inferentialism*. Harvard University Press.

- Brentano, F. (1874/1973): *Psychology from an Empirical Standpoint* (translation A. Pancurello, D. T., & L. McAllister). New York: Humanities Press.
- Brooks, R. (1991a): "Intelligence without Representation", *Artificial Intelligence* 47.
- Brooks, R. (1991b): "Intelligence without Reason". *Proceedings of the 12th International Joint Conference on Artificial Intelligence*. Morgan Kaufman.
- Brown, H. (1992): "Direct Realism, Indirect Realism, and Epistemology", *Philosophy and Phenomenological Research*, Vol. LII, no. 2.
- Bruner, J. & Klein, G. (1960): "The Functions of Perceiving: New Look Retrospective" στο Kaplan, B. & Wapner, S. (eds) *Perspectives in Psychological Theory*, International Universities Press.
- Bruner, J. (1957): "On Perceptual Readiness", *Psychological Review*, 64, σελ. 123-154.
- Burge, T. (1986): "Individualism and Psychology", *Philosophical Review*, 95.
- Butterfield, J. (ed.) (1986): *Language, Mind and Logic*. Cambridge University Press.
- Campbell, J. (1994): *Past, Space, & Self*. A Bradford Book- The MIT Press.
- Campbell, J. (1999): "Schizophrenia, the Space of Reasons, and Thinking as a Motor Process", *The Monist*, vol. 82, no.4.
- Casey, E. (1983): "Keeping the Past in Mind", *Review of Metaphysics* 37, σελ. 77-95.
- Cassam, Q. (ed.) (1994): *Self-Knowledge*. Oxford University Press.
- Castaneda, H. (1966): "'He' A Study in the Logic of Self-Consciousness", *Ratio*, 8.
- Caudill M. & Butler C. (1992): *Understanding Neural Networks Computer Explorations Volume 1: Basic Networks*. A Bradford Book - The MIT Press.
- Caudill, M. & Butler, C. (1992): *Understanding Neural Networks Computer Explorations Volume 2: Advanced Networks*. A Bradford Book - The MIT Press.
- Chapman, D. (1991): *Vision, Instruction, & Action*. The MIT Press.
- Chisholm, M. R. (1979): *The First Person, An Essay on Reference & Intentionality*. Royal Institute of Philosophy.
- Chomsky, N. (1957): *Syntactic Structures*. Mouton
- Chomsky, N. (1957/1980): "A Review of B.F. Skinner's *Verbal Behavior*" στο Block 1980.

- Chrisley, R. (1993): "Connectionism, Cognitive Maps and the Development of Objectivity", *Artificial Intelligence Review*, 7, σελ. 329-354.
- Chrisley, R. (1996): *Non-Conceptual Psychological Explanation: Content & Computation*. Διδακτορική διατριβή.
- Chrucky, A. (1990): *Critique of Wilfrid Sellars' Materialism*. Διδακτορική διατριβή.
- Churchland, P. M. (1979): *Scientific Realism and the Plasticity of Mind*. Cambridge University Press.
- Churchland, P. M. (1981): "Eliminative Materialism and Propositional Attitudes", *Journal of Philosophy* 78, no. 2.
- Churchland, P. M. (1988): "Perceptual Plasticity and Theoretical Neutrality". *Philosophy of Science*, 55 (2), σελ. 167-188.
- Churchland, P. M. (1988): *Matter & Consciousness - A Contemporary Introduction to the Philosophy of Mind*. A Bradford Book - The MIT Press.
- Churchland, P. M. (1989): *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. The MIT Press.
- Churchland, P. M. (1989a): "Some Reductive Strategies in Cognitive Neurobiology" στο Churchland 1989.
- Churchland, P. M. (1995): *The Engine of Reason, the Seat of the Soul*. A Bradford Book - The MIT Press.
- Churchland, P. M. (1998): "Conceptual Similarity and Sensory/Neural Diversity", *The Journal of Philosophy*, Vol. XCV, no.1.
- Churchland, P. S. & Grush, R. (1999): "Computation and the Brain" στο Wilson & Keil 1999.
- Churchland, P.S. & Sejnowski J. T. (1992): *The Computational Brain*. A Bradford Book - The MIT Press.
- Clark, A. & Chalmers, D. J. (1998): "The extended mind", *Analysis* 58, σελ. 7-19.
- Clark, A. (1989): *Microcognition: Philosophy, Cognitive Science, & Parallel Distributed Processing*. A Bradford Book - The MIT Press.
- Clark, A. (1993): *Associative Engines: Connectionism, Concepts, & Representational Change*. A Bradford Book - The MIT Press.
- Clark, A. (1997): *Being There - Putting Brain, Body, & World Together Again*. The MIT Press.
- Clark, A. (ed.) (1994): *Synthese-Special Issue on Connectionism & the Frontiers of AI*. *Synthese* 101, σελ. 305-345.
- Cobb-Stevens R. (1990): *Husserl & Analytic Philosophy*. Kluwer Academic Publishers.
- Coffa, J. A. (1991): *The Semantic Tradition from Kant to Carnap*. Cambridge University Press.

- Copeland, J. & Sylvan R. (1999): "Beyond the Universal Turing Machine", *Australasian Journal of Philosophy*, vol. 77, σελ. 46-66.
- Copeland, J. (1993): *Artificial Intelligence - A Philosophical Introduction*. Blackwell.
- Copeland, J. (1996): "What is Computation?", *Synthese*, vol. 108, σελ.335-359.
- Copeland, J. (1997): "The Church-Turing Thesis". Stanford Encyclopedia of Philosophy. Πηγή Internet.
- Cottrell, G. & Metcalfe, J. (1991): "EMPATH: Face, Emotion, and Gender Recognition Using Holons", στο Lippman, R., Moody, J., & Touretzky, D. (eds): *Advances in Neural Information Processing Systems*, Vol. 3. San Mateo, CA: Morgan Kaufmann.
- Cottrell, G. (1991): "Extracting Features from Faces Using Compression Networks: Face, Identity, Emotions and Gender Recognition Using Holons", στο Touretzky, D., Elman, J., Sejnowski, T. & Hinton, G. (eds): *Connectionist Models: Proceedings of the 1990 Summer School*. San Mateo, CA: Morgan Kaufmann.
- Crayling, C. A. (1988): *Wittgenstein*. Oxford University Press.
- Crossley, N.J.- Ash, J.C. - Brickhill, J. C. - Stillewett C.J. - Williams H.N. (1972): *What Is Mathematical Logic*. Oxford University Press.
- Cummins, R. (1975): "Functional Explanation", *The Journal of Philosophy*, LXXII, 20, σελ. 741-764.
- Cummins, R. (1983): *The Nature of Psychological Explanation*. The MIT Press.
- Cummins, R. (1989): *Meaning & Mental Representation*. A Bradford Book - The MIT Press.
- Cummins, R. (1996): *Representation, Targets, & Attitudes*. A Bradford Book- The MIT Press.
- Cummins, R. (1996b): "Systematicity". *Journal of Philosophy* 93, σελ. 591-614.
- Cussins, A. (1990): "The Connectionist Construction of Concepts" στο Boden 1990.
- Cussins, A. (1992): "Content, Embodiment and Objectivity: The Theory of Cognitive Trails", *Mind* 101.
- Cussins, A. (2000): "Experience, Thought and Activity" ανακοίνωση στο συνέδριο με τίτλο *Experience and Knowledge*, Ρέθυμνο, Κρήτης, 26-30 Οκτωβρίου 2000.
- Cyrułnik, B. (1995): *Η Γέννηση του Νοήματος*. Εκδόσεις "Κάτοπτρο".
- Dahlbom, B. (1993): *Dennett and his Critics, Demystifying the Mind*. Blackwell.
- Damasio, A. (1994): *Descartes' Error - Emotion, Reason, & the Human Brain*. A Grosset/Putnam Book.

- Dancy, J. & Sosa, E. (eds.) (1992): *A Companion to Epistemology*. Blackwell.
- Dancy, J. (1985): *An Introduction to Contemporary Epistemology*. Blackwell.
- Davidson, D. (1980): *Essays on Actions and Events*. Oxford: Clarendon Press.
- Davidson, D. (1980a): "Mental Events" στο Davidson 1980.
- Davidson, D. (1985): "Reply to Quine on Events" στο Lepore & McLaughlin 1985.
- Davidson, D. (1989): "The Myth of the Subjective" στο Krausz, M. (1989): *Relativism: Interpretation and Confrontation*. University of Notre Dame Press.
- Davidson, D. (1994): "Knowing One's Own Mind" στο Cassam 1994.
- Davies, M. & Humphreys, W. G. (eds) (1993): *Consciousness - Readings in Mind & Language*. Blackwell.
- Davies, M. & Stone, T. (eds) (1995): *Mental Simulation*. Cambridge: Blackwell.
- Dayan, P. (1999): "Unsupervised Learning" στο Wilson. & Keil 1999.
- Dennett, D. (1969/1981): "The Nature of Images and the Introspective Trap" στο Block 1981.
- Dennett, D. (1975): "Why the Law of Effect will not Go Away" στο Lycan 1990.
- Dennett, D. (1978a): *Brainstorms*. Montgomery: Bradford Books.
- Dennett, D. (1978b): "Intentional Systems" στο Dennett 1978a.
- Dennett, D. (1987): *The Intentional Stance*. The MIT Press.
- Dennett, D. (1990): "Cognitive Wheels", στο Boden 1990.
- Dennett, D. (1991a): *Consciousness Explained*. Penguin Books.
- Dennett, D. (1991b): "Granny's Campaign for Safe Science" στο Loewer & Rey 1991.
- Descartes R. (1984): *The Philosophical Writings of Descartes*. Cambridge University Press.
- Devitt, M. (1989): "A Narrow Representational Theory of the Mind" στο Lycan 1990.
- Devitt, M. (1991): "Why Fodor Can't Have It Both Ways" στο Loewer & Rey 1991.
- Dewey J. (1922): *Human Nature and Conduct. An Introduction to Social Psychology*, London: George Allen and Unwin.
- Diederich J. (1990): *Artificial Neural Networks*. IEEE Computer Society Press.
- Dillon, C. M. (1988): *Merleau-Ponty's Ontology*. Indiana University Press.

- Donald, M. (1991): *Origins of the Modern Mind - Three Stages in the Evolution of Culture & Cognition*. Harvard University Press.
- Dretske, F. (1969): *Seeing and Knowing*, The University of Chicago Press.
- Dretske, F. (1981): *Knowledge & the Flow of Information*. The MIT Press.
- Dretske, F. (1986): "Misrepresentation" στο Lycan 1990.
- Dretske, F. (1988): *Expaining Behavior - Reasons in a World of Causes*. A Bradford Book - The MIT Press.
- Dretske, F. (1994): "Dretske" στο Guttenplan 1994.
- Dretske, F. (1995): *Naturalizing the Mind*. A Bradford Book - The MIT Press.
- Dreyfus, H. (1980): "Holism and Hermeneutics", *Review of Metaphysics* 34, σελ. 3-23.
- Dreyfus, H. (1991): *Being-In-The-World*. The MIT Press.
- Dreyfus, H. (1992): *What Computers Still Can't Do - A Critique of Artificial Reason*. The MIT Press.
- Dreyfus, H. (1996): "Response to my Critics", *Artificial Intelligence*, 80, σελ. 171-191.
- Dreyfus, H. (2001): "The Primacy of Phenomenology over Logical Analysis". Αδημοσίευτο. Πηγή: Internet.
- Dreyfus, H. (ed.) (1982): *Husserl, Intentionality and Cognitive Science*. The MIT Press.
- Dreyfus, H. and Dreyfus, S. (1986): *Mind over Machine*. Macmillan, Free Press.
- Dreyfus, H. and Dreyfus, S. (1990): "Making a Mind versus Modelling a Brain" στο Boden 1990.
- Dummett, M. (1981): *Frege: The Philosophy of Language*. Duckworth, 2nd edition.
- Dummett, M. (1993): *Origins of Analytical Philosophy*. Harvard University Press.
- Dunn, M. (1990): "Relevant Predication 2: Intrinsic Properties and Internal Relations", *Philosophical Studies* 60, σελ. 177-206.
- Edelman, M. G. (1992): *Bright Air, Brilliant Fire - On the Matter of the Mind*. Basic Books A Division of Harper Collins Publishers.
- Edwards, S. (1994): *Externalism in the Philosophy of Mind*. Avebury.
- Eliasmith, C. (1997): "The Third Contender: A Critical Examination of the Dynamicist theory of Cognition", *Journal of Philosophical Psychology*, 9 (4), σελ. 441-463.
- Erneling, E. C. (1993): *Understanding Language Acquisition - The Framework of Learning*. State University of New York Press.
- Evans, G. (1980): "Comment on Fodor", *The Behavioral and Brain Sciences*, 3.

- Evans, G. (1982): *The Varieties of Reference* (Edited by J. McDowell). Clarendon Press.
- Eysenck, M., Keane, M. (1995): *Cognitive Psychology*. Psychology Press.
- Farah, M. (1990): *Visual Agnosia - Disorders of Object Recognition and What they Tell us about Normal Vision*. A Bradford Book - The MIT Press.
- Farah, M. (1994): "Neuropsychological Inference with an Interactive Brain: A Critique of the "Locality" Assumption". *Behavioral and Brain Sciences*, 17, σελ. 43-104.
- Farah, M., Soso, M. & Dasheiff, R. (1992): "Visual Angle of the Mind's Eye before and after unilateral occipital lobectomy", *Journal of Experimental Psychology: Human Perception and Performance* 18(1), σελ. 241-246.
- Feigenbaum E. & Feldman J. (1963): *Computers and Thought*, McGraw-Hill Company.
- Feigl, H. (1958): "The 'Mental' and the 'Physical'" στο Feigl, H., Scriven, M. & Maxwell, G. (eds) (1958): *Concepts, Theories, and the Mind-Body Problem (Minnesota Studies in the Philosophy of Science 2)*. University of Minnesota Press.
- Feyerabend P. (1963): "Materialism and the Mind-Body Problem", *Review of Metaphysics* 17.
- Field, H. (1981): "Mental Representation" στο Block 1981.
- Flanagan, O. (1992): *Consciousness Reconsidered*. A Bradford Book - The MIT Press.
- Fodor, J. & Lepore E. (1992): *Holism: A Shopper's Guide*. Blackwell.
- Fodor, J. & Lepore E. (1994): "Why Meaning (Probably) Isn't Conceptual Role" στο Stich & Warfield 1994.
- Fodor, J. & McLaughlin, B. (1991): "Connectionism and the Problem of Systematicity" στο Horgan & Tienson 1991.
- Fodor, J. & Pylyshyn, Z (1988): "Connectionism and Cognitive Architecture: A Critical Analysis" στο Beakley & Ludlow 1992.
- Fodor, J. (1968): *Psychological Explanation: An Introduction to the Philosophy of Psychology*. Random House.
- Fodor, J. (1975): *The Language of Thought*. Crowell.
- Fodor, J. (1980): "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology", *Brain and Behavioral Sciences*, vol. 3, no. 1. Επανεκδόθηκε στο Fodor 1981.
- Fodor, J. (1981): *Representations - Philosophical Essays On the Foundations of Cognitive Science*. The MIT Press.
- Fodor, J. (1981a): "Propositional Attitudes" στο Fodor 1981.
- Fodor, J. (1981b): "Imagistic Representation" στο Block 1981.
- Fodor, J. (1983): *The Modularity of Mind*. The MIT Press.

- Fodor, J. (1985a): "Fodor's Guide to Mental Representation" στο Fodor 1990a.
- Fodor, J. (1985b): "Narrow Content and Meaning Holism", *αδημοσίευτο*.
- Fodor, J. (1987): *Psychosemantics - The Problem of Meaning in the Philosophy of Mind*. A Bradford Book -The MIT Press.
- Fodor, J. (1990a): *A Theory of Content & Other Essays*. A Bradford Book-The MIT Press.
- Fodor, J. (1990b): "Psychosemantics or: Where do Truth Conditions Come From?" στο Lycan 1990.
- Fodor, J. (1990c): "Observation Reconsidered" στο Fodor, J. 1990a.
- Fodor, J. (1990d): "Precis of the Modularity of Mind" στο Fodor, J. 1990a.
- Fodor, J. (1990e): "Why Should the Mind Be Modular?" στο Fodor, J. 1990a.
- Fodor, J. (1990f): "A Theory of Content, I: The Problem" στο Fodor, J. 1990a.
- Fodor, J. (1990g): "A Theory of Content, II: The Theory" στο Fodor, J. 1990a.
- Fodor, J. (1991): "Replies" στο Loewer & Rey 1991.
- Fodor, J. (1994): *The Elm & the Expert - Mentalese & Its Semantics*. A Bradford Book - The MIT Press.
- Ford, K., Glymour C. & Hayes P. (1995): *Android Epistemology*. The MIT Press.
- Frege, G. (1884/1959): *The Foundations of Arithmetic: A Logico-Mathematical Enquiry into the Concept of Number* (Translation J. L. Austin). Blackwell.
- Frege, G. (1892/1960): "On Sense and Reference" στο Geach, P. & Black, M. (eds) (1960): *Philosophical Writings of Gottlob Frege*. Blackwell.
- Frege, G. (1892/1977): "Νόημα και Αναφορά", (μετάφραση Ιόλη Πατέλη), *Δευκαλίων* 17, σελ. 19-40.
- Frege, G. (1959): *Foundations of Arithmetic*. (Translation J. Austin) Northwestern University Press
- Frege, G. (1976). "The Thought: a Logical Inquiry". (Translation A. and M. Quinton) στο P. F. Strawson (ed.) (1976): *Philosophical Logic*. Oxford University Press.
- Frege, G. (1979): *Posthumous Writings* (Μετάφραση P. Long & R. White). Blackwell.
- Frege, G. (1988). "Thoughts". (Μετάφραση P. Geach & R.H. Stoothoff). στο Frege, G. (1984): *Collected Papers on Mathematics, Logic, and Philosophy*. B. McGuinness (ed). Blackwell. Επανεέκδοση στο Salmon & Soames 1988, σελ. 33-55.
- French, A. P. (ed.) (1994): *Mindwest Studies in Philosophy Volume XIX*. University of Notre Dame Press.

- French, R. M. (1999): "Catastrophic Forgetting in Connectionist Networks". *Trends in Cognitive Sciences*, Vol. 3. No 4, 1999, σελ. 128-135.
- Frisch, K. von (1967): *The Dance Language and Orientation of Bees*. Harvard University Press
- Gasché, R. (1986): *The Tain of Mirror - Derrida and the Philosophy of Reflection*. Harvard University Press.
- Gibson, J. J. (1986): *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, Publishers.
- Gibson, J.J. (1972): "A theory of direct visual perception" στο Royce J.R. and Rozeboom W.W. (eds.) (1972): *The Psychology of Knowing*. Gordon & Breach.
- Givon, T. (1989): *Mind, Code & Context - Essays in Pragmatics*. Lawrence Erlbaum Associates, Publishers.
- Glock, H.-G. (1996): *A Wittgenstein Dictionary*. Blackwell.
- Glymour, C. (1998): "Γνωσιολογία των ανδροειδών: Υπολογισμός, Τεχνητή Νοημοσύνη και Φιλοσοφία της Επιστήμης" στο Salmon et alia 1998.
- Godfrey-Smith, P. (1989): "Misinformation", *Canadian Journal of Philosophy*, 19 (4), σελ. 535-550.
- Godfrey-Smith, P. (1994): "A Continuum of Semantic Optimism" στο Stich & Warfield 1994.
- Goldman, A. (1986): *Epistemology & Cognition*. Harvard University Press.
- Goldman, A. (1993): *Philosophical Applications of Cognitive Science*. Westview Press.
- Goldman, A. (ed.) (1993): *Readings in Philosophy & Cognitive Science*. A Bradford Book - The MIT Press.
- Gombrich, E. H.J. (1960): *Art and Illusion - A Study in the Psychology of Pictorial Representation*. Phaidon-Press
- Goodman, N. (1968): *Languages of Art - An Approach to a Theory of Symbols*. The Bobbs - Merrill Company, Inc..
- Goodman, N. (1988): "Representation Re-presented" στο Goodman & Elgin, 1988.
- Goodman, N. Elgin, C. (1988): *Reconceptions in Philosophy and Other Arts and Sciences*. Hackett.
- Gordon, J. (1985): "Dream-world or Life-world? A Phenomenological Solution to an Ancient Puzzle", *Husserl Studies* 2, σελ. 169-191.
- Graham, G. & Lynn, S. (1994): *Philosophical Psychopathology*. A Bradford Book -The MIT Press.
- Graham, G. & Lynn, S. (1994a): "Mind and Mine" στο Graham & Lynn 1994.
- Graham, G. (1993): *Philosophy of Mind: An Introduction*. Blackwell.

- Green, W. D. & Others (eds) (1996): *Cognitive Science*. Blackwell.
- Gregory, D. (1987): *Heidegger & Artificial Intelligence*. U.M.I. Dissertation Services.
- Gregory, R. (1970): *The Intelligent Eye*. Weidenfeld & Nicolson, London.
- Grene, M. (1974): *The Understanding of Nature- Essays in the Philosophy of Biology*. D. Reidel Publishing Company.
- Grice, P. (1957): "Meaning" στο Grice 1989.
- Grice, P. (1961): "The Causal Theory of Perception" στο Grice 1989.
- Grice, P. (1989): *Studies in the Way of Words*. Harvard University Press.
- Griffin, R. D. (1992): *Animal Minds*. The University of Chicago Press.
- Grush, R. (1995): *Emulation & Cognition*. Διδακτορική διατριβή.
- Grush, R. (1997): "Yet Another Design for the Brain?", *Philosophical Psychology* 10(2), σελ. 233-242.
- Guignon B. C. (1983): *Heidegger & the Problem of Knowledge*. Hackett Publishing Company.
- Guignon, B. C. (ed.) (1993): *The Cambridge Companion to Heidegger*. Cambridge University Press.
- Guttenplan, S. (ed.) (1994): *A Companion to the Philosophy of Mind*. Blackwell.
- Hacking, I. (1983): *Representing & Intervening*. Cambridge University Press.
- Hamerof R. S., Kaszniak W. A., Scott C. A. (eds) (1996): *Toward a Science of Consciousness: The First Tucson Discussions & Debates*. A Bradford Book - The MIT Press.
- Hardwick, C. (ed.) (1977): *Semiotics and Significs: The Correspondence between Charles S. Peirce and Victoria Lady Welby*. Indiana University Press.
- Harnad, S. (1990): "The Symbol Grounding Problem". *Physica D.*, 42, σελ. 335-346.
- Hartshorne, C. & Weiss, P. (eds) (1960): *Collected Papers of Charles Sanders Peirce Volume 1 Principles of Philosophy & Volume 11 Elements of Logic*. The Belknap Press of Harvard University Press.
- Haugeland, J. (1981): "Analog and Analog" στο *Philosophical Topics*, Vol. 12, no 1.
- Haugeland, J. (1982/1998): "Weak Supervenience" στο Haugeland 1998.
- Haugeland, J. (1985): *Artificial Intelligence, the Very Idea*. A Bradford Book - The MIT Press.
- Haugeland, J. (1985/1992): *Τεχνητή Νοημοσύνη - Σχεδιάζοντας τη Νόηση: Από την Υπολογιστική Θεωρία στις Σύγχρονες Ευφυείς Μηχανές*. Εκδόσεις "Κάτοπτρο".
- Haugeland, J. (1991): "Representation Genera" στο Haugeland 1998.

- Haugeland, J. (1995): "Mind Embedded and Embodied" στο Haugeland 1998.
- Haugeland, J. (1998): *Having Thought*. Harvard University Press.
- Haugeland, J. (1998a): "Truth and Rule-following" στο Haugeland 1998.
- Heidegger, M. (1927/1962): *Being & Time*. Harper San Francisco- A Division of Harper Collins Publishers.
- Heidegger, M. (1977): *The Question Concerning Technology*. Harper Colophon Books.
- Heidegger, M. (1979): *History of the Concept of Time*. Indiana University Press Bloomington.
- Heidegger, M. (1982): *The Basic Problems of Phenomenology* (Translation Hofstadter, A.). Indiana University Press.
- Heidegger, M. (1985): *What is a Thing?* (Translation Barton Jr, W.B. & Deutsch, V.). University Press of America.
- Heidegger, M. (1999): *Τα Βασικά Προβλήματα της Φαινομενολογίας*. (Μετάφραση Π. Κόντος). Εκδόσεις "Γνώση".
- Heil J. & Mele A. (eds) (1995): *Mental Causation*. Clarendon Press-Oxford.
- Heil, J. (1993): *The Nature of True Minds*. Cambridge Studies in Philosophy.
- Heil, J. (1998): *Philosophy of Mind*. Routledge.
- Held, R. & Hein, A. (1963): "Movement produced stimulation in the development of visually guided behavior". *Journal of Comparative and Physiological Psychology*, 56, σελ. 873-876.
- Hempel C.G. (1935/1980): "The Logical Analysis in Psychology" στο Block 1980.
- Hempel, C. (1950): "Problems and Changes in the Empiricist Criterion of Meaning", *Review Internationale de Philosophie* 11, σελ. 41-63.
- Hendriks-Jansen, H. (1996): *Catching Ourselves in the Act - Situated Activity, Interactive Emergence, Evolution, and Human Thought*. The MIT Press.
- Henrich, D. (1971): "Self-Consciousness: A Critical Introduction to a Theory" στο *Man and World* 4, σελ. 3-28.
- Henrich, D. (1982): "Fichte's Original Insight" in *Contemporary German Philosophy*, vol.1, Christensen D.E. et al. (eds), University Park: Pennsylvania State University Press, 1982, pp. 15-53.
- Hinton, G. & Sejnowski, T (1986): "Learning and Relearning in Boltzmann Machines" στο Rumelhart, McClelland and the PDP Research Group 1986, κεφάλαιο 7.
- Hinton, G., McClelland, J. & Rumelhart, D. (1986): "Distributed Representations" στο Rumelhart, McClelland and the PDP Research Group 1986. Επανεκδοση στο Boden 1990.

- Hobson, P. (1993): "Through feeling and sight to self and symbol" στο Neisser 1993.
- Hofstadter, R. D. & Dennett, C. D. (επιμέλεια) (1993): *Το Εγώ της Νόησης - Φαντασίες και Στοχασμοί για τον Εαυτό και την Ψυχή*. Κατοπτρο.
- Hookway, C. & Peterson D. (1993): *Philosophy & Cognitive Science*. Cambridge University Press.
- Hopfield, J.J. (1982) "Neural Networks and Physical Systems with emergent collective computational abilities", *Proceedings of the National Academy of Sciences*, 79, σελ. 2554-8.
- Hopkins, R. (1995): "Explaining Depiction", *Philosophical-Review*, 104(3)
- Horgan, T. & Tienson, J. (eds) (1991): *Connectionism and the Philosophy of Mind*. Kluwer Academic Publishers.
- Hornsby, J. (1985): "Physicalism, Events and Part-Whole Relations" στο Lepore & McLaughlin 1985.
- Humphrey, N. (1992): *A History of the Mind*. Vintage.
- Husserl, E. (1900/1970): *Logical Investigations* (Trans. J. N. Findlay). (Δύο τόμοι), London: Routledge & Kegan Paul.
- Husserl, E. (1983): *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy* (First Book) (Translation Kersten, Fred). Nijhoff
- Husserl, Edmund (1954/1970): *The Crisis of European Sciences & Transcendental Phenomenology*. Northwestern University Press.
- Hyman, J. (1997): "Words and Pictures" στο Preston 1997.
- Jackendoff, R. (1983): *Semantics & Cognition*. The MIT Press.
- Jackson, F. (1977): *Perception*. Cambridge University Press.
- Jackson, F. (1982): "Epiphenomenal Qualia", *Philosophical Quarterly*, 32.
- Jacob, P. (1997): *What Minds Can Do: Intentionality in a non-Intentional World*. Cambridge University Press.
- Jacob, P. (1999): "Can Selection Explain Content?". Πηγή Internet.
- James, W. (1892/1961) *Psychology: The Briefer Course*. Notre Dame: University of Notre Dame Press.
- Johnson, M. (1987): *The Body in the Mind - The Bodily Basis of Meaning, Imagination, & Reason*. The University of Chicago Press.
- Johnson-Laird, P. (1983): *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge University Press.
- Johnson-Laird, P. (1988): *The Computer and the Mind*, Harvard University Press.
- Johnson-Laird, P. (1989): "Mental Models" στο Posner 1989.
- Judge, B. (1983): *Thinking about Things: A Philosophical Study of Representation*. Scottish Academic Press.

- Kalish, D & Montague, R. (1964): *Logic Techniques of Formal Reasoning*. Harcourt, Brace & World, Inc..
- Kant, I. (1929): *Critique of Pure Reason*. Macmillan Education.
- Kant, I. (1956): *Critique of Practical Reason*. The Bobbs-Merrill Company, Inc..
- Kant, I. (1983): *Προλεγόμενα σε κάθε Μελλοντική Μεταφυσική*. (Εισαγωγή, μετάφραση, σχόλια: Γ. Τζαβάρας). Εκδόσεις "Δωδώνη".
- Kant, I. (1983α): *Κριτική του Καθαρού Λόγου - Σύστημα Υπερβατικών Ιδεών, Παραλογισμοί Υπερβατικής Ψυχολογίας*, Τεύχος Πρώτο. (Μετάφραση Μ. Δημητρακόπουλος).
- Kant, I. (1985): *Κριτική του Καθαρού Λόγου - Η Αντινομία του Καθαρού Λόγου*, Τεύχος Δεύτερο. (Μετάφραση Μ. Δημητρακόπουλος).
- Kant, I. (1986): *Κριτική του Καθαρού Λόγου - Η Υπερβατική Θεολογία*, Τεύχος Τρίτο. (Μετάφραση Μ. Δημητρακόπουλος).
- Kant, I. (1987): *Κριτική του Καθαρού Λόγου - Η Υπερβατική Μεθοδολογία*, Τεύχος Τέταρτο. (Μετάφραση Μ. Δημητρακόπουλος).
- Kant, I. (υπό έκδοση): *Κριτική της Κριτικής Δύναμης*. Επιμέλεια: Γ. Ξηροπαίδης, Κ. Ψυχοπαίδης. Εκδόσεις Πόλις.
- Kim, J. (1993): *Supervenience and Mind*. Cambridge University Press.
- Kim, J. (1996): *The Philosophy of Mind*. Westview.
- Kohonen, T. (1984): *Self-Organization and Associative Memory*. New York: Springer-Verlag.
- Kolak, D. & Martin, R. (eds.) (1991): *Self & Identity*. Macmillan Publishing Company.
- Kosslyn M. S. (1980): *Image & Mind*. Harvard University Press.
- Kosslyn M. S. (1981): "The Medium and the Message in Mental Imagery: A Theory". *Psychological Review* 88(1), σελ. 46-66.
- Kosslyn M. S. (1983): "Mental Representation" στο Anderson, J. & Kosslyn, S. (eds.) (1984): *Tutorials in Learning and Memory, Essays in Honor of Gordon Bower*. W.H. Freeman.
- Kosslyn M. S. (1983): *Ghost's in the Minds Machine*. Norton.
- Kosslyn M. S. (1990): "Mental Imagery" στο Osherson, Kosslyn & Hollerbach 1990.
- Kosslyn S. (1994): *Image & Brain: The Resolution of the Imagery Debate*. A Bradford Book - The MIT Press.
- Kosslyn, S. & Hatfield, G. (1984): "Representation without Symbol Systems", *Social Research*, 51, σελ. 1019-1045.
- Kosslyn, S., & Koenig, Olivier (1992): *Wet Mind - The New Cognitive Neuroscience*. The Free Press.
- Kosslyn, S., Ball, T. & Reiser, B. (1978): "Visual Images Preserve Metric Spatial Information: Evidence from Studies of Mental Scanning",

- Journal of Experimental Psychology: Human Perception and Performance* 4, σελ. 47-60.
- Kripke, S. (1972): "Naming and Necessity", στο Davidson, D. & Harman, G. (1972): *Semantics of Natural Language*. Dordrecht: Reidel, σελ. 253-355.
- Kuhn, T. (1981): *Η Δομή των Επιστημονικών Επαναστάσεων* (Εισαγωγή, επιμέλεια Β. Κάλφας, Μετάφραση Γ. Γεωργακόπουλος, Β. Κάλφας). Σύγχρονα Θέματα.
- Lacan, J. (1949/1984): "Το στάδιο του καθρέφτη σαν διαμορφωτής της λειτουργίας του εγώ" στο Μπάλλας 1984.
- Lakoff, G. (1987): *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. The University of Chicago Press.
- Langton, G. C. (1995): *Artificial Life: An Overview*. The MIT Press.
- Lashley, K. S. (1963 /1929): *Brain mechanisms and intelligence: A quantitative study of injuries to the brain*. Dover.
- Leiber, J. (1991): *An Introduction to Cognitive Science*. Blackwell.
- Leibniz, G.W. (1985): *The Monadology and Other Philosophical Writings* (Translator, editor: Latta, R.). Garland
- Leidlmair, K. (1991): *Künstliche Intelligenz und Heidegger - Über den Zwiespalt von Natur und Geist*. Wilhelm Fink Verlag.
- Lenat, D. and Feigenbaum, E. (1991): "On the thresholds of knowledge", *Artificial Intelligence*, 47 (1-3), σελ. 185-230.
- Lennon, K. (1990): *Explaining Human Action*. Open Court.
- Lepore, E. & Gulick V. R. (eds) (1991): *John Searle & his Critics*. Blackwell.
- Lepore, E. & McLaughlin, B. (eds.) (1985): *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Blackwell.
- Lewis, D. (1972): "Psychophysical and Theoretical Identifications", *Australasian Journal of Philosophy* 50, επανεκδόθηκε στο Block 1980.
- Lewis, D. (1983): "Extrinsic Properties," *Philosophical Studies* 44, σελ. 197-200.
- Loar, B (1981): *Mind and Meaning*. UK: Cambridge University Press.
- Locke, J. (1984): *Essay Concerning Human Understanding*. (Frazer, A. (ed.)). Oxford University Press.
- Loewer, B. & Rey, G. (eds) (1991): *Meaning in Mind: Fodor and his Critics*. Blackwell.
- Lormand, E. (1999): "Frame Problem" στο Wilson & Keil 1999.
- Lycan, W. (1980): "The Functionalist Reply". *Behavioral and Brain Sciences* 3, σελ. 434-435.
- Lycan, W. (1987): *Consciousness*. A Bradford Book - The MIT Press.
- Lycan, W. (ed.) (1990): *Mind & Cognition*. Blackwell.

- Lycan, W. (1999): *Philosophy of Language: A Contemporary Introduction*. Routledge.
- Lyons, W. (1986): *The Disappearance of Introspection*. The MIT Press.
- Lyons, W. (1995): *Approaches to Intentionality*. Clarendon Press.
- Machamer, P. (1998): "Φιλοσοφία της Ψυχολογίας" στο Salmon, H. M. et alia 1998.
- Magee, B. (ed.) (1987): *The Great Philosophers - An Introduction to Western Philosophy*. Oxford University Press.
- Mallin, B. S. (1979): *Merleau-Ponty's Philosophy*. New Haven & London Yale University Press.
- Maloney, J. C. (1989): *The Mundane Matter of the Mental Language*. Cambridge University Press.
- Marbach, E. (1993): *Mental Representation and Consciousness: Towards a Phenomenological Theory of Representation and Reference*. Kluwer Academic Publishers.
- Marr, D. (1982): *Vision*. W.H. Freeman & Co.
- McClamrock, R. (1995): *Existential Cognition - Computational Minds in the World*. The University of Chicago Press.
- McClelland L. J. & Rumelhart E. D. (eds) (1988): *Explorations in Parallel Distributed Processing*. A Bradford Book - The MIT Press.
- McClelland, J. & Kawamoto, A. (1986): "Mechanisms of Sentence Processing: Assigning Roles to Constituents" στο McClelland, Rumelhart & the PDP Research Group 1986b.
- McClelland, L. J., Rumelhart, E. D. & Hinton, G. (1986): "The Appeal of PDP" στο Rumelhart, McClelland, and the PDP Research Group 1986, σελ. 3-44.
- McClelland, L. J., Rumelhart, E. D. & the PDP Research Group (1986): *Parallel Distributed Processing, Explorations in the Microstructure of Cognition, Volume 2: Psychological & Biological Models*. The MIT Press.
- McCulloch, G. (1995): *The Mind & its World*. London: Routledge.
- McCulloch, W. & Pitts, W. (1943): "A Logical Calculus of the Ideas Immanent in Nervous Activity" στο Boden 1990.
- McDowell, J. (1977): "On the Sense and Reference of a Proper Name", *Mind*, 86, σελ. 159-184.
- McDowell, J. (1983): "Criteria, Defeasibility, and Knowledge", *Proceedings of the British Academy*, LXVIII, σελ. 455-479.
- McDowell, J. (1984): "De Re Senses", *Philosophical Quarterly*, 34.
- McDowell, J. (1986): "Singular Thought and the Extend of Inner Space" στο Pettit & McDowell 1986.
- McDowell, J. (1992): "Putnam on Mind and Meaning", *Philosophical*

- McDowell, J. (1994): *Mind & World*. Harvard University Press.
- McDowell, J. (1994a): "The content of perceptual experience" *The Philosophical Quarterly*, 44, σελ. 190-205.
- McDowell, J. (1995): "Ο Wittgenstein περί του να ακολουθεί κανείς έναν κανόνα", *Δευκαλιών* 13, τεύχος 2-3, σελ. 175-214.
- McDowell, J. (2000): "Transcendental Empiricism", ανακοίνωση στο συνέδριο με τίτλο *Experience and Knowledge*, Ρέθυμνο, Κρήτης, 26-30 Οκτωβρίου 2000.
- McGinn C. (1991): *The Problem of Consciousness: Essays Towards a Resolution*. Blackwell.
- McGinn, Colin (1989): *Mental Content*. Blackwell
- McLeod, P., Plunkett, K. & Rolls, E. (1998): *Introduction to Connectionist Modelling of Cognitive Processes*. Oxford University Press.
- McRae, R. (1965): "'Idea' as a Philosophical Term in the Seventeenth Century", *Journal of the History of Ideas*, 26.
- Mead, G. H. (1982): *The Individual & the Social Self - Unpublished Work of George Herbert Mead*. The University of Chicago Press.
- Mellor, D.H. (1988): "I and Now", *Proceedings of the Aristotelian Society*, 89, σελ. 79-94.
- Merleau-Ponty, M. (1945): *Phenomenologie de la Perception*. Gallimard.
- Merleau-Ponty, M. (1945/1962): *Phenomenology of Perception*. Routledge
- Merleau-Ponty, M. (1964): *Signs*. Northwestern University Press.
- Merleau-Ponty, M. (1964): *The Primacy of Perception & Other Essays: On Phenomenological Psychology, The Philosophy of Art, History & Politics*. Northwestern University Press.
- Merleau-Ponty, M. (1973): *Consciousness and the Acquisition of Language*, (Translation H. Silverman). Northwestern University Press.
- Merleau-Ponty, M. (1977): Προοίμιο στην "Φαινομενολογία της Αντίληψης". (Μετάφραση Φ. Καλλίας). Εκδόσεις Έρασμος.
- Metzinger T. (ed.) (1995): *Conscious Experience*. Imprint Academic Schoningh.
- Michaels, C.F. & Carello, C. (1981): *Direct Perception*. Prentice Hall.
- Miller, A. (1998): *Philosophy of Language*. UCL Press.
- Millikan, G. R. (1984): *Language, Thought, & Other Biological Categories - New Foundations For Realism*. A Bradford Book - The MIT Press.
- Millikan, G. R. (1993a): *White Queen Psychology and Other Essays for Alice*. A Bradford Book - The MIT Press.
- Millikan, G. R. (1993b): "On Mentalese Orthography" στο Millikan 1993a.
- Millikan, G. R. (1993c): "Biosemantics" στο Millikan 1993a.

- Millikan, G. R. (1993d): "Naturalist Reflections on Knowledge" στο Millikan 1993a.
- Millikan, G. R. (1993e): "Thoughts without Laws" στο Millikan 1993a.
- Millikan, G. R. (1996): "Pushmi-pullyu Representations", Πηγή Internet. Το άρθρο εκδόθηκε στο J. Tomberlin, ed. (1996): *Philosophical Perspectives* vol. IX, Ridgeview Publishing, pp. 185-200.
- Millikan, G. R. (1998): "Wings, Spoons, Pills and Quills; A Pluralist Theory of Functions," *The Journal of Philosophy* 96, 4, σελ.191-206.
- Millikan, G. R. (1999): "The Language-Thought Partnership: A Bird's Eye View", Πηγή Internet. Το άρθρο εκδόθηκε στο *Language and Communication* - special issue on Language and Intentionality, H. J. Glock (ed.). 1999.
- Minsky, M. (1975): "A Framework for Representing Knowledge" στο Winston, P. (ed.) (1975): *The Psychology of Computer Vision*. McGraw-Hill.
- Minsky, M.A. & Pappert, S. (1969): *Perceptrons*. The MIT Press.
- Mitchell, W.J.T. (1994): *Picture Theory*. The University of Chicago Press.
- Moore, G.E. (1913): "The Status of Sense-Data", *Proceedings of the Aristotelian Society* 14, σελ. 355-381.
- Moore, J. & Newell, A. (1973): "How can MERLIN understand?" στο Gregg, L. (ed.): *Knowledge and Cognition*. Lawrence Erlbaum Associates.
- Moravia, S. (1986): *The Enigma of the Mind - The Mind-Body Problem in Contemporary Thought*. Cambridge University Press.
- Morick, H. (ed.) (1967): *Wittgenstein and the Problem of Other Minds*. McGraw-Hill.
- Morris M.G.R. (1989): *Parallel Distributed Processing-Implications for Psychology & Neurobiology*. Clarendon Press.
- Morris, C. (1938): "Foundation of the Theory of Signs". *International Encyclopaedia of Unified Science* 1-2. The University of Chicago Press.
- Morris, C. (1946): *Signs, Language and Behavior*. Prentice-Hall.
- Mulhall, S. (1990): *On Being in the World - Wittgenstein & Heidegger on Seeing Aspects*. Routledge.
- Mulhall, S. (1996): *Heidegger & Being & Time*. Routledge.
- Murdock, B. B. (1979): "Convolution and correlation in perception and memory". In L. G. Nilsson (Ed.), *Perspectives on Memory Research*. Hillsdale NJ: Lawrence Erlbaum Associates, σελ. 609-626.
- Nagel, E., Newman R.J. (1958): *Το Θεώρημα του Godel*. Τροχάλια.
- Nagel, T. (1974): "What it is like to be a Bat?", *Philosophical Review*, 83.
- Nagel, T. (1986): *The View From Nowhere*. Oxford University Press .

- Nazir, T. A., & O'Regan, J. K. (1990): "Some Results on Translation Invariance in the Human Visual System", *Spatial Vision*, 5(2), 81-100.
- Neander, K. (1995): "Misrepresenting and Malfunctioning", *Philosophical Studies*, 79, no.2, σελ. 109-141.
- Neisser, U. (ed.) (1993): *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*. Cambridge University Press.
- Nemirow L.: "Physicalism and the Cognitive Role of Acquaintance" στο Lycan 1990.
- Newell, A. & Simon, H. (1972): *Human Problem Solving*. Prentice-Hall.
- Nichols, R. (1995): *Classical Cryptography Course*. Aegean Park Press. Πηγή Internet.
- Nicolacopoulos, P. (ed.) (1990): *Greek Studies in the Philosophy and History of Science*. Kluwer Academic Publishers.
- Noeth, W. (1995): *Handbook of Semiotics*. Indiana University Press.
- Nozick, R. (1981): *Philosophical Explanations*. Harvard University Press.
- Nunn, C. (1996): *Awareness. What it Is - What it Does*. Routledge.
- O'Callaghan, J. (1997): "The Problem of Language and Mental Representation in Aristotle and St. Thomas", *The Review of Metaphysics* 50, 499-545.
- Olafson, A. F. (1987): *Heidegger & the Philosophy of Mind*. Yale University Press.
- O'Regan, J. K. & Noë, A. (2001): "A sensorimotor Account of Vision and Visual Consciousness", *Behavioral and Brain Sciences* 24 (5)
- Osherson N. D. & Lasnik H. (eds) (1990): *Language An Invitation to Cognitive Science Volume 1*. The MIT Press.
- Osherson N. D., Kosslyn M. S., & Hollerbach M. Z. (eds) (1990): *Visual Cognition & Action: An Invitation to Cognitive Science, Volume 2*. The MIT Press.
- Osherson N.D. & Smith E. E. (eds) (1990): *Thinking- An Invitation to Cognitive Science Volume 3*. The MIT Press.
- Otto, R. H. & Tuedio, A. J. (eds) (1987): *Perspectives on Mind*. D. Reidel Publishing Company.
- Paivio, A. (1971): *Imagery and Verbal Processes*. Holt, Rinehart & Winston.
- Papert, S. & Goldstein, I. (1975): "Artificial Intelligence, Language and the Study of Knowledge", AI Memo No. 337
- Papineau, D. (1987): *Reality and Representation*. Blackwell.
- Papineau, D. (1993): *Philosophical Naturalism*. Blackwell
- Papineau, D. (1994): "Representation and Explanation", *Philosophy of Science*, 51, σελ. 550-572.

- Papineau, D. (1998): "Teleosemantics and Indeterminacy". *Australasian Journal of Philosophy*, Vol. 76, No 1, σελ. 1-14.
- Peacocke, C. (1986): "Analogue Content", *Aristotelian Society*, Suppl. 60, σελ. 1-18.
- Peacocke, C. (1986): *Thoughts: An Essay on Content*. Blackwell.
- Peacocke, C. (1992): *A Study of Concepts*. A Bradford Book - The MIT Press.
- Peirce, C.S. (1967): *Collected Papers*. (C. Hartshorne and P. Weiss (eds)). Harvard University Press.
- Penrose, R. (1989): *The Emperor's New Mind*. Oxford University Press.
- Perry, J. (1979): "The Problem of Essential Indexicals", *Nous*, 13.
- Pettit, P. & McDowell, J. (eds.) (1986): *Subject, Thought and Context*. Clarendon Press.
- Piaget, J. (1946/1962): *Play, Dreams, and Imitation in Childhood*. New York: Norton.
- Place U. T. (1956): "Is Consciousness a Brain Process?", *British Journal of Psychology*, 47.
- Place, U. T. (2000): "The Role of the Hand in the Evolution of Language: Target Article on Language Origins". *Psychology* 11 (007)
- Plate, T. A. (1993): "Holographic Recurrent Networks". in C. L. Giles, S. J. Hanson and J. D. Cowan (eds.), *Advances in Neural Processing Systems 5 (NIPS92)*. Morgan Kaufmann.
- Pollack, J. (1990): "Recursive Distributed Representations". *Artificial Intelligence* 46, σελ. 77-105.
- Port, R. & van Gelder, T. (eds) (1995): *Mind as Motion: Explorations in the Dynamics of Cognition*. The MIT Press.
- Posner, I. M. (ed.) (1989): *Foundations of Cognitive Science*. A Bradford Book The MIT Press.
- Preston, E. (1988): *Representational & Non-Representational Intentionality Husserl, Heidegger & Artificial Intelligence*. UMI Dissertation.
- Preston, E. (1994): "Behaviorism and Mentalism: Is There a Third Alternative?", *Synthese*, 100, 5, σελ. 167-196.
- Preston, E. (1998): "A Pluralist Theory of Function", *The Journal of Philosophy*, XCV, 5, σελ. 215-254.
- Preston, J. (ed.) (1997): *Thought and Language*. Cambridge University Press.
- Priest, S. (1991): *Theories of the Mind*. Penguin Books.
- Prieto, L. (1986): *Μηνύματα και Σήματα*. Εκδόσεις "Νεφέλη".

- Putnam H. (1967): "The mental Life of Some Machines", H.N.Castaneda (ed), *Intentionality, Minds, and Perception*, Wayne State University Press.
- Putnam H. (1967/1980): "The Nature of Mental States" στο Block 1980.
- Putnam, H. (1975a): *Mind Language & Reality*. Cambridge University Press.
- Putnam, H. (1975b): "The Meaning of 'Meaning'" στο Putnam 1975a.
- Putnam, H. (1978): *Meaning & the Moral Sciences*. Routledge & Kegan Paul.
- Putnam, H. (1981): *Reason, Truth, & History*. Cambridge University Press.
- Putnam, H. (1981a): "Brains in a Vat" στο Putnam 1981.
- Putnam, H. (1988): *Representation & Reality*. A Bradford Book. The MIT Press.
- Putnam, H. (1994): *Words & Life*. Harvard University Press.
- Putnam, H. (1994a): "Sense, Nonsense, and the Senses: An Inquiry into the Powers of the Human Mind", *Journal of Philosophy*, 91, σελ. 445-517.
- Pylyshyn Z. (1978): "Imagery and Artificial Intelligence" στο Block 1981.
- Pylyshyn Z. (1984): *Computation & Cognition: Towards a Foundation for Cognitive Science*. The MIT Press.
- Pylyshyn, Z. (1981): "The Imagery Debate: Analogue Media versus Tacit Knowledge", *Psychological Review*, 88, σελ. 16-45.
- Pylyshyn, Z. (2001): "Mental Imagery: In search of a theory", *Brain and Behavioral Sciences*, υπό δημοσίευση.
- Quine, W.V O. (1985): "Events and Reification" στο Lepore & McLaughlin 1985.
- Quine, W.V.O. (1951): "Two dogmas of empiricism" στο Quine 1953.
- Quine, W.V.O. (1953): *From a Logical Point of View*. Harvard University Press.
- Quine, W.V.O. (1960): *Word & Object*. The MIT Press.
- Ramsey, F. P. (1927): "Facts and Propositions" στο Mellor, D. (1990): *Philosophical Papers: P.F. Ramsey*. Cambridge University Press.
- Ramsey, W., Stich, S.P., Rumelhart D. (eds) (1991): *Philosophy and Connectionist Theory*. Erlbaum.
- Recanati, F. (1993): *Direct Reference: From Language to Thought*. Blackwell.
- Reid, T. (1785/1969): *Essays on the Intellectual Powers of Man*. Baruch, Brody (ed.). The MIT Press.
- Reisberg, D. & Chambers, D. (1991): "Neither Pictures nor Propositions: What Can we Learn from a Mental Image?", *Canadian Journal of Psychology*, 45(3), σελ. 336-352.

- Rich, E. & Knight, K. (1991): *Artificial Intelligence*. International edition: McGraw-Hill.
- Richardson, J. (1986): *Existential Epistemology*. Clarendon Press.
- Ricoeur, P. (1986): *Περί Ερμηνείας* (Μετάφραση Σ. Ροζάνη). Εκδόσεις Ερασμός.
- Ricoeur, P. (1996): *Η Ζωντανή Μεταφορά* (Μετάφραση Κ. Παπαγιώργης). Εκδόσεις Κριτική.
- Rorty R. (1965): "Mind-Body Identity, Privacy and Categories", *Review of Metaphysics* 17.
- Rorty, O. A. (1975): *The Identities of Persons*. University of California Press.
- Rosenfield, I. (1992): *The Strange, the Familiar, and Forgotten: An Anatomy of Consciousness*. Picador.
- Rosenthal, D. M. (1986): "Two concepts of consciousness", *Philosophical Studies* 49, σελ. 329-59.
- Rumelhart D., Smolensky P., McClelland, J. & Hinton, G. (1986): "Schemata and Sequential Thought Processes in PDP Models" στο McClelland, Rumelhart & the PDP Research Group 1986b.
- Rumelhart, D., McClelland, J. and the PDP Research Group (1986): *Parallel Distributed Processing: Explorations in the Microstructures of Cognition. Volume 1: Foundations*. The MIT Press.
- Ryle G. (1957): "The Theory of Meaning" στο Mace C.A. (ed.) (1957): *British Philosophy in Mid Century*. Allen and Unwin..
- Ryle, G. (1949): *The Concept of Mind*. Hutchinson's University Library.
- Ryle, G. (1971). "A Puzzling Element in the Notion of Thinking", *Collected Papers vol.II: Collected Essays 1929-1968*. Hutchinson.
- Rynasiewicz, R. (1990): *An Introduction to Symbolic Logic*. Ανέκδοτες σημειώσεις.
- Sacks, O. (1990): *Ο Άνθρωπος που Μπέρδεψε τη Γυναίκα του με ένα Καπέλο*. (Μετάφραση Κ. Πόταγας). Εκδόσεις Καστανιώτη.
- Sacks, O. (1995): *An Anthropologist on Mars*. A. Knopf.
- Sacks, O. (1995a): "To See and not to See" στο Sacks 1995.
- Salmon, H. M. et alia (επιμέλεια) (1998): *Εισαγωγή στη Φιλοσοφία της Επιστήμης*. Πανεπιστημιακές Εκδόσεις Κρήτης.
- Salmon, N. & Soames, S. (eds) (1988): *Propositions & Attitudes*. Oxford University Press.
- Saussure, F. (1979): *Μαθήματα Γενικής Γλωσσολογίας* (Μετάφραση, σχόλια προλογικό σημείωμα: Φ.Δ. Αποστολόπουλος). Εκδόσεις Παπαζήση.
- Schank, R. & Abelson, R. P. (1977): *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum Associates.

- Schank, R. (1973): "Identification of conceptualizations underlying natural language" στο Schank, R. & Colby, K. (eds) (1973): *Computer Models of Thought and Language*. Freeman.
- Schier, F. (1986): *Deeper into Pictures*. Cambridge University Press.
- Schiffer, S. (1987): *Remnants of Meaning*. A Bradford Book - The MIT Press.
- Searle J. (1980): "Minds, Brains, and Programs", *The Behavioral and Brain Sciences* 3.
- Searle, J. (1983): *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press.
- Searle, J. (1984/1994): *Νους, Εγκέφαλος και Επιστήμη*. Πανεπιστημιακές Εκδόσεις Κρήτης.
- Searle, J. (1992): *The Rediscovery of the Mind*. The MIT Press.
- Searle, J. (1995): *The Construction of Social Reality*. The Free Press.
- Sebeok, T. (1994): *Sings: An Introduction to Semiotics*. University of Toronto Press
- Sellars, W. (1963): *Science, Perception, and Reality*. Humanities Press.
- Sellars, W. (1963a): "Some Reflections on Language Games", στο Sellars 1963.
- Sellars, W. (1980): "Inference and Meaning" στο Sicha 1980.
- Shallice, T. (1988): *From Neuropsychology to Mental Structure*. Cambridge University Press.
- Shanker, G.S. (1988): *Wittgenstein's Remarks on the Significance of Godel's Theorem*. Routledge.
- Shanon, B. (1993): *The Representational & the Presentational - An Essay on Cognition & the Study of Mind*. Harvester Wheatsheaf.
- Shepard, R. & Metzler, J. (1971/1992): "Mental Rotation of Three-Dimensional Objects" στο Beakley & Ludlow 1992.
- Shoemaker, S. (1963): *Self-Knowledge & Self-Identity*. Cornell University Press.
- Shoemaker, S. (1963/1994): "Self-Reference and Self-Awareness" στο Cassam 1994.
- Shoemaker, S. (1984): *Identity, Cause, & Mind: Philosophical Essays*. Cambridge University Press.
- Shoemaker, S. (1986/1994): "Introspection and the Self" στο Cassam 1994.
- Shoemaker, S. (1992): "Self-Knowledge and Self-Identity" στο Dancy & Sosa 1992.
- Shoemaker, S. (1994): *The First-Person Perspective and Other Essays*. Cambridge University Press
- Sicha, J. (ed.) (1980): *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfred Sellars*. Ridgeview Publishing.

- Sider, T. (υπό έκδοση): "Intrinsic Properties", *Philosophical Studies*.
- Simon, H. (1969): *The Sciences of the Artificial*. The MIT Press.
- Simons, P.M. (1987): *Parts: A Study in Ontology*. Clarendon Press.
- Skinner, B. F. (1957): *Verbal Behavior*. Englewood Cliffs, Prentice-Hall.
- Skinner, B. F. (1974): *About Behaviorism*. A. Knopf.
- Sloman, A. (1978): *The Computer Revolution in Philosophy: Philosophy, Science & Models of Mind*. The Harvester Press.
- Smart J.J.C. (1959): "Sensations and Brain Processes", *Philosophical Review* 68. (Επανεκδόθηκε στο Rosenthal, D. (ed.) (1990): *The Nature of Mind*. Oxford University Press).
- Smith, B. & Smith, W. D. (eds) (1995): *The Cambridge Companion to Husserl*. Cambridge University Press.
- Smith, B. (1991): "The Owl and the Electric Encyclopedia". *Artificial Intelligence*, 47 (1-3), σελ. 251-288
- Smolensky, P. (1988a): "On the Proper Treatment of Connectionism" *The Behavioral and Brain Sciences*, 11 (1). Επανάδοση στο Goldman 1993.
- Smolensky, P. (1988b): "The Constituent Structure of Connectionist Mental States: A Reply to Fodor and Pylyshyn" στο Beakley & Ludlow 1992.
- Smolensky, P. (1991): "Connectionism and the Language of Thought" Loewer & Rey 1991.
- Snowdon, P. (1981): "Perception, Vision and Causation", *Proceedings of the Aristotelian Society*, σελ. 175-191.
- Snyder, S. (1974): *Madness and the Brain*. McGraw Hill.
- Stalnaker, R. (1986): "Verificationism and a Causal Account of Meaning", *Synthese* 69, σελ. 107-137.
- Stample, D. (1979): "Towards a Causal Theory of Linguistic Representation" στο French, P., Uehling, Th., & Wettstein H. (eds) (1979): *Contemporary Perspectives in the Philosophy of Language*. University of Minnesota Press.
- Sterleny, K. (1989): "The Imagery Debate" στο Lycan 1990.
- Sterleny, K. (1990): *The Representational Theory of Mind - An Introduction*. Blackwell.
- Stewart, H. (1997): *The Ontology of Mind - Events, Processes, and States*. Clarendon Press.
- Stich S. (1983): *From Folk Psychology to Cognitive Science: The Case against Belief*. The MIT Press.
- Stich, S. & Warfield, T. (eds) (1994): *Mental Representation: A Reader*. Blackwell.

- Stroud, B. (1984): *The Significance of Philosophical Skepticism*. Clarendon Press.
- Swoyer, C. (1991): "Structural Representation", *Synthese* 87, σελ. 449-508.
- Taylor, C. (1985): *Human Agency & Language - Philosophical Papers 1*. Cambridge University Press.
- Taylor, C. (1985a): "What Is Involved in a Genetic Psychology?" στο Taylor 1985.
- Taylor, C. (1985b): "Language and Human Nature" στο Taylor 1985.
- Thelen, E. & Smith, L.B. (1993): *A Dynamic Systems Approach to the Development of Cognition and Action*. The MIT Press.
- Tootell, R., Silverman, M., Switkes, E. & De Valois, R. (1982): "Deoxyglucose analysis of retinotopic organization in primate striate cortex", *Science* 218, σελ. 902-904.
- Touretzky, D. & Hinton, G. (1988): "A Distributed Connectionist Production System", *Cognitive Science*, 12, σελ. 423-466.
- Tsinorema, S. (1990): "Wittgenstein, Rationality and Relativism" στο Nicolakopoulos 1990.
- Tugendhat, E. (1976): *Traditional & Analytical Philosophy*. Cambridge University Press.
- Tugendhat, E. (1986): *Self-Consciousness & Self-Determination*. The MIT Press.
- Turing A. (1950): "Computing Machinery and Intelligence", *Mind* 59. Επανεκδόθηκε στο Boden 1990.
- Tye, M. (1991): *The Imagery Debate*. A Bradford Book - The MIT Press.
- Tye, M. (1995): *Ten Problems of Consciousness - A Representational Theory of the Phenomenal Mind*. A Bradford Book - The MIT Press.
- van Gelder, T. & Port, R. (1995a): "It's about Time: An Overview of the Dynamical Approach to Cognition" στο van Gelder & Port 1995.
- van Gelder, T. & Port, R. (eds) (1995): *Mind as Motion: Explorations in the Dynamics of Cognition*. The MIT Press.
- van Gelder, T. (1990): "Compositionality: A Connectionist Variation on a Classical Theme", *Cognitive Science*, 14, σελ. 355-384.
- van Gelder, T. (1991): "What is the "D" in "PDP"? A Survey of the Concept of Distribution" στο Ramsey, Stich & Rumelhart 1991.
- van Gelder, T. (1992): "Defining 'Distributed Representation'", *Connection Science*, vol. 4, Nos. 3 & 4, σελ. 175-191.
- van Gelder, T. (1995): "What Might Cognition Be if not Computation?", *Journal of Philosophy* 92, no. 7, σελ. 345-381.
- van Gelder, T. (1998): "The Dynamical Hypothesis in Cognitive Science", *Behavioral and Brain Sciences*, 21, 1-14.

- Sider, T. (υπό έκδοση): "Intrinsic Properties", *Philosophical Studies*.
- Simon, H. (1969): *The Sciences of the Artificial*. The MIT Press.
- Simons, P.M. (1987): *Parts: A Study in Ontology*. Clarendon Press.
- Skinner, B. F. (1957): *Verbal Behavior*. Englewood Cliffs, Prentice-Hall.
- Skinner, B. F. (1974): *About Behaviorism*. A. Knopf.
- Sloman, A. (1978): *The Computer Revolution in Philosophy: Philosophy, Science & Models of Mind*. The Harvester Press.
- Smart J.J.C. (1959): "Sensations and Brain Processes", *Philosophical Review* 68. (Επανεκδόθηκε στο Rosenthal, D. (ed.) (1990): *The Nature of Mind*. Oxford University Press).
- Smith, B. & Smith, W. D. (eds) (1995): *The Cambridge Companion to Husserl*. Cambridge University Press.
- Smith, B. (1991): "The Owl and the Electric Encyclopedia". *Artificial Intelligence*, 47 (1-3), σελ. 251-288
- Smolensky, P. (1988a): "On the Proper Treatment of Connectionism" *The Behavioral and Brain Sciences*, 11 (1). Επανεκδόση στο Goldman 1993.
- Smolensky, P. (1988b): "The Constituent Structure of Connectionist Mental States: A Reply to Fodor and Pylyshyn" στο Beakley & Ludlow 1992.
- Smolensky, P. (1991): "Connectionism and the Language of Thought" Loewer & Rey 1991.
- Snowdon, P. (1981): "Perception, Vision and Causation", *Proceedings of the Aristotelian Society*, σελ. 175-191.
- Snyder, S. (1974): *Madness and the Brain*. McGraw Hill.
- Stalnaker, R. (1986): "Verificationism and a Causal Account of Meaning", *Synthese* 69, σελ. 107-137.
- Stampe, D. (1979): "Towards a Causal Theory of Linguistic Representation" στο French, P., Uehling, Th., & Wettstein H. (eds) (1979): *Contemporary Perspectives in the Philosophy of Language*. University of Minnesota Press.
- Sterenly, K. (1989): "The Imagery Debate" στο Lycan 1990.
- Sterenly, K. (1990): *The Representational Theory of Mind - An Introduction*. Blackwell.
- Stewart, H. (1997): *The Ontology of Mind - Events, Processes, and States*. Clarendon Press.
- Stich S. (1983): *From Folk Psychology to Cognitive Science: The Case against Belief*. The MIT Press.
- Stich, S. & Warfield, T. (eds) (1994): *Mental Representation: A Reader*. Blackwell.

- Stroud, B. (1984): *The Significance of Philosophical Skepticism*. Clarendon Press.
- Swoyer, C. (1991): "Structural Representation", *Synthese* 87, σελ. 449-508.
- Taylor, C. (1985): *Human Agency & Language - Philosophical Papers 1*. Cambridge University Press.
- Taylor, C. (1985a): "What Is Involved in a Genetic Psychology?" στο Taylor 1985.
- Taylor, C. (1985b): "Language and Human Nature" στο Taylor 1985.
- Thelen, E. & Smith, L.B. (1993): *A Dynamic Systems Approach to the Development of Cognition and Action*. The MIT Press.
- Tootell, R., Silverman, M., Switkes, E. & De Valois, R. (1982): "Deoxyglucose analysis of retinotopic organization in primate striate cortex", *Science* 218, σελ. 902-904.
- Touretzky, D. & Hinton, G. (1988): "A Distributed Connectionist Production System", *Cognitive Science*, 12, σελ. 423-466.
- Tsinorema, S. (1990): "Wittgenstein, Rationality and Relativism" στο Nicolakopoulos 1990.
- Tugendhat, E. (1976): *Traditional & Analytical Philosophy*. Cambridge University Press.
- Tugendhat, E. (1986): *Self-Consciousness & Self-Determination*. The MIT Press.
- Turing A. (1950): "Computing Machinery and Intelligence", *Mind* 59. Επανεκδόθηκε στο Boden 1990.
- Tye, M. (1991): *The Imagery Debate*. A Bradford Book - The MIT Press.
- Tye, M. (1995): *Ten Problems of Consciousness - A Representational Theory of the Phenomenal Mind*. A Bradford Book - The MIT Press.
- van Gelder, T. & Port, R. (1995a): "It's about Time: An Overview of the Dynamical Approach to Cognition" στο van Gelder & Port 1995.
- van Gelder, T. & Port, R. (eds) (1995): *Mind as Motion: Explorations in the Dynamics of Cognition*. The MIT Press.
- van Gelder, T. (1990): "Compositionality: A Connectionist Variation on a Classical Theme", *Cognitive Science*, 14, σελ. 355-384.
- van Gelder, T. (1991): "What is the "D" in "PDP"? A Survey of the Concept of Distribution" στο Ramsey, Stich & Rumelhart 1991.
- van Gelder, T. (1992): "Defining 'Distributed Representation'", *Connection Science*, vol. 4, Nos. 3 & 4, σελ. 175-191.
- van Gelder, T. (1995): "What Might Cognition Be if not Computation?", *Journal of Philosophy* 92, no. 7, σελ. 345-381.
- van Gelder, T. (1998): "The Dynamical Hypothesis in Cognitive Science", *Behavioral and Brain Sciences*, 21, 1-14.

- van Gelder, T. (1998a): "Disentangling Dynamics, Computation and Cognition", *Behavioral and Brain Sciences*, 21
- van Gelder, T. (1999a): "Distributed Representations" στο Wilson & Keil 1999.
- Varela, F., Thompson E., and E. Rosch (1991): *The Embodied Mind - Cognitive Science and Human Experience*. The MIT Press.
- Virvidakis, St. (1984): *Transcendental Arguments, Transcendental Idealism & Skepticism (Remarks on the Transformation of the Transcendental Approach)*. Διδακτορική Διατριβή.
- Von Eckardt, B. (1992): *What Is Cognitive Science?* A Bradford Book - The MIT Press.
- Vosniadou, S. & Brewer, W. (1992): "Mental Models of the Earth: A Study of Conceptual Change in Childhood", *Cognitive Psychology*, 24, σελ. 535-585.
- Warner, R. & Szubka, T. (eds) (1994): *The Mind - Body Problem - A Guide to the Current Debate*. Blackwell.
- Watson, J. (1913a): "Psychology as the Behaviorist views it", *Psychological review*, 20, σελ. 158-177.
- Watson, J. (1913b): "Image and Affection in Behavior", *Journal of Philosophy, Psychology and Scientific Methods*, vol. 10, no. 16, 1913, σελ. 421-424.
- Webb, B (1994): "Robotic Experiments in Cricket Phonotaxis" στο Cliff, D.T., Husbands, P., Meyer, J.-A., and Wilson, S. W. (eds): *From Animals to Animats 3: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*. The MIT Press.
- Weizenbaum, J. (1966): "ELIZA- A Computer Program for the Study of Natural Language Communication between Man and Machine", *Communications of the ACM* 9 (1), σελ. 36-44.
- Weizenbaum, J. (1976): *Computer Power & Human Reason: From Judgment to Calculation*. Penguin Books.
- Wheeler, M. (1994): "Active Perception in Meaningful Worlds". Cognitive science research paper 327, University of Sussex.
- Wheeler, M. (1995): *The Philosophy of Situated Activity*. Sussex. Διδακτορική Διατριβή.
- Whyte, J. (1990): "Success Semantics". *Analysis* 50, σελ. 149-157.
- Whyte, J.T. (1991): "The Normal Rewards of Success", *Analysis*, 51.
- Wiener, N. (1948/1961): *Κυβερνητική ή Έλεγχος και Επικοινωνία στα Ζώα και στις Μηχανές*. (Μετάφραση Ι. Σιατίτσα, Π. Παπαχρήστου). Εκδόσεις Καστανιώτης.
- Wilson, R. & Keil, Fr. (eds.) (1999): *The MIT Encyclopedia of Cognitive Sciences*. The MIT Press.

- Wilson, R. (1995): *Cartesian Psychology & Physical Minds - Individualism & the Sciences of the Mind*. Cambridge University Press.
- Winograd, T. - Flores, F. (1986): *Understanding Computers & Cognition - A New Foundation for Design*. Addison - Wesley Publishing Company.
- Winograd, T. (1972): *Understanding Natural Language*. Academic Press.
- Winograd, T. (1976): "Artificial Intelligence and Language Comprehension" στο *Artificial Intelligence and Language Comprehension*, National Institute of Education.
- Wittgenstein, L. (1921/1971): *Tractatus*. *Δευκαλίων*, 7/8, 1971, σελ. 183-271.
- Wittgenstein, L. (1953): *Philosophical Investigations* (translator G.E.M. Anscombe). Blackwell.
- Wittgenstein, L. (1953/1977): *Φιλοσοφικές Έρευνες* (Μετάφραση Π. Χριστοδουλίδης). Εκδόσεις Παπαζήση.
- Wittgenstein, L. (1958): *The Blue and Brown Books*. Blackwell.
- Wittgenstein, L. (1967): "Wittgenstein's Lectures in 1930-1933" στο Morick 1967.
- Wittgenstein, L. (1980): *Culture and Value*, (επιμελητής: von Wright, G.H.). Blackwell.
- Wittgenstein, L. (1984): *Το Μπλε και το Καφέ Βιβλίο*. (Μετάφραση Κ. Κωβαίος). Εκδόσεις Καρδαμίτσα.
- Wittgenstein, L. (1986): *On Certainty* (Edited By G.E.M. Anscombe & G.H.Von Wright). Harper Torchbooks.
- Wittgenstein, L. (1986): *Πολιτισμός και Αξίες*, (Μετάφραση Δραγώνα-Μονάχου, Μ. & Κωβαίος, Κ.). Εκδόσεις Καρδαμίτσα.
- Wollheim, R. (1980): *Art and its Objects*. Cambridge University Press.
- Αναπολιτάνος, Α. Δ. (1985): *Εισαγωγή στη Φιλοσοφία των Μαθηματικών*. Εκδόσεις "Νεφέλη".
- Βιτγκενστάιν, Λ. (1993): *Φιλοσοφικές Παρατηρήσεις* (Εισαγωγή, μετάφραση, σχόλια: Κ. Μ. Κωβαίος). Εκδόσεις "Γνώση".
- Βυγκότσκι, Λ. (1934): *Σκέψη και Γλώσσα*. Εκδόσεις "Γνώση".
- Έκο, Ου. (1988): *Θεωρία Σημειωτικής*. Εκδόσεις "Γνώση".
- Καγγελάρης, Φ. (1985): *Η Διαδικασία της Αποπροσωποποίησης στη Σχιζοφρένεια - Το Βλέμμα και το Είναι στη Ψύχωση*. Εκδόσεις Καστανιώτη.
- Κάλφας, Β. (1983): *Επιστημονική Πρόοδος και Ορθολογικότητα*. Σύγχρονα Θέματα.
- Καντ Ι. (1976): *Κριτική του Καθαρού Λόγου* (Τόμος Α') (Εισαγωγή, μετάφραση, σχόλια: Α. Γιανναράς). Εκδόσεις Παπαζήση.
- Καντ Ι. (1979): *Κριτική του Καθαρού Λόγου* (Τόμος Β') (Εισαγωγή - Μετάφραση - Σχόλια Α. Γιανναράς). Εκδόσεις Παπαζήση.

- Καραγιάννης, Γ. (1989): *Σήματα και Συστήματα*. Εκδόσεις Ε.Μ.Π..
- Καραγιάννης, Γ., Σταυγιάουερ. (1991): *Αναγνώριση Πρότυπων και Μηχανές που Μαθαίνουν*. Εκδόσεις Συμεών.
- Κατή, Δ., Κονδύλη, Μ. & Νικηφορίδου, Κ. (επιμέλεια) (1999): *Γλώσσα και Νόηση: Επιστημονικές και Φιλοσοφικές Προσεγγίσεις*. Εκδόσεις Αλεξάνδρεια.
- Καφετζόπουλος, Ε. (1995): *Εγκέφαλος, Συνείδηση και Συμπεριφορά, Μια Ιστορική Εισαγωγή στη Νευροψυχολογία*. Εξάντας -Τριανφς Λόγος.
- Κιντή, Β. (1995): *Kuhn & Wittgenstein- Φιλοσοφική Έρευνα της Δομής των Επιστημονικών Επαναστάσεων*. Εκδόσεις Σμίλη.
- Κολέτσος, Γ. (1992): "Εισαγωγή στη Θεωρία Μοντέλων". Ανέκδοτες σημειώσεις.
- Κουτούγκος, Α. (1996): "'Thoughts' in the Intelligent Control of Behaviour and the Capacity to Feel: Comments on Fred Dretske's Machines, Plants, and Animals - When does Behaviour Become Intelligent Action?", Ανακοίνωση στο πρώτο Αθήνα - Pittsburgh Συμπόσιο για την Ιστορία και τη Φιλοσοφία της Επιστήμης και της Τεχνολογίας, Δελφοί, Μάιος 28-30, 1996.
- Κουτούγκος, Α. (2000): "'The Experience of Value' Comments on Michael Tye's 'The Transparency of Experience'", στο τρίτο Αθήνα - Pittsburgh Συμπόσιο με τίτλο *Experience and Knowledge*, Ρέθυμνο, 26-30 Οκτωβρίου 2000.
- Μαραγκός, Γ. (1996): *Γνωσιολογισμοί στο μεταίχμιο της Φιλοσοφίας της Επιστήμης και της Γνωσιοεπιστήμης*. Εκδόσεις Οδυσσέας.
- Μπάλλας, Κ. (Μετάφραση, επιμέλεια) (1984): *Ψυχανάλυση: Φρόντ και Λακάν*. Εκδόσεις Imago.
- Μπαλτάς, Α. (1999): "Ο Wittgenstein και το a priori της γλώσσας" στο Κατή, Κονδύλη, Νικηφορίδου 1999.
- Μπαλτάς, Α. (2000): *Αντικείμενα και Ώψεις Εαυτού*. Εστία
- Μπασάκος, Π. (1999): *Επιχείρημα και Κρίση*. Νήσος
- Μπενβενίστ, Μπάρτ, Ντεριντά, Πιρς, Φουκώ (1981): *Κείμενα Σημειολογίας*. Εκδόσεις "Νεφέλη".
- Ξηροπαίδης Γ. (1995): *Ο Heidegger και το Πρόβλημα της Οντολογίας*. Κριτική.
- Ξηροπαίδης Γ. (1995a): *Ονοματοκρατία και Αυτοσυνείδητη Υποκειμενικότητα - Κριτική Εισαγωγή σε μια Σημασιολογική Προσέγγιση των Φαινομένων της Αυτοσχεσίας*. Ανέκδοτες σημειώσεις μεταπτυχιακού σεμιναρίου στο πρόγραμμα "Ιστορίας και Φιλοσοφίας των επιστημών και της Τεχνολογίας", Τομέας Α.Κ.Ε.Δ., Ε.Μ.Π. 1994-1995.
- Παγωνδιώτης Κ. (1990): *"Η Κριτική της Επιστήμης από τον Χάιντεγκερ"*.

- Παγωνδιώτης, Κ. (1995): "Το Φαινόμενο Συνείδηση", *Δευκαλίων*, 14/1.
- Παγωνδιώτης, Κ. (1999): "Υπάρχει Γλώσσα της Σκέψης;" στο Κατή, Κονδύλη & Νικηφορίδου 1999.
- Πατέλλη, Ι. (1991): *Νόημα και Πράξη*. Εξάντας.
- Πεντζοπουλου-Βαλαλα, Τ. (1985): *Θέματα Σύγχρονης Γερμανικής Φιλοσοφίας - Εισαγωγή στον Husserl*. Εκδοτικός Οίκος Αφων Κυριακίδη.
- Πεντζοπουλου-Βαλαλα, Τ. (1991): *Heidegger ο Φιλόσοφος του Λόγου και της Σιωπής - Φιλοσοφική Έρευνα*. Εκδόσεις Βάνιας.
- Πετρουνάκος, Σ. (1999): "Ιδιωτικό ή Προσωπικό Νόημα;" στο Κατή, Κονδύλη & Νικηφορίδου 1999.
- Πλάτων (1975): *Πλάτωνος Διάλογοι*, τομ. 4. Πάπυρος.
- Πλάτων (1975a): *Θεαίτητος* στο Πλάτων 1975.
- Ράσσελ, Β. (1963): *Τα Προβλήματα της Φιλοσοφίας*. Εκδόσεις Αρσενίδης.
- Ρικέρ, Π. (1990): *Δοκίμια Ερμηνευτικής*. Μορφωτικό Ινστιτούτο Αγροτικής Τράπεζας.
- Συμεωνίδου-Χριστίδου, Τ. (1998): *Εισαγωγή στη Σημασιολογία*. University Studio Press.
- Τσινόρεμα, Β. (1995): "Ο Wittgenstein και η Συζήτηση γύρω από τη Γλωσσική Κανονιστικότητα", *Δευκαλίων* 13, τεύχος 2-3, σελ. 137-173.
- Φρέγκε, Γκ. (1984): *Τα Θεμέλια της Αριθμητικής* (Μετάφραση Γ. Ρουσσόπουλος). Εκδόσεις "Νεφέλη".
- Χαΐνης, Ι.Θ. (1984): *Μαθήματα Ανώτερων Μαθηματικών -τόμος Β₁*. Οργανισμός Εκδόσεως Διδακτικών Βιβλίων.
- Χάιντεγκερ, Μ. (1927). *Είναι και Χρόνος* (Μετάφραση Γ. Τζαβάρας [Α' τόμος 1978 και Β' τόμος 1985]). Εκδόσεις "Δωδώνη".
- Χάιντεγκερ, Μ. (1979): *Φαινομενολογία και Θεολογία* (Μετάφραση Γ. Τζαβάρα). Εποπτεία.
- Χούσσερλ, Ε. (1986): *Δεύτερη Λογική Έρευνα - Η Ιδεατή Ενότητα του Είδους και οι Νεότερες Θεωρίες για την Αφαίρεση*. (Εισαγωγή, μετάφραση, σημειώσεις: Ν.Μ. Σκουτερόπουλος). Εκδόσεις "Γνώση".