

Εισαγωγή στη
Θεωρία και Μεθοδολογία
της Γνωσιακής Επιστήμης

Αθανάσιος Χρ. Πρωτόπαπας

Ιανουάριος 2016

Πρόγραμμα μεταπτυχιακών σπουδών «γνωσιακή επιστήμη»
Τμήμα Μεθοδολογίας, Ιστορίας και Θεωρίας της Επιστήμης
Πανεπιστήμιο Αθηνών

Τρίτη αναθεώρηση (2011), συμπληρωμένη και διορθωμένη (2016).
Για παρατηρήσεις, υποδείξεις, διορθώσεις: aprotopapas@phs.uoa.gr

Περιεχόμενα

Εισαγωγή	1
1 Εισαγωγή στη μελέτη της νόησης	2
1.1 Νους, νόηση, νοητική λειτουργία	2
1.2 Τεχνητή και φυσική νοημοσύνη	4
1.3 Μοντέλα	7
1.4 Αναπαραστάσεις	9
1.5 Επίπεδα περιγραφής	11
1.6 Βούληση και αυτοματισμός	12
Θεωρητικές προσεγγίσεις	15
2 Κλασική (συμβολική) γνωσιακή επιστήμη	15
2.1 Σύμβολα και υπολογισμοί	15
2.2 Συμβολικές καταστάσεις	17
2.3 Παραγωγικότητα και συνδυαστικότητα	20
2.4 Αρθρώματα και τοπικότητα	22
2.5 Περίληψη	24
3 Συνδετισμός	25
3.1 Θεμελιώδεις ιδιότητες βιολογικών νευρώνων	25
3.2 Μαθηματικοί νευρώνες και συνδετιστικά δίκτυα	26
3.3 Προσδιορισμός αναπαραστάσεων και ρύθμιση των συνδέσεων	29
3.4 Είδη και ιδιότητες συνδετιστικών μοντέλων	32
3.5 Περίληψη	35
4 Δυναμικά συστήματα	37
4.1 Τροχιές στο χώρο των φάσεων	37
4.2 Συνεχής και διακριτός χρόνος	40
4.3 Χαρακτηριστικά της δυναμικής προσέγγισης	42
4.4 Σύγκριση συμβολικών και δυναμικών συστημάτων	44
5 Ενσώματη νόηση	49
5.1 Νοητικές αναπαραστάσεις του σώματος	49
5.2 Αρχές της ενσώματης νόησης	52
5.3 Εννοιολογική μεταφορά και αφαιρετικότητα	55
5.4 Οικολογική ψυχολογία	57

6	Πιθανοκρατική (μπεϊζιανή) προσέγγιση	59
6.1	Πιθανοκρατική νόηση	60
6.2	Αναθεώρηση πεποιθήσεων	64
6.3	Ορθολογικότητα	66
6.4	Επίπεδο ανάλυσης	68
7	Εξελικτική ψυχολογία	69
7.1	Εξέλιξη και φυσική επιλογή	69
7.2	Εμφυτότητα και περιβάλλον	72
7.3	Λειτουργικός κατακερματισμός	74
7.4	Παλαιοανθρωπολογία	76
7.5	Διαφυσισιμότητα	78
7.6	Περίληψη	80
8	Σύνοψη των θεωρητικών προσεγγίσεων	80
	Μεθοδολογικές προσεγγίσεις	83
9	Γνωστική ψυχολογία	84
9.1	Θεματικές ενότητες	85
9.2	Θεωρία και πείραμα	86
9.3	Το γενικό γραμμικό μοντέλο και η μηδενική υπόθεση	88
9.4	Γενίκευση και ο μέσος νους	90
9.5	Μοντέλα	92
10	Νευροεπιστήμες	93
10.1	Επίπεδα οργάνωσης και αναγωγισμός	93
10.2	Εντοπιστική υπόθεση	95
10.3	Μεθοδολογικά χαρακτηριστικά των νευροεπιστημών	97
10.4	Νευροαπεικόνιση	99
11	Φιλοσοφία και γλωσσολογία	102
11.1	Γλωσσική ικανότητα	103
11.2	Μεθοδολογία της διαίσθησης	105
11.3	Μάθηση εναντίον έμφυτης γνώσης	108
11.4	Φύση των εννοιών	109
12	Συμπέρασμα: Συμπληρωματικότητα	111

Υπολογιστικά μοντέλα νόησης	113
13 Ο αόριστος των αγγλικών ρημάτων	114
13.1 Εισαγωγή στο πρόβλημα	114
13.2 Η συμβολική προσέγγιση	114
13.3 Το ζήτημα της αναπτυξιακής παλινδρόμησης	116
13.4 Συνδεικτικά μοντέλα του αοριστικού συστήματος	117
14 ACT-R	118
14.1 Απλή δυαδική πρόσθεση στο ACT-R	118
14.2 Ο αόριστος των ρημάτων στο ACT-R	121
15 tlearn	122
15.1 Απλή δυαδική πρόσθεση στο tlearn	123
15.2 Ο αόριστος των ρημάτων στο tlearn	124
Παράρτημα	129
16 Συμπληρωματική βιβλιογραφία	129

Εισαγωγή

Η γνωσιακή επιστήμη έχει ως αντικείμενο τη μελέτη της νόησης. Το θεμελιώδες ερώτημα «τι είναι νους;» και το πιο συγκεκριμένο «ποια είναι η αρχιτεκτονική του νου που επιτελεί τις νοητικές λειτουργίες;» μπορούν να προσεγγιστούν από διαφορετικές αφητηρίες, ανάλογα με τις θεωρητικές προτιμήσεις και το υπόβαθρο κάθε ερευνητή. Η γενική - και ως εκ τούτου αρκετά ασαφής - απάντηση είναι ότι «ο νους είναι μηχανή». Αυτό σημαίνει πως ο νους είναι ένα σύστημα προβλέψιμων σχέσεων με το περιβάλλον και συστηματικής λειτουργίας, βασισμένης σε μέρη με καθορισμένους ρόλους που αλληλεπιδρούν μεταξύ τους με συγκεκριμένους τρόπους. Οι διαφορετικές προσεγγίσεις στο θέμα αφορούν κυρίως στο τι είδους μηχανή είναι ο νους, καθώς αυτό καθορίζει το είδος των μερών από τα οποία μπορεί να αποτελείται και το είδος των μεταξύ τους σχέσεων.

Η υπολογιστική παράδοση της μαθηματικής λογικής, της τεχνητής νοημοσύνης και της «κυβερνητικής» εκφράζεται στη σύγχρονη γνωσιακή επιστήμη με τη *συμβολική* κατεύθυνση για την αρχιτεκτονική του νου, όπου συμβολικές αναπαραστάσεις συμμετέχουν σε συντακτικές διεργασίες σε αλγοριθμικό επίπεδο, ανεξάρτητα από το περιεχόμενό τους. Πρόκειται για την κατεύθυνση που έχει επηρεαστεί και καθοριστεί σε μεγάλο βαθμό από την ανάπτυξη των ηλεκτρονικών υπολογιστών και της πληροφορικής.

Η παράδοση του συνειρισμού και του εμπειρισμού εκφράζεται στο σύγχρονο *συνδεισισμό*, μέσα από ριζικά αναδιαρθρωμένες θεωρήσεις που (ξεφεύγοντας από τον κλασικό συνειρισμό και τον εμπειρισμό) πετυχαίνουν να προσομοιώσουν διεργασίες υψηλής πολυπλοκότητας χωρίς να θυσιάζουν τις βασικές αρχές της παράλληλης κατανομημένης επεξεργασίας και της ευέλικτης μάθησης σε δίκτυα απλών διασυνδεδεμένων κόμβων. Η κατεύθυνση αυτή βασίζεται σε βιολογικής έμπνευσης αρχές λειτουργίας, βασισμένες σε απλουστευμένα νευρωνικά δίκτυα.

Η *δυναμική* προσέγγιση έρχεται να αντιμετωπίσει κατά μέτωπο την παραγνωρισμένη διάσταση του χρόνου στις νοητικές διεργασίες, εισάγοντας μια συνεχή, μη αναπαραστασιακή θεώρηση βασισμένη σε μαθηματική γλώσσα. Ως μετεξέλιξη της, η θεώρηση της *ενσώματης* νόησης δίνει πρωταρχικό ρόλο στην ύπαρξη και τη συμμετοχή του σώματος, αφενός στην προέλευση, τον προσδιορισμό και την ερμηνεία των νοητικών αναπαραστάσεων και αφετέρου στην αλληλεπίδραση με το περιβάλλον και με άλλους ενσώματους νούς.

Η ταχύτητα ανερχόμενη τα τελευταία χρόνια *πιθανοκρατική* προγέννηση δίνει έμφαση στην ορθολογικότητα της νοητικής λειτουργίας, με την έννοια της λήψης αποφάσεων στις οποίες η νόηση καταλήγει πιθανοκρατικά συνδυάζοντας όλες τις διαθέσιμες πληροφορίες με το θεωρητικά βέλτιστο τρόπο. Αντίθετα με τις προηγούμενες προσεγγίσεις δεν εστιάζει στο γνωσιακό μηχανισμό εκτέλεσης των υπολογισμών αλλά στο αποτέλεσμά τους, σε σχέση με τα κριτήρια και το περιβάλλον του οργανισμού.

Στο πλαίσιο της ανάπτυξης αυτών των εναλλακτικών προσεγγίσεων, παρουσιάζεται και η *εξελικτική* θεώρηση, σύμφωνα με την οποία ο νους πρέπει να ιδωθεί στο πλαίσιο εξέλιξης του ανθρώπου ως είδους με συγκεκριμένες ιδιότητες, ανάγκες και περιορισμούς

που απορρέουν από το περιβάλλον στο οποίο επιβίωσε ως διακριτό βιολογικό είδος. Έτσι, οι νοητικές λειτουργίες είναι αποτέλεσμα εξελικτικών πιέσεων για τη λύση συγκεκριμένων επιτακτικών προβλημάτων επιβίωσης και αναπαραγωγής που αντιμετώπισαν οι μικρές κοινωνίες των μακρινών προγόνων μας στις σαβάνες της Αφρικής.

Στο πρώτο μέρος των σημειώσεων αυτών, μετά από μια γενική εισαγωγή στα προβλήματα της μελέτης της νόησης και τις κοινές επιδιώξεις των γνωσιακών επιστημών, περιγράφονται και συγκρίνονται μεταξύ τους οι βασικές αυτές κατευθύνσεις στην προσέγγιση, κατανόηση και περιγραφή των νοητικών φαινομένων. Ως μια πρώτη εισαγωγή στη θεωρία και φιλοσοφία της γνωσιακής επιστήμης, το μέρος αυτό χρησιμεύει καθοδηγητικά για την κατανόηση των θεωρητικών καταβολών των διαφόρων προσεγγίσεων και μελετών στη βιβλιογραφία. Θα ήταν ιδιαίτερα επιτυχημένο αν κατάφερε να προσανατολίσει τον αναγνώστη να αναγνωρίζει τα βασικά χαρακτηριστικά κάθε προσέγγισης και να ευαισθητοποιείται στα εκάστοτε δυνατά και αδύνατα σημεία της, ώστε να αξιολογεί καλύτερα τα διάφορα μοντέλα και τα συμπεράσματα που εξάγονται από αυτά.

1 Εισαγωγή στη μελέτη της νόησης

1.1 Νους, νόηση, νοητική λειτουργία

Στο ερώτημα «τι είναι νους;» δεν μπορούμε εύκολα να απαντήσουμε με κάποιον ουδέτερο ορισμό που να μην υπονοεί συγκριμένες θεωρητικές προτιμήσεις. Μπορούμε όμως σε πρώτη προσέγγιση να παραθέσουμε έναν κατάλογο από περιπτώσεις που εμπίπτουν σε αυτό που εννοούμε νόηση, διερευνώντας έτσι τα όρια μεταξύ δεργασιών που εμπίπτουν στη νόηση και άλλων που δεν τις θεωρούμε νοητικές, ακόμα κι αν μοιράζονται σημαντικές ιδιότητες με τις νοητικές. Είναι σημαντικό να κατανοήσουμε εξ αρχής πως υπάρχει ένα εύρος δεργασιών, από τη μαθηματική σκέψη για τη λύση ενός προβλήματος, για παράδειγμα, που όλοι συμφωνούμε πως είναι αμιγώς νοητική διεργασία, ως το βάδισμα ή τη διατήρηση της ισορροπίας μας σε όρθια θέση, που δεν είναι τόσο σαφές αν πρέπει να χαρακτηριστούν νοητικά φαινόμενα και να μελετηθούν ως τέτοια. Υπάρχουν βέβαια και φαινόμενα όπως η ρύθμιση της αναπνοής ή της θρέψης που κατά κοινή ομολογία ανήκουν στο χώρο της φυσιολογίας και όχι της νόησης, παρότι αλληλεπιδρούν ισχυρά με νοητικές διεργασίες.¹

Οι περιπτώσεις που χαρακτηρίζουμε νοητικές διεργασίες περιλαμβάνουν μια τεράστια ποικιλία δραστηριοτήτων: διασχίζω ένα ρυάκι πηδώντας βράχο-βράχο, διαλέγω το μελλοντικό μου επάγγελμα, εξηγώ σ' ένα φίλο μια ιδέα μου. Οι δραστηριότητες αυτές δεν μοιάζουν να έχουν τίποτα κοινό μεταξύ τους, μοιράζονται όμως την ανάγκη της υπόθεσης

¹Για παράδειγμα, ο ρυθμός της αναπνοής επηρεάζεται από τη συγκινησιακή κατάσταση και άρα έμμεσα από τα ερεθίσματα που μας επηρεάζουν συναισθηματικά. Όταν πεινάμε, όχι μόνο είναι δύσκολο να συγκεντρωθούμε σε ένα νοητικό έργο, αλλά είναι και πιο πιθανό να φανταστούμε θέματα σχετικά με φαγητό, επηρεάζοντας έτσι τη σκέψη μας.

του νου για την κατανόηση και την περιγραφή τους. Ακόμα και η επιλογή ενός βράχου από τους άλλους που βρίσκονται λιγότερο ή περισσότερο μέσα στο νερό του ρυακιού, η εκτίμηση ότι αυτός ο βράχος βρίσκεται εντός της ακτίνας ασφαλούς άλματος, ο υπολογισμός και η εκτέλεση της απαραίτητης μυικής δραστηριότητας, ίσως και η τελική ανακούφιση ότι τελικά έμεινα στεγνός, όλο αυτό αποτελεί καθαρά νοητική δραστηριότητα. Μπορεί κανείς να προσπαθήσει να μειώσει το νοητικό βάθος μιας τέτοιας δραστηριότητας, σε σύγκριση π.χ. με την απόδειξη ενός μαθηματικού θεωρήματος, αποδίδοντάς της μια «μηχανιστική» εξήγηση. Στην πραγματικότητα όμως μια τέτοια θεώρηση υποτιμά την πολυπλοκότητα και την υποκειμενικότητα της δραστηριότητας, αλλά υποτιμά επίσης και το φυσικό ντετερμινισμό που αναγκαστικά κυριαρχεί στην αιτιοκρατική λειτουργία του φυσικού υποβάθρου του νου, όπως θα αναλυθεί στη συνέχεια.

Ας εξετάσουμε λίγο περισσότερο το πρώτο παράδειγμα νοητικής διεργασίας, το πέρασμα ενός ρυακιού με άλματα από βράχο σε βράχο. Τη στιγμή πριν από κάθε άλμα, τι είναι αυτό που κάνει ένα βράχο να ξεχωρίζει από το φυσικό περιβάλλον, το οποίο είναι γεμάτο βράχους, και να γίνεται στόχος μιας σχεδιαζόμενης τροχιάς και τελικά βάθρο στήριξης; Οι φυσικές ιδιότητες του περιβάλλοντος δεν περιλαμβάνουν την ειδοποιό διαφορά που διακρίνει το βράχο-στόχο από τους άλλους. Μόνο στο νου ορίζεται η επιλογή και στοχοθέτηση, όπως στο νου επιτελείται και ο απαραίτητος υπολογισμός για το άλμα. Μα, θα πει κανείς, δεν υπολογίζω συνειδητά τις δυνάμεις που θα ασκήσουν οι μυς, απλώς κοιτάζω το βράχο και πηδώ. Όμως η επίγνωση δεν είναι απαραίτητη για τη νοητική λειτουργία, παρότι ορισμένοι πιστεύουν ότι το πιο ενδιαφέρον κομμάτι του νου είναι εκείνο που σχετίζεται με τη φευγαλέα εστίαση της προσοχής από τη μια σκέψη στην άλλη, σχηματίζοντας αυτό που αντιλαμβανόμαστε ως «ειρμό των σκέψεων». Η απουσία επίγνωσης, που συχνά σχετίζεται με το βαθμό αυτοματοποίησης μιας διεργασίας, στα πλαίσια της γνωστικής ψυχολογίας έχει να κάνει με τη διαθεσιμότητα γνωσιακών πόρων και όχι με το νοητικό χαρακτήρα μιας διεργασίας. Το αν ένα έργο εκτελείται συνειδητά ή αυτόματα δεν επηρεάζει το χαρακτηρισμό του ως νοητικού ή άλλου τύπου.

Όταν επιτελούμε μια νέα, άγνωστη ή ασυνήθιστη διεργασία, πρέπει να συγκεντρωθούμε σ' αυτήν και να κατευθύνουμε συνειδητά τη σκέψη και τις δράσεις μας, έχοντας ίσως δυσκολία στο να συγκρατήσουμε και να συνυπολογίσουμε συγχρόνως όλες τις απαραίτητες όψεις της. Με την επανάληψη και την τελειοποίηση έρχεται η συνήθεια, η απελευθέρωση δηλαδή των γνωσιακών πόρων και η ανάθεση της διεργασίας σε μη συνειδητούς μηχανισμούς. Σταδιακά, η επίγνωση μειώνεται και δίνει τη θέση της στην αυτοματοποίηση, η οποία περιλαμβάνει τη βελτίωση, τη σταθεροποίηση και την επιτάχυνση της διεργασίας. Το κλειδί είναι η διαβάθμιση της επίγνωσης κατά τη διάρκεια της αυτοματοποίησης, διότι εξαιτίας της διαβάθμισης γίνεται σαφές ότι δεν μπορούμε να διαχωρίσουμε σαφώς κατηγορίες μεταξύ αυτόματης και συνειδητής λειτουργίας. Άρα δεν μπορούμε να ξεχωρίσουμε τις συνειδητές από τις αυτόματες λειτουργίες και να θεωρήσουμε μόνο τις πρώτες ως νοητικές ενώ τις άλλες να τις απορρίψουμε ως άνευ νοητικού ενδιαφέροντος. Επανερχόμενοι στον αυτόματο υπολογισμό του άλματος προς το βράχο, είμαστε υποχρεωμένοι να δεχτούμε

ότι αποτελεί νοητική λειτουργία, ανεξάρτητα από το πόσο μηχανικά γίνεται κάθε φορά, εφόσον για την εκτέλεσή του χρειαζόμαστε την υπόθεση του νου και των υπολογιστικών του δυνατοτήτων.

Το δεύτερο παράδειγμα νοητικής λειτουργίας, η εξήγηση μιας ιδέας σε κάποιο φίλο, ίσως είναι προφανέστερα νοητικό, εφόσον μια ιδέα δεν έχει πεδίο ορισμού άλλο από το νου. Τι μπορεί να είναι μια ιδέα αν όχι το περιεχόμενο ενός νου; Ουσιαστικά πρόκειται για ένα κυκλικό παράδειγμα, γιατί περιέχει τη νοητική λέξη «ιδέα». Το τρίτο όμως παράδειγμα είναι πιο διαφωτιστικό: η επιλογή μελλοντικού επαγγέλματος είναι σαφώς μια νοητική λειτουργία, ο λόγος όμως δεν είναι ότι είναι ήδη διατυπωμένη νοητικά. Εισερχόμενοι στην ουσία της λειτουργίας, διαπιστώνουμε ότι για την επιλογή είναι απαραίτητο να λάβουμε υπόψη τις *επιθυμίες* μας και τις *πεποιθήσεις* μας. Για παράδειγμα, τι *μας* *αρέσει* να κάνουμε αλλά και τι *επιθυμούμε* να πετύχουμε στο μέλλον, πού *πιστεύουμε* ότι υπάρχει προσφορά εργασίας, ποια εργασία *νομίζουμε* ότι προσφέρει υψηλούς μισθούς ή κάποιο συγκεκριμένο επαγγελματικό περιβάλλον κλπ. Οι επιθυμίες και οι πεποιθήσεις είναι κατεξοχήν νοητικές έννοιες και αποτελούν ένα πολύ ενδιαφέρον εννοιολογικό σύστημα που απασχολεί ιδιαίτερα τους φιλοσόφους της νόησης.

Επιχειρώντας μια γενική απάντηση στο ερώτημα «τι είναι νους;», μπορούμε να θεωρήσουμε ως «νοητική» οποιαδήποτε λειτουργία απαιτεί την υπόθεση του νου για να περιγραφεί και να εξηγηθεί. Αν μια λειτουργία περιγράφεται με ικανοποιητικό τρόπο σε επίπεδο φυσικής, χημείας, ή φυσιολογίας, τότε δεν χρειάζεται να τη θεωρήσουμε νοητική. Αν όμως για να την κατανοήσουμε είναι απαραίτητο να την περιγράψουμε σε νοητικό επίπεδο τότε θεωρούμε ότι είναι νοητική λειτουργία. Άλλωστε και η ίδια η έννοια του νου είναι θεωρητική, καθώς ο νους δεν είναι δημόσια παρατηρήσιμος. Στη γνωσιακή επιστήμη, ο νους είναι μια υπόθεση που κάνουμε, υπό τη μορφή κάποιας μηχανής, για να εξηγήσουμε φαινόμενα συμπεριφοράς που φαίνεται πως απαιτούν σύνθετες εσωτερικές λειτουργίες για την αντικειμενική περιγραφή τους.

1.2 Τεχνητή και φυσική νοημοσύνη

Μια κεντρική θέση της κλασικής γνωσιακής επιστήμης θέλει το νου να είναι συγκρίσιμος με έναν ηλεκτρονικό υπολογιστή, και σε πολλά ουσιαστικά σημεία παρόμοιος με αυτόν. Η θέση αυτή αναλύεται σε επόμενες ενότητες. Προς το παρόν μας ενδιαφέρει το είδος της σύγκρισης που εννοείται και η σχέση της λεγόμενης «τεχνητής νοημοσύνης» με τη νόηση, δηλαδή τη «φυσική νοημοσύνη», αν μπορεί κανείς να χρησιμοποιήσει έναν τέτοιο όρο. Ένα παράδειγμα που χρησιμοποιείται συχνά αναφορικά με τις νοητικές δυνατότητες των ηλεκτρονικών υπολογιστών είναι το σκάκι. Μπορεί κανείς να διαβάσει για σύγκριση ανθρώπινης και μηχανικής νοημοσύνης, για τη μάχη μεταξύ ανθρώπου και μηχανής, για τις απεριόριστες δυνατότητες των μηχανών που αύριο-μεθαύριο θα κατακτήσουν τον κόσμο κλπ. Η πραγματικότητα είναι πολύ πιο πεζή και ο άνθρωπος μάλλον δεν κινδυνεύει από τους ηλεκτρονικούς υπολογιστές στο διανοητικό πεδίο πουθενά στο προβλέψιμο μέλλον. Απλώς το σκάκι είναι κάκιστο παράδειγμα νόησης και παντελώς ακατάλληλο πεδίο σύγ-

κρισης. Είναι περίπου σα να συγκρίνουμε έναν μαθητή με το κομπιουτεράκι του στην αριθμητική πράξη της πρόσθεσης και να συμπεραίνουμε ότι τα κομπιουτεράκια είναι οι μαθητές του μέλλοντος και μάλιστα θα είναι καλύτεροι από τα παιδιά.

Το σκάκι είναι ένα κλειστό, ιδιαίτερα περιορισμένο και καλά προκαθορισμένο σύστημα-«παιχνίδι», στο οποίο είναι κάθε φορά επιτρεπτό μόνο ένα μικρό (καλά καθορισμένο) σύνολο διακριτών κινήσεων, δηλαδή μεταβάσεων από την τρέχουσα κατάσταση σε μια επόμενη. Μια παρτίδα αποτελείται από μια αλληλουχία διακριτών καταστάσεων (θέσεων σκακιέρας), ξεκινώντας πάντα από την ίδια αρχική κατάσταση και στοχεύοντας σε μια συγκεκριμένη, πάντα την ίδια, λειτουργικά ορισμένη κατάσταση (το ματ). Κάθε κατάσταση ορίζεται πλήρως από τις διακριτές θέσεις ενός συγκεκριμένου συνόλου συμβόλων (μια και το συγκεκριμένο σχήμα ή το ακριβές χρώμα των κομματιών δεν παίζει ρόλο). Με δυο λόγια, το σκάκι μοιάζει σχεδιασμένο για να αναπαρίσταται σε έναν ηλεκτρονικό υπολογιστή και όχι σε έναν ανθρώπινο νου. Για το λόγο αυτό είναι και ιδιαίτερα δύσκολο για τον ανθρώπινο νου, αποτελεί μια νοητική πρόκληση για τους λίγους που ασχολούνται μ' αυτό, ενώ οι πραγματικά καλοί σκακιστές αποτελούν ασήμαντο ποσοστό του γενικού πληθυσμού. Κανένας δεν μαθαίνει σκάκι από μόνος του σε βρεφική ή πρώτη παιδική ηλικία ούτε μπορεί να παίξει σκάκι χωρίς κόπο και συγκέντρωση.

Εξαιτίας των περιορισμών της σκακιέρας και των επιτρεπτών κινήσεων των κομματιών, ο αριθμός των δυνατών παρτίδων είναι πεπερασμένος (εξαιρώντας επαναλήψεις καταστάσεων και τις συνεπαγόμενες ατέρμονες κυκλικότητες). Μπορεί να είναι ένας ασύλληπτα τεράστιος αριθμός, αλλά παραμένει κάποιος αριθμός. Αν υπήρχε αρκετή αποθηκευτική δυνατότητα στο σύμπαν θα ήταν δυνατό κάποιος να καταγράψει και να αποθηκεύσει όλες τις δυνατές παρτίδες σε έναν υπολογιστή και να κατασκευάσει την τέλεια σκακιστική μηχανή που θα επιλέγει κάθε φορά τη συνέχεια εκείνη της παρτίδας που θα οδηγεί με μεγαλύτερη πιθανότητα στη νίκη.² Ο ανθρώπινος νους, επειδή διαθέτει ευέλικτες στρατηγικές, μπορεί και ξεπερνά το πρόβλημα της χωρητικότητας αντιμετωπίζοντας το παιχνίδι εν μέρει ως μια άσκηση αναγνώρισης σχέσεων και προτύπων. Παρακάμπτει δηλαδή την αντικειμενική δυσκολία, εφαρμόζοντας μια εντελώς διαφορετική στρατηγική από αυτή που θα υπαγόρευε η μαθηματική ανάλυση του προβλήματος. Ο ηλεκτρονικός υπολογιστής, επειδή δεν διαθέτει τέτοιου είδους στρατηγικές, προσεγγίζει όλο και καλύτερα, με την πρόοδο της τεχνολογίας, την τέλεια σκακιστική μηχανή αυξάνοντας τη χωρητικότητά του και την ταχύτητα επεξεργασίας ενός μεγάλου αριθμού πιθανών θέσεων (μελλοντικών κινήσεων). Πολύ απλά, η ουσία της νοητικής διεργασίας του ανθρώπου σκακιστή δεν έχει καμία σχέση με την ουσία της υπολογιστικής διεργασίας της μηχανής.

Τελικά μια σκακιστική μηχανή κερδίζει (ή θα κερδίζει μετά από λίγη ακόμα τεχνολογική πρόοδο) τους καλύτερους σκακιστές. Αλλά και κάθε ταπεινό κομπιουτεράκι κερδίζει όλους τους μαθηματικούς στον υπολογισμό τετραγωνικών ριζών (για να μην πούμε για λο-

²Επειδή η απαιτούμενη χωρητικότητα είναι όντως ασύλληπτη, στην πράξη ένας ηλεκτρονικός υπολογιστής συνδυάζει δύο τακτικές: μνήμη ενός πολύ μεγάλου αριθμού παρτίδων και τη δυνατότητα εξαντλητικής διερεύνησης (και ευριστικής αξιολόγησης) των πιθανών συνεχειών μιας παρτίδας, έξι, επτά ή και περισσότερες κινήσεις μπροστά.

γάριθμους). Αυτό όμως δεν προκαλεί κάποιον να υποστηρίξει ότι το κομπιουτεράκι είναι εξυπνότερο από τους μαθηματικούς.

Η ανθρώπινη νόηση είναι καταρχήν φτιαγμένη (για την ακρίβεια, εξελιγμένη) έτσι ώστε να αντιμετωπίζει και να επιλύει τα διαρκή προβλήματα της επιβίωσης και αναπαραγωγής στο συγκεκριμένο περιβάλλον όπου εξελίχθηκε. Είναι προσαρμοσμένη, για παράδειγμα, στην οπτική αντίληψη και ταχύτατη αναγνώριση και αξιολόγηση των αντικειμένων του οπτικού πεδίου, υπό οποιοσδήποτε συνθήκες φωτισμού, ακόμα κι αν μερικά από αυτά κρύβονται από άλλα. Είναι προσαρμοσμένη στην κατασκευή εργαλείων και την εξεύρεση οδών και μεθόδων για την προσέγγιση τροφής και θηραμάτων και την αποφυγή θηρευτών, καθώς και στην επίλυση προβλημάτων στέγασης, θέρμανσης, επιλογής ζευγαριού και αναπαραγωγής.³ Είναι προσαρμοσμένη στην εσωτερική λεπτομερειακή αναπαράσταση του περιβάλλοντος χώρου και των γύρω περιοχών και στον προσανατολισμό, όπως και στη συγκρότηση διαρκούς εικόνας για το άμεσο περιβάλλον συνδυάζοντας στοιχεία από διάφορες αισθητήριες οδούς. Είναι προσαρμοσμένη στη λεκτική (αλλά και τη μη λεκτική) επικοινωνία, στην επεξεργασία και παραγωγή γλωσσικών συμβόλων, στις κοινωνικές αλληλεπιδράσεις κλπ. Σε όλα αυτά τα προβλήματα η επίδοση του καλύτερου διαθέσιμου εξειδικευμένου ηλεκτρονικού υπολογιστή είναι μάλλον απογοητευτική, και έτσι θα παραμείνει για πολύ καιρό ακόμα.

Οι φυσικές λειτουργίες του νου είναι τόσο φυσικές που δεν μας φαίνεται πόσο περίπλοκες μπορεί να είναι από υπολογιστική σκοπιά, ή πόσο δύσκολα είναι τα προβλήματα που διαρκώς με μεγάλη ευκολία επιλύει οποιοσδήποτε νους, χωρίς ειδική άσκηση ή διδασκαλία. Κανένας φυσιολογικός άνθρωπος δεν χρειάζεται να διδαχθεί πώς να βλέπει, να μιλά ή να τρέχει, ούτε απαιτείται ιδιαίτερη εξάσκηση γι' αυτά (ενώ σκεφτείτε πόσοι καλοί σκακιστές υπάρχουν στο σύνολο του πληθυσμού και πόσο έχουν κοπιάσει για να φτάσουν στο επίπεδο αυτό). Αυτές είναι οι διεργασίες που θα έπρεπε να ενδιαφέρουν οποιονδήποτε μελετά τη νόηση ή θέλει να κατασκευάσει τεχνητή νόηση. Αντίθετα, τα προβλήματα που για το νου είναι αφύσικα, όπως για παράδειγμα τα μαθηματικά, κάθε φυσιολογικός άνθρωπος τα βρίσκει δύσκολα (ακόμα και αν τα θεωρεί ελκυστικά) και μαθαίνει να τα λύνει μετά από πολύχρονες σπουδές και εξάσκηση. Αυτά δεν αποτελούν πεδία φυσικής νόησης αλλά είναι προβλήματα άσχετα προς τη λειτουργία (και τους σκοπούς) του νου. Παρότι τραβούν το ενδιαφέρον των μηχανικών το μόνο που μπορούν να προσφέρουν είναι η οριοθέτηση του φυσικού πεδίου νου, από την εξωτερική πλευρά: μας δείχνουν μέχρι πού μπορεί να φτάσει η αποτελεσματικότητα του νου αν τον πιέσουμε και τον σπρώξουμε στα όριά του.

³ Δεν υπονοείται ότι κάθε εξειδικευμένη συμπεριφορά ή ικανότητα είναι εξελικτικά προσαρμοσμένη. Αυτή είναι μια πολύ ισχυρή εκδοχή που προωθείται από την προσέγγιση της εξελικτικής ψυχολογίας (βλ. Κεφ. 7). Εδώ υποστηρίζεται η πολύ ασθενέστερη θέση ότι τελικά ο οργανισμός θα πρέπει να είναι συνολικά προσαρμοσμένος ώστε να αντιμετωπίζει με επιτυχία τα βασικά προβλήματα επιβίωσης και αναπαραγωγής.

1.3 Μοντέλα

Στην επιστημονική ορολογία, η λέξη μοντέλο χρησιμοποιείται για να υποδηλώσει ένα τεχνητό σύστημα που κατασκευάζουμε με στόχο την προσομοίωση και την κατανόηση ενός φυσικού συστήματος. Βασικά χαρακτηριστικά ενός μοντέλου αποτελούν η αφαιρετικότητα και η προβλεπτική ισχύς. Και τα δύο αφορούν στη σχέση του μοντέλου με το πρωτότυπο (φυσικό) σύστημα αναφοράς.

Η αφαιρετικότητα έχει να κάνει με την επιλογή των χαρακτηριστικών εκείνων του φυσικού συστήματος τα οποία κρίνει ο μελετητής του ότι είναι ουσιώδη και για το λόγο αυτό τα διατηρεί και τα ενσωματώνει στο μοντέλο. Αντίθετα, χαρακτηριστικά που θεωρούνται δευτερεύοντα, συμπτωματικά, ή περιφερειακά, παραλείπονται με στόχο την απλούστευση του μοντέλου και την καλύτερη κατανόηση του συστήματος αναφοράς. Για παράδειγμα, στην περίπτωση ενός αεροπλάνου, το χρώμα μπορεί να θεωρηθεί συμπτωματικό χαρακτηριστικό, ενώ το σχήμα και η πιητική ικανότητα μάλλον ουσιώδη χαρακτηριστικά. Αποτέλεσμα της καλής επιλογής ουσιωδών χαρακτηριστικών είναι η επιτυχημένη προσομοίωση από το μοντέλο της επιθυμητής συμπεριφοράς ή ιδιότητας του συστήματος αναφοράς. Έτσι, στο παράδειγμα του αεροπλάνου, η διατήρηση των ουσιωδών αεροδυναμικών χαρακτηριστικών μπορεί να δώσει στο μοντέλο πιητικά χαρακτηριστικά παρόμοια με αυτά του πραγματικού αεροπλάνου, οπότε και το μοντέλο θεωρείται πετυχημένο. Αντίθετα, η λανθασμένη επιλογή (ή λανθασμένη μεταφορά στις συνθήκες του μοντέλου) μπορεί να έχει ως αποτέλεσμα αδυναμία πτήσης ή πιητικά χαρακτηριστικά διαφορετικά από αυτά του πραγματικού αεροπλάνου, οπότε και το μοντέλο θεωρείται αποτυχημένο.

Η αφαιρετικότητα στην επιλογή ουσιωδών χαρακτηριστικών δεν αφορά μόνο στην κατασκευή αλλά και στη λειτουργία του μοντέλου. Ο μελετητής ενός συστήματος επιλέγει τις λειτουργίες του συστήματος που τον ενδιαφέρουν και στη συνέχεια κατασκευάζει ένα μοντέλο επιλέγοντας χαρακτηριστικά του συστήματος που θεωρεί σημαντικά με στόχο να επιτύχει την ακριβή προσομοίωση των λειτουργιών εκείνων που επέλεξε. Αυτές οι δύο ομάδες επιλογών είναι στενά συνδεδεμένες μεταξύ τους. Στο παράδειγμα του αεροπλάνου, το χρώμα είναι δευτερεύον χαρακτηριστικό επειδή η λειτουργία-στόχος του μοντέλου είναι η πτήση. Αν η λειτουργία-στόχος ήταν η ορατότητα ή η αντανάκλαστικότητα του φωτός, τότε πιθανότατα το χρώμα να ήταν ουσιώδες χαρακτηριστικό ενώ οι αεροδυναμικές ιδιότητες δευτερεύουσας σημασίας.

Το ζήτημα της προβλεπτικής ισχύος έχει κεντρική θέση στην έννοια του μοντέλου. Εν μέρει έχει να κάνει με την επιλογή της λειτουργίας εκείνης του φυσικού συστήματος που μας ενδιαφέρει. Προφανώς το μοντέλο θα πρέπει, δεδομένων των κατάλληλων παραμέτρων, να «προβλέπει» σωστά τη συμπεριφορά του συστήματος αναφοράς, να πετυχαίνει δηλαδή με ακρίβεια τη λειτουργία-στόχο. Πιο σημαντικό ακόμα, όμως, είναι το μοντέλο να προβλέπει χαρακτηριστικά της λειτουργίας του συστήματος αναφοράς που δεν έχουν προβλεφθεί στη σχεδιάσή του ή, ακόμα, που είναι άγνωστα στο μελετητή. Με τον τρόπο αυτό το μοντέλο χρησιμοποιείται ως μέσο πειραματισμού με την έννοια της επιστημονικής επαλήθευσης και έτσι μεγιστοποιεί τη χρησιμότητά του ως διανοητικό κατασκεύασμα.

Για παράδειγμα, μπορεί το μοντέλο ενός αεροπλάνου να παραμορφώνεται ή να καταστρέφεται εν πτήση υπό ορισμένες συνθήκες ταχύτητας-θερμοκρασίας- ανέμου κλπ., ενώ τέτοιες συνθήκες δεν έχουν παρατηρηθεί στο πραγματικό αεροπλάνο. Αν εκ των υστέρων επαληθευτεί η «πρόβλεψη» του μοντέλου, αν δηλαδή προκύψουν (φυσικά ή τεχνητά) οι κατάλληλες συνθήκες για το πραγματικό αεροπλάνο και το αποτέλεσμα είναι πράγματι η παραμόρφωση ή καταστροφή, τότε είναι προφανές ότι το μοντέλο είναι όχι μόνο ακριβές όσον αφορά στη λειτουργία-στόχο, δηλαδή την πτήση υπό συνηθισμένες συνθήκες, αλλά και ένα σημαντικό εργαλείο μέσα από το οποίο μπορούμε να μάθουμε περισσότερα για το ίδιο το σύστημα αναφοράς (το πραγματικό αεροπλάνο). Η τελευταία ιδιότητα είναι πολύ σημαντική διότι οι επιθυμητοί χειρισμοί μπορεί να είναι αδύνατοι ή ανεπίτρεπτοι στο σύστημα αναφοράς, ή μπορεί η πολυπλοκότητά τους να είναι απαγορευτική. Είναι προφανής η επιστημολογική συγγένεια του επιτυχημένου μοντέλου με την επιτυχημένη επιστημονική θεωρία, με τη διαφορά ότι ένα διανοητικό κατασκεύασμα, για να είναι μοντέλο και όχι θεωρία, θα πρέπει να έχει συγκεκριμένη υπόσταση (π.χ., υλοποίηση σε φυσικό μέσο ή λογισμικό) ώστε να επιτρέπει άμεση και συγκεκριμένη αλληλεπίδραση με τον κατασκευαστή του.

Επανερχόμενοι στο ζήτημα της τεχνητής νοημοσύνης ως μοντέλου του νου, μπορούμε να διακρίνουμε δύο κατηγορίες συστημάτων, οι οποίες αντιστοιχούν σε δύο αποκλίνουσες προσεγγίσεις τεχνητής νοημοσύνης. Η λεγόμενη *ασθενής τεχνητή νοημοσύνη* (weak AI) αφορά σε μοντέλα που στοχεύουν στην αναπαραγωγή της περιγραφής μιας συμπεριφοράς, δηλαδή μιας σχέσης εισροής-εκροής, χωρίς να δίνεται ιδιαίτερη σημασία στον τρόπο που επιτυγχάνεται η συμπεριφορά. Μια σκακιστική μηχανή είναι κατά κανόνα μοντέλο τέτοιου τύπου, εφόσον πρωταρχικό (αν όχι μοναδικό) μέλημα στη σχεδίασή της είναι να παίζει τόσο καλά σκάκι όσο οι καλύτεροι σκακιστές και μόνο μέτρο επιτυχίας της είναι να τους κερδίζει. Αντίθετα, η λεγόμενη *ισχυρή τεχνητή νοημοσύνη* (strong AI) αφορά σε μοντέλα που έχουν στόχο την προσομοίωση όχι μόνο του αποτελέσματος αλλά και της εσωτερικής λειτουργίας, σε συμβολικό τουλάχιστον επίπεδο. Στη γνωσιακή επιστήμη στόχος δεν είναι η προσομοίωση της συμπεριφοράς αλλά της νοητικής λειτουργίας καθαυτής, ενώ για ένα μηχανικό τεχνητής νοημοσύνης το ζητούμενο συνήθως είναι η κατασκευή ενός συστήματος που θα πετυχαίνει κάποιο συγκεκριμένο αποτέλεσμα. Για το λόγο αυτό η γνωσιακή επιστήμη έχει μόνο ιστορική σχέση με την τεχνητή νοημοσύνη.

Ο ρόλος των μοντέλων στη γνωσιακή επιστήμη είναι εργαλειώδης. Δεν προσπαθούμε να κατασκευάσουμε τεχνητή νόηση αλλά να διερευνήσουμε τις συνέπειες των θεωριών μας και να τις βελτιώσουμε. Η κατασκευή ενός μοντέλου, καθώς αναγκαστικά εμπεριέχει την απλοποίηση του φυσικού συστήματος αναφοράς (του νου), βασίζεται σε ρητή και συγκεκριμένη υλοποίηση κάποιων θεωρητικών αρχών. Η αλληλεπίδραση του ερευνητή με το μοντέλο αποκαλύπτει τις λειτουργικές επιπτώσεις των αρχών αυτών στις προσομοιωμένες πειραματικές συνθήκες. Έτσι, ο ερευνητής κατανοεί καλύτερα την ίδια τη θεωρία του, τις αδυναμίες της, και το ρόλο επιμέρους στοιχείων της. Παράλληλα, μπορεί να παρεμβαίνει διορθωτικά στο μοντέλο, τροποποιώντας τις θεωρητικές προκείμενες, ώστε αυτό να προσεγ-

γίζει καλύτερα τη φυσική συμπεριφορά και να γεννά νέες προβλέψεις για συνθήκες που δεν έχουν ακόμα διερευνηθεί πειραματικά. Έτσι το μοντέλο του γνωσιακού επιστήμονα είναι ένα εργαλείο θεωρητικής διερεύνησης και ανάπτυξης.

Η κατασκευή μοντέλων παίζει κεντρικό ρόλο στη γνωσιακή επιστήμη, σε συνδυασμό με άλλες διερευνητικές και πειραματικές μεθοδολογίες. Επειδή η πολυπλοκότητα του συστήματος-στόχου, δηλαδή του νου, είναι τεράστια, προσπαθούμε να κατανοήσουμε τη λειτουργία του συνδυάζοντας ευρήματα από διάφορες προσεγγίσεις, όπως ψυχολογικά πειράματα, μετρήσεις στο νευρικό σύστημα, φιλοσοφικές και γλωσσολογικές θεωρίες. Ο συνδυασμός υλοποιείται σε συγκεκριμένα μοντέλα όπου προσπαθούμε να περιγράψουμε όσο ακριβέστερα μπορούμε τις επιμέρους νοητικές διεργασίες που πιστεύουμε ότι είναι σημαντικές για κάποιο έργο. Κατασκευάζοντας και χρησιμοποιώντας τα μοντέλα καταλαβαίνουμε καλύτερα τι είδους περαιτέρω πειραματικές διερευνήσεις είναι απαραίτητες, ενώ συγχρόνως μελετάμε τη φύση των περιορισμών του μοντέλου και των ιδιοτήτων που απορρέουν από αυτό. Κάθε μοντέλο είναι αναγκαστικά περιορισμένο στην περιγραφή κάποιας πολύ συγκεκριμένης λειτουργίας ή ενός έργου. Όμως στο πλαίσιο της γνωσιακής επιστήμης μας ενδιαφέρουν τα μοντέλα που κατασκευάζονται στα πλαίσια γενικότερων θεωρήσεων για το νου. Έτσι, διάφοροι ερευνητές χρησιμοποιούν παρόμοιους τρόπους για να κατασκευάσουν μοντέλα για τις διεργασίες που μελετούν, με τελικό στόχο την προσέγγιση ενός «γενικού μοντέλου του νου», δηλαδή την κατανόηση των βασικών αρχών της νοητικής λειτουργίας. Ελπίζουμε με τον τρόπο αυτό να κατανοήσουμε όχι μόνο την εκτέλεση επιμέρους διεργασιών αλλά το συνολικό τρόπο διάρθρωσης και λειτουργίας του νου. Για το λόγο αυτό, αναφερόμαστε στις συνολικές προσεγγίσεις που προσδιορίζουν τις βασικές αρχές κατασκευής μοντέλων με τον όρο «αρχιτεκτονική του νου». Όπως θα δούμε στη συνέχεια, υπάρχουν ορισμένες κύριες εναλλακτικές προσεγγίσεις στην αρχιτεκτονική του νου, οι οποίες έχουν πολύ διαφορετικές θεωρητικές αφετηρίες και πρακτικές συνέπειες, και προσδιορίζουν πολύ διαφορετικούς τύπους μοντέλων.

1.4 Αναπαραστάσεις

Στη γνωσιακή επιστήμη, ο όρος *αναπαράσταση* αναφέρεται σε λειτουργικό χαρακτηριστικό του νου με σημασιακό περιεχόμενο. Ανα-παρίσταται εντός του νου κάτι το οποίο υπάρχει ή συμβαίνει εκτός του νου. Η έννοια της αναπαράστασης είναι γενικότερη από την έννοια του συμβόλου, όπως αυτή χρησιμοποιείται τόσο από τη γλωσσολογία όσο και από την κλασική γνωσιακή επιστήμη, και χρησιμοποιείται περισσότερο και με λιγότερους περιορισμούς. Ανεξάρτητα από γνωσιολογικές δεσμεύσεις για τη φύση του σημασιολογικού περιεχόμενου του νου, οι νοητικές αναπαραστάσεις είναι απαραίτητες για την κατανόηση της αιτιότητας σε νοητικό επίπεδο, αλλά και για την έκφραση των νοητικών γενικεύσεων από την ποικιλομορφία του φυσικού υποβάθρου.

Για παράδειγμα, ας υποθέσουμε ότι επιλέγω μια επαγγελματική απασχόληση επειδή χρειάζομαι χρήματα και πιστεύω ότι αυτή η εργασία παρέχει υψηλές απολαβές και ευκαιρίες αύξησης των αποδοχών με διάφορους τρόπους. Η επιθυμία και η πεποίθηση

αποτελούν νοητικές έννοιες της λεγόμενης «*απλοϊκής ψυχολογίας*» (ή «δημόδους» ή «του κοινού νου», στα αγγλικά folk psychology), της κοινής θεωρίας δηλαδή στο μυαλό του καθενός, οι οποίες «εξηγούν» την επιλογή της εργασίας και τη συμπεριφορά που ακολουθεί. Ο καθένας χρησιμοποιεί τέτοιους όρους για να κατανοήσει, να περιγράψει και να εξηγήσει τη συμπεριφορά τόσο τη δική του όσο και των άλλων. Στον κοινό νου θεωρούμε ότι οι επιθυμίες και οι πεποιθήσεις αποτελούν όχι μόνο αιτιολογικά στοιχεία που δικαιολογούν τη συμπεριφορά αλλά και αιτιακούς παράγοντες που την προκαλούν.

Όμως για ένα μοντέλο του νου δεν είναι εύκολο να περιγραφεί μια αντίστοιχη αιτιακή σχέση που να δικαιολογεί την κανονικότητα της συμπεριφοράς, διότι οι έννοιες που υπεισέρχονται στην εξήγηση δεν έχουν φυσική υπόσταση. Όταν λέμε ότι «επιθυμώ χρήματα» δεν εννοούμε αυτά τα χάρτινα αντικείμενα με την καθορισμένη αξία. Εννοούμε κάτι γενικότερο, αφηρημένο, που μπορεί να περιλαμβάνει οικόπεδα, μετοχές και καταθέσεις σε τράπεζες, γενικά ένα οικονομικό δυναμικό που καθορίζεται περισσότερο από τις δραστηριότητες που επιτρέπει παρά από κάποια συγκεκριμένα απτά χαρακτηριστικά του φυσικού κόσμου. Η έννοια αυτή των «χρημάτων», ή του πλουτισμού γενικότερα, είναι λοιπόν μια γενίκευση πάνω σε ένα ευρύ σύνολο καταστάσεων του φυσικού κόσμου. Επιπλέον, είναι αυτή η γενικευμένη έννοια που δίνει επεξηγηματική ισχύ στο νοητικό «μοντέλο», διότι η επιλογή της εργασίας δεν εξηγείται από τα φυσικά χαρακτηριστικά των χαρτονομισμάτων αλλά από την έννοια του πλούτου ως δυνατότητας. Συνεπώς το νοητικό μοντέλο, για να εξηγήσει ικανοποιητικά τη συμπεριφορά που μας ενδιαφέρει, θα πρέπει να συμπεριλάβει με κάποιον τρόπο τη γενικευμένη αυτή έννοια με την επεξηγηματική ισχύ.

Ακόμα και για πιο απτές περιπτώσεις, συναντάμε την ανάγκη να ορίσουμε τέτοιες νοητικές οντότητες προκειμένου να εξηγήσουμε τις παρατηρούμενες συμπεριφορές. Για παράδειγμα, ας υποθέσουμε ότι κάποιος στρίβει δεξιά σε μια διασταύρωση με πρόθεση να συναντήσει κάποιο φίλο του που βρίσκεται σε εκείνη την κατεύθυνση. Το παρατηρούμενο φαινόμενο είναι η στροφή, ένα αδιαμφισβήτητο φυσικό γεγονός. Ποια είναι η αιτία της στροφής; Δεν μπορεί να είναι ο φίλος, διότι εκείνος βρίσκεται αλλού και δεν μπορεί να δράσει ως ερέθισμα ούτε με φυσική αλληλεπίδραση. Για να εξηγήσουμε το γεγονός αυτό πρέπει να υποθέσουμε ότι στο νου του υποκειμένου μας βρίσκεται κάτι που αντιστοιχεί στο φίλο, το οποίο προκαλεί αιτιακά τη στροφή προς την κατάλληλη κατεύθυνση. Αυτό το κάτι το ονομάζουμε νοητική αναπαράσταση.

Με το σκεπτικό αυτό καταλήγουμε στο ότι για να εξηγήσουμε νοητικές διεργασίες είναι απαραίτητο να συμπεριλάβουμε στα μοντέλα μας αναπαραστάσεις νοητικών εννοιών. Ως αναπαράσταση εδώ θεωρείται μια οντότητα που έχει νοητική υπόσταση και αιτιακή ισχύ μέσα στο μοντέλο, καθώς και σημασιακό περιεχόμενο, δηλαδή αντιστοιχεί σε κάποια συγκεκριμένη έννοια. Η αναπαράσταση των εννοιών στο νου είναι αναπόφευκτη παραδοχή για οποιαδήποτε απόπειρα εξήγησης των νοητικών λειτουργιών που βασίζεται σε αποδοχή των νοητικών περιγραφών. Με την ενσωμάτωση αναπαραστάσεων στο μοντέλο του νου, και κατ'επέκταση με την υπόθεση των αναπαραστάσεων για το νου, διευθετείται θεωρητικά το ζήτημα του εννοιολογικού περιεχομένου. Χάρη στις αναπαραστάσεις δίνεται υπόσταση στις

ουσιώδεις γενικεύσεις του γνωστικού επιπέδου: το «χρήματα» είναι μια αναπαράσταση, μια οντότητα μέσα στο νου, η οποία αντιστοιχεί στις γενικευμένες ιδιότητες του πλούτου και όχι στα φυσικά χαρακτηριστικά των χαρτονομισμάτων, και η οποία μπορεί να συμμετέχει σε διεργασίες που αφορούν στην αντίστοιχη γενικευμένη έννοια. Αντίστοιχα, το «θέλω» είναι μια νοητική κατάσταση ειδικού τύπου η οποία μπορεί να αναφέρεται στην αναπαράσταση «χρήματα» και με τη σχέση αυτή να προκαλεί συμπεριφορές πλουτισμού. Έτσι, χάρη στις αναπαραστάσεις παρέχεται και ένα όχημα για τη σημασιακή ερμηνεία των νοητικών λειτουργιών. Δεν αρκούν όμως οι αναπαραστάσεις από μόνες τους για την κατανόηση των νοητικών διεργασιών. Χρειάζεται και ένας μηχανισμός ο οποίος θα χειρίζεται και θα παράγει αναπαραστάσεις, υλοποιώντας έτσι τις νοητικές λειτουργίες.

Το θέμα του μηχανισμού εμπίπτει στην αρχιτεκτονική δομή του νου και θα μας απασχολήσει στη συνέχεια. Υπάρχουν διάφορες εκδοχές για τη μορφή που μπορεί να έχει ένας τέτοιος μηχανισμός, ενώ από τις υποθέσεις για το μηχανισμό εξαρτάται σε μεγάλο βαθμό και η φυσική υπόσταση των αναπαραστάσεων. Ακολουθώντας θα εξεταστεί η συμβολική αρχιτεκτονική, κατά την οποία ο νους είναι ένας συντακτικός επεξεργαστής συμβόλων. Θα αναπτυχθεί η δυναμική προσέγγιση για το νοητικό σύστημα ως κατεξοχήν μη συμβολική πρόταση. Επίσης θα παρουσιαστούν και θα αντιπαραβληθούν συνδυαστικές προσεγγίσεις για το μηχανισμό του νου οι οποίες καλύπτουν ένα ευρύ φάσμα μοντέλων και εφαρμογών.

1.5 Επίπεδα περιγραφής

Καθώς στη γνωσιακή επιστήμη μας απασχολεί ο νους, δηλαδή το νοητικό «επίπεδο περιγραφής», είναι πολύ σημαντικό να αποσαφηνίσουμε την έννοια του επιπέδου περιγραφής. Αναφορικά με το παράδειγμα του ανθρώπου που στρίβει δεξιά στη γωνία προκειμένου να συναντήσει ένα φίλο του, αντιλαμβανόμαστε ότι συνυπάρχουν πολλές διαφορετικές περιγραφές του φαινομένου. Ένας φυσικός θα μπορούσε να περιγράψει τις θέσεις και τις τροχιές όλων των υποατομικών σωματιδίων (πρωτονίων, νετρονίων, ηλεκτρονίων) που συναπαρτίζουν το παρατηρούμενο σώμα, κάθε χρονική στιγμή. Ή θα μπορούσε να περιγράψει ένα φυσικό σώμα με συγκεκριμένη μάζα και ταχύτητα, το οποίο μεταβάλλει την κατεύθυνση της κίνησής του εξαιτίας των δυνάμεων τριβής που ασκούνται μεταξύ των πελμάτων του και του οδοστρώματος. Ένας βιολόγος θα μπορούσε να περιγράψει κάθε κύτταρο του οργανισμού, συμπεριλαμβάνοντας την έκκριση και απορρόφηση νευροδιαβαστών από τα νευρικά κύτταρα και τις νευρικές ώσεις στον εγκέφαλο και το περιφερικό νευρικό σύστημα που οδηγούν στη σύσπαση των μυϊκών ινών.

Αυτές οι επιστημονικά ακριβείς περιγραφές περιγράφουν το ίδιο ακριβώς φυσικό φαινόμενο από διαφορετικές οπτικές γωνίες. Χρησιμοποιούν διαφορετικές επεξηγηματικές οντότητες και σχέσεις μεταξύ τους, λ.χ. στον κόσμο της υποατομικής φυσικής υπάρχουν μόνο σωματίδια, ενώ στον κόσμο της κυτταρικής βιολογίας υπάρχουν μόνο κύτταρα. Αυτό σημαίνει πως οι περιγραφές της υποατομικής φυσικής μπορούν να δώσουν απαντήσεις μόνο αναφορικά με σωματίδια ενώ οι περιγραφές της κυτταρικής βιολογίας μπορούν να δώσουν απαντήσεις μόνο αναφορικά με κύτταρα. Καμία από αυτές τις εξηγήσεις δεν περι-

λαμβάνει την έννοια του νοήμονος ατόμου ως υποκειμένου και και έτσι, θεμελιωδώς, καμία δεν μπορεί να απαντήσει στο ερώτημα γιατί αυτός έστριψε προς εκείνη την κατεύθυνση.

Γίνεται λοιπόν αντιληπτό ότι από όλες τις ταυτόχρονες σωστές περιγραφές κάποιου φυσικού γεγονότος, στη γνωσιακή επιστήμη μας ενδιαφέρουν εκείνες που αναφέρονται σε νοητικές οντότητες (όπως οι αναπαραστάσεις) και ιδιότητες (όπως οι επιθυμίες και οι πεποιθήσεις). Μόνο τέτοιου είδους εξηγήσεις διατηρούν τις επεξηγηματικά χρήσιμες γενικεύσεις (όπως π.χ. την έννοια των χρημάτων που προαναφέρθηκε). Η επεξηγηματική επάρκεια είναι ο λόγος που δεν μας καλύπτει το φυσικό ή το βιολογικό επίπεδο και δεν μπορούμε να αρκεστούμε σε περιγραφές της εγκεφαλικής λειτουργίας. Το νοητικό επίπεδο περιγραφής είναι ένα θεωρητικό επίπεδο, καθώς οι έννοιες που περιλαμβάνει είναι όλες υποθετικές—δηλαδή, υποθέτουμε την ύπαρξή τους, παρότι δεν μπορούμε να τις παρατηρήσουμε άμεσα, προκειμένου να εξηγήσουμε τις παρατηρούμενες συμπεριφορές.

Με το κριτήριο αυτό μπορούμε να αποφασίσουμε πότε πρέπει να μιλάμε για νου, νόηση και νοήμονα συμπεριφορά: Όταν το νοητικό επίπεδο περιγραφής είναι απαραίτητο για την περιγραφή των φαινομένων με επαρκή γενικευσιμότητα, τότε το χρησιμοποιούμε και θεωρούμε ότι πρόκειται για νόηση. Όταν αρκεί κάποιο χαμηλότερο, απλούστερο επίπεδο περιγραφής, τότε δεν μιλάμε για νόηση και δεν επικαλούμαστε νοητική αιτιότητα. Για παράδειγμα, όταν ο θερμοστάτης ανάβει το καλοριφέρ, δεν χρειάζεται να υποθέσουμε ότι κρύωσε και θέλει να ζεσταθεί. Αρκεί να περιγράψουμε το φυσικό σύστημα (ηλεκτρονικό ή μηχανικό) και τον τρόπο με τον οποίο επηρεάζεται από τη θερμοκρασία. Έτσι, δεν υποθέτουμε ότι η λειτουργία του θερμοστάτη είναι νοητική. Αντίθετα, για να περιγράψουμε πότε ο άνθρωπος του παραδείγματος θα στρίψει δεξιά, δεν μας καλύπτει η περιγραφή του φυσικού συστήματος του σώματός του, αλλά θα πρέπει να γνωρίζουμε αν θέλει να συναντήσει το φίλου του και πού πιστεύει ότι βρίσκεται ο φίλος του. Για την περίπτωση αυτή, λοιπόν, χρησιμοποιούμε εξηγήσεις στο νοητικό επίπεδο, με νοητικές έννοιες, και μιλάμε για νόηση παρότι αναγνωρίζουμε τη χρησιμότητα και την ορθότητα των φυσικών και βιολογικών περιγραφών για άλλους σκοπούς.

1.6 Βούληση και αυτοματισμός

Από την πολυμελετημένη γλωσσική επικοινωνία ως το παρθένο έδαφος της ηθικής, η διεπιστημονική μελέτη των νοητικών φαινομένων έχει στόχο την κατανόηση και περιγραφή τους μέσα από μηχανιστικά μοντέλα. Την κατασκευή δηλαδή τεχνητών συστημάτων, βάσει των αντίστοιχων θεωριών, τα οποία υλοποιούν τις κατάλληλες αναπαραστάσεις και διεργασίες ώστε να επιδεικνύουν εξωτερικά τη «σωστή» συμπεριφορά, περιγράφοντας ικανοποιητικά τη λειτουργία του νου. Εδώ μπορεί να τεθεί το ερώτημα κατά πόσον ένα τέτοιο σύστημα μπορεί να προσεγγίσει συνολικά το νου ή αν αφήνει υποχρεωτικά εκτός σημαντικές όψεις και εκφάνσεις του. Αν η αδυναμία πλήρους κατανόησης είναι εκ των προτέρων δεδομένη, ίσως θα έπρεπε να αναζητήσουμε άλλες μεθοδολογίες και να στραφούμε σε αναζήτηση άλλων τρόπων μελέτης και περιγραφής του νου. Μια πιθανή αντίρρηση μπορεί να προέλθει από δυϊστικού τύπου αφετηρίες. Να θεωρούμε δηλαδή το νου ως μη φυσική οντότητα, ως

«πνεύμα» ή «ψυχή», κάτι εκτός του φυσικού κόσμου που δεν υπόκειται στους φυσικούς νόμους και δεν μπορεί να γίνει κατανοητό με επιστημονικές μεθόδους. Μια τέτοια θέση δεν μπορεί βεβαίως να αντιμετωπιστεί στα πλαίσια του επιστημονικού διαλόγου. Απλά, όποιος το πιστεύει αυτό δεν ασχολείται με τη γνωσιακή επιστήμη.

Υπάρχουν όμως και διαφορετικού τύπου αντιρρήσεις, λιγότερο ριζικές αλλά εξίσου σημαντικές για τη μελέτη του νου. Η «ελεύθερη βούληση» είναι σίγουρα μια τέτοια περίπτωση. Λέγοντας «ελεύθερη βούληση» στον κοινό λόγο απηχούμε συνήθως τον αποτροπιασμό μας για το ενδεχόμενο ο νους να είναι «απλώς» μια μηχανή, δηλαδή να είναι ένα σύστημα που λειτουργεί αιτιοκρατικά με βάση τους φυσικούς νόμους και μόνο, και του οποίου τη συμπεριφορά θα μπορούσαμε να προβλέψουμε πλήρως αν γνωρίζαμε σε επαρκή λεπτομέρεια όλες τις φυσικές παραμέτρους του. Ίσως πιστεύουμε ότι αν ο νους εμπεριέχει αιτιακά τη μελλοντική του λειτουργία τότε είναι λιγότερο ελεύθερος και άρα εμείς λιγότερο σημαντικοί ως αυθύπαρκτες οντότητες και ενδεχομένως λιγότερο υπεύθυνοι για τις πράξεις μας: «δεν το ήθελα κύριε πρόεδρε, ο εγκέφαλός μου το έκανε»! Επειδή η αντίρρηση αυτή είναι διαδεδομένη όσο και παραπλανητική, αξίζει να την εξετάσουμε.

Καταρχήν ως συμφωνήσουμε ότι «βούληση» είναι η υποκειμενική αίσθηση ότι μια πράξη μας είναι συνεπής με την πρόθεσή μας. Με απλούστερα λόγια, βούληση ασκούμε όταν νιώθουμε πως κάνουμε κάτι επίτηδες. Είναι μια αναμφισβήτητη έκφανση της νοητικής φαινομενολογίας, δεδομένου ότι συνήθως δεν έχουμε καμία αμφιβολία αν π.χ. κουνήσαμε το χέρι μας επίτηδες ή μας το κούνησε κάποιο εξωτερικό αίτιο (σπρώξιμο) ή άλλη αιτία χωρίς δική μας πρόθεση (αθέλητη μυϊκή σύσπαση). Η ιδέα ότι κάνουμε κάτι επίτηδες είναι θεμελιώδης για την αυτοαντίληψή μας καθώς και για την ηθική μας υπόσταση στα πλαίσια της ατομικής υπευθυνότητας. Σε διαγραμματική μορφή, μια εμπρόθετη πράξη χαρακτηρίζεται στην απλοϊκή κοινή θεώρηση από ένα σχήμα σαν αυτό:

Βούληση → Άλλες νοητικές διεργασίες → Πράξη

Αντίθετα, μια πράξη που έγινε «αυτόματα» (χωρίς τη βούλησή μας αλλά όχι απαραίτητα παρά τη βούλησή μας), μπορεί να περιγραφεί απλούστερα:

Άλλες νοητικές διεργασίες → Πράξη

Αν έχουν έτσι τα πράγματα, τότε το πεδίο της γνωσιακής επιστήμης ως μελέτης των νοητικών διεργασιών θα πρέπει να περιοριστεί στην επεξεργασία του αποτελέσματος των βουλητικών επιλογών και στη μελέτη των αυτόματων πράξεων. Τι μπορεί να είναι όμως η βούληση, αν δεν είναι και αυτή προϊόν μηχανιστικών νοητικών διεργασιών; Ας απορρίψουμε καταρχήν, ως εκτός επιστήμης, το δυϊστικό ενδεχόμενο να είναι κάποιο ψυχικό φαινόμενο έξω από την ύλη. Η φυσική όχι μόνο δεν αναγνωρίζει «ψυχικές» δυνάμεις αλλά δεν αφήνει και κανένα περιθώριο ως προς τη δυνατότητα νέων δυνάμεων να επηρεάζουν την ύλη. Η κατανόηση και περιγραφή των φυσικών φαινομένων στο μακροσκοπικό επίπεδο από την επιστήμη είναι αρκετά πλήρης και η πιθανότητα να επηρεάζεται η ύλη από κάτι που μένει ακόμα άγνωστο είναι μικρή. Όμως οι νοητικές διεργασίες είναι εγκεφαλική

δραστηριότητα, στο φυσικό επίπεδο, και οδηγούν σε παρατηρήσιμες πράξεις, επίσης στο φυσικό επίπεδο. Συνεπώς δεν μπορούν να προκληθούν από μη φυσικές «πνευματικές» δυνάμεις, εκτός αν καταρριφθεί όλη η φυσική.

Η βούληση λοιπόν δεν μπορεί να προέρχεται έξω από το φυσικό κόσμο. Μέσα στο φυσικό κόσμο υπάρχει το μηχανιστικό σύστημα του νου, το οποίο υλοποιείται στο φυσικό (και μηχανιστικό) σύστημα του εγκεφάλου. Το σύστημα του νου αναπτύσσεται μέσα στο άτομο στο οποίο εντάσσεται και συγκεντρώνει το αποτέλεσμα των εμπειριών του ατόμου αυτού, βάσει των οποίων προγραμματίζει μηχανιστικά τις σκέψεις και τις πράξεις του και, γενικότερα, τις αλληλεπιδράσεις του με το περιβάλλον. Οι ιδιότητες και οι συγκεκριμένες ενέργειες του μηχανιστικού νου αντιστοιχούν απολύτως στα συγκεντρωτικά χαρακτηριστικά του ατόμου και είναι συνεπείς με το γενικότερο χαρακτήρα του, το παρελθόν του, τις ιδιαιτερότητες, τις εμπειρίες και τις προτιμήσεις του. Συνεπώς, ο μηχανιστικός νους είναι μια ιδανική πηγή βούλησης για το άτομο, καθώς εξασφαλίζει ότι οι «επίτηδες» πράξεις του θα είναι απόρροια της σύστασης, της εμπειρίας και της ιδιοσυγκρασίας του, άρα θα μπορεί ως άτομο και ως υποκείμενο να νιώθει υπευθυνότητα γι' αυτές.

Τι εναλλακτικές περιπτώσεις υπάρχουν; Συχνά θεωρούμε την ελευθερία της βούλησης ως αποδέσμευση από το προκαθορισμένο. Όμως το προκαθορισμένο αντανάκλα τον εαυτό, ενώ οποιαδήποτε αποδέσμευση από αυτό θα είναι είτε τυχαία είτε ακατάλληλη. Θέλουμε για ελεύθερη βούληση κάτι που δεν εκφράζει το χαρακτήρα μας, όπως αυτός έχει διαμορφωθεί από την ανάπτυξη και την εμπειρία μας; Θέλουμε την ελευθερία μας να είναι ένα καπρίτσιο της τυχειότητας; Προφανώς όχι. Στην πραγματικότητα κανείς δεν θέλει την ελεύθερη βούληση να είναι απρόβλεπτη, διότι τότε δεν τεκμηριώνεται ακεραιότητα του χαρακτήρα, ούτε να είναι τυχαία, διότι τότε πάει περίπατο η ατομική υπευθυνότητα. Θέλουμε την ελεύθερη βούληση να είναι ακριβώς το προϊόν των ατομικών νοητικών διεργασιών, και όσο πιο μηχανιστικές είναι αυτές τόσο καλύτερα εκφράζουν αυτό που φανταζόμαστε ότι είναι ο εαυτός μας.

Κατανοώντας ότι η περιγραφή του νου με μηχανιστικά μοντέλα δεν μειώνει ούτε τα άτομα ούτε τις επιλογές τους, γίνεται σαφές ότι το δυνητικό επεξηγηματικό πεδίο της γνωσιακής επιστήμης μπορεί να συμπεριλάβει όλες τις νοητικές διεργασίες και τις αιτιότητες που αποδίδονται σε αυτές. Έτσι η διερεύνηση των «αυτόματων» νοητικών διεργασιών δεν είναι μια άσκηση σε επιμέρους, κι ενδεχομένως ασήμαντες, ιδιότητες του νου, αλλά η δυνατότητα προσέγγισης του κεντρικού χαρακτήρα και όλων των εκφάνσεων και λειτουργιών της νόησης.

Θεωρητικές προσεγγίσεις

2 Κλασική (συμβολική) γνωσιακή επιστήμη

Η κλασική γνωσιακή επιστήμη θέτει συγκεκριμένες προδιαγραφές για τη φύση των αναπαραστάσεων και τη σχέση τους με τους μηχανισμούς που τις χειρίζονται και τις παράγουν. Συγκεκριμένα, προτείνει ότι οι αναπαραστάσεις είναι αφηρημένα σύμβολα, τα οποία έχουν δύο ειδών χαρακτηριστικά: τα *συντακτικά*, με βάση τα οποία γίνεται η επεξεργασία τους μέσα στο σύστημα, και τα *σημασιακά*, τα οποία έχουν να κάνουν με αυτό που συμβολίζει η αναπαράσταση. Η κρίσιμη διάκριση είναι ότι το σημασιακό περιεχόμενο είναι λειτουργικά αδρανές, δηλαδή το συμβολικό νοητικό σύστημα δεν επεξεργάζεται νοητικές σημασίες αλλά νοητικές αλληλουχίες τυπικών συμβόλων. Το αποτέλεσμα είναι μια θεώρηση του νου που μοιάζει πολύ με τον ηλεκτρονικό υπολογιστή ή, καλύτερα, με τη θεωρητική υπολογιστική μηχανή του Turing και τις διάφορες ισοδύναμες εκδοχές της. Η ουσία της υπόθεσης ότι ο νους είναι υπολογιστής αφορά στις συντακτικές ιδιότητες των αναπαραστάσεων-συμβόλων, χάρη στις οποίες αυτά τα σύμβολα αλληλεπιδρούν και μετέχουν σε διεργασίες επεξεργασίας τους, ενώ το σημασιακό τους περιεχόμενο αποτελεί προϊόν ερμηνείας από τον παρατηρητή του συστήματος που έχει πρόσβαση στο εξωτερικό περιβάλλον του νου και στις σχέσεις μεταξύ του νου και του περιβάλλοντος αυτού.

2.1 Σύμβολα και υπολογισμοί

Ας υποθέσουμε⁴ ότι κατασκευάζουμε μια μηχανή η οποία διαθέτει τέσσερις καταχωρητές, δηλαδή θέσεις προσωρινής αποθήκευσης, και ότι σε καθένα καταχωρητή μπορεί ανά πάσα στιγμή να υπάρχει είτε ένα \times είτε ένα \circ είτε τίποτα. Η μηχανή αυτή είναι έτσι σχεδιασμένη ώστε να ακολουθεί τους εξής κανόνες:

- K1: Αν στον καταχωρητή 1 υπάρχει \times και στον καταχωρητή 2 υπάρχει \circ τότε στον καταχωρητή 3 τοποθετείται ένα \times και στον καταχωρητή 4 τοποθετείται ένα \circ .
- K2: Αν στον καταχωρητή 1 υπάρχει \circ και στον καταχωρητή 2 υπάρχει \times τότε στον καταχωρητή 3 τοποθετείται ένα \times και στον καταχωρητή 4 τοποθετείται ένα \circ .
- K3: Αν στον καταχωρητή 1 υπάρχει \times και στον καταχωρητή 2 υπάρχει \times τότε στον καταχωρητή 3 τοποθετείται ένα \circ και στον καταχωρητή 4 τοποθετείται ένα \times .
- K4: Αν στον καταχωρητή 1 υπάρχει \circ και στον καταχωρητή 2 υπάρχει \circ τότε στον καταχωρητή 3 τοποθετείται ένα \circ και στον καταχωρητή 4 τοποθετείται ένα \circ .

Είναι σημαντικό να αναγνωρίσουμε εξαρχής ότι η μηχανή αυτή μπορεί να είναι υδραυλική, μηχανική, ηλεκτρομηχανική, οπτοηλεκτρονική, ένα θεωρητικό σχέδιο στο τετράδιο ή ένα πρόγραμμα στον υπολογιστή, χωρίς να αλλάζει τίποτα από την ουσία του ζητήματος. Η

⁴Το παράδειγμα βασίζεται στο Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

ακριβής μορφή και τυχόν φυσική υπόσταση των \times και \circ δεν επηρεάζει ούτε την ισχύ ούτε την ενδεχόμενη μετέπειτα ερμηνεία των κανόνων αυτών.

Παρατηρώντας το σύστημα αυτό, διαπιστώνουμε ότι, αντικαθιστώντας το \times με 1 και το \circ με 0, η μηχανή αυτή υλοποιεί την πρόσθεση μονοψήφιων αριθμών στο δυαδικό σύστημα: $1 + 0 = 01$, $0 + 1 = 01$, $0 + 0 = 00$, $1 + 1 = 10$.⁵ Για να ισχύει αυτή η ερμηνεία, εκτός από την αντικατάσταση των συμβόλων, πρέπει να αντιστοιχίσουμε τους καταχωρητές 1 και 2 στους προσθετέους, τον καταχωρητή 3 στο μονοψήφιο άθροισμα και τον καταχωρητή 4 στο κρατούμενο. Μέσα από αυτό το πλέγμα αντιστοιχιών προς το σύστημα δυαδικής πρόσθεσης προκύπτει η *ερμηνεία* της λειτουργίας του συστήματος. Υπό αυτήν την έννοια, μπορούμε να πούμε ότι το \times αναπαριστά τη μονάδα και το \circ αναπαριστά το μηδέν, είναι όμως ιδιαίτερα σημαντικό ότι δεν υπάρχει καμία ιδιότητα των \times και \circ που να αντιστοιχεί στα 1 και 0 εκτός από τις σχέσεις μεταξύ τους και με τη λειτουργία του συστήματος. Θα μπορούσε εξίσου (με αντίστοιχη τροποποίηση των κανόνων K1–K4) το \times να αναπαριστά το μηδέν και το \circ τη μονάδα, ή να έχουμε δύο άλλα, εντελώς διαφορετικά «σύμβολα».

Το σύστημα αυτό λειτουργεί σύμφωνα με τους τέσσερις κανόνες, ανεξάρτητα από οποιαδήποτε ερμηνεία μπορεί να δοθεί στη λειτουργία του, λαμβάνοντας υπόψη μόνο ιδιότητες των συμβόλων που αφορούν στις σχέσεις μεταξύ τους και με τους κανόνες. Αυτές οι—μη σημασιακές—ιδιότητες των συμβόλων, που έχουν να κάνουν με το πώς αυτά συν-τάσσονται και επηρεάζουν τη λειτουργία του συστήματος, ονομάζονται *συντακτικές ιδιότητες*. Κατ'επέκταση και το σύστημα αυτό ονομάζεται σύστημα συντακτικής συμβολικής επεξεργασίας, εφόσον η λειτουργία του βασίζεται σε συντακτικές ιδιότητες των αναπαραστάσεών του και όχι στο σημασιακό τους περιεχόμενο. Οι αναπαραστάσεις οι ίδιες ονομάζονται σύμβολα, καθιστώντας σαφέστερο ότι ο ρόλος τους στη λειτουργία του συστήματος είναι συντακτικός και όχι αναπαραστασιακός με αναφορά στο περιεχόμενο. Γενικότερα, οι αναπαραστάσεις στη συμβολική προσέγγιση του νου ονομάζονται «σύμβολα». Τα σύμβολα έχουν συντακτικές ιδιότητες, βάσει των οποίων γίνονται οι υπολογισμοί, καθώς και σημασιακό περιεχόμενο, το οποίο δεν επηρεάζει τους υπολογισμούς. Λέγοντας «υπολογισμοί», εννοούμε κάθε επεξεργασία των συμβόλων, δηλαδή την εφαρμογή και εκτέλεση των συντακτικών κανόνων του συστήματος. Συνεπώς, ένα συμβολικό σύστημα περιλαμβάνει σύμβολα και κανόνες, και περιγράφει τη νοητική λειτουργία υπολογιστικά, ως εφαρμογή των κανόνων στα σύμβολα. Οι νοητικές διεργασίες περιγράφονται ως αλληλουχίες εκτέλεσης κανόνων, ενώ οι νοητικές αναπαραστάσεις περιγράφονται ως σύμβολα.

⁵Στο δεκαδικό σύστημα, που χρησιμοποιούμε συνήθως για να συμβολίσουμε αριθμητικές ποσότητες με δέκα ψηφία (0 έως 9), κάθε ψηφίο εκφράζει πλήθος μιας δύναμης του 10: από δεξιά, το πρώτο ψηφίο λέει πόσες μονάδες ($1 = 10^0$), το δεύτερο πόσες δεκάδες ($10 = 10^1$), το τρίτο εκατοντάδες ($100 = 10^2$), το τέταρτο χιλιάδες ($1000 = 10^3$) κ.ο.κ. Αντίστοιχα, στο δυαδικό σύστημα, που χρησιμοποιείται στους ηλεκτρονικούς υπολογιστές και έχει μόνο δύο ψηφία (0 και 1), κάθε ψηφίο εκφράζει πλήθος μιας δύναμης του 2: το δεξιότερο ψηφίο λέει πόσες μονάδες ($1 = 2^0$), το δεύτερο πόσες δυάδες ($2 = 2^1$), το τρίτο τετράδες ($4 = 2^2$), το τέταρτο οχτάδες ($8 = 2^3$) κ.ο.κ. Έτσι, για παράδειγμα ο αριθμός δύο στο δυαδικό σύστημα είναι 10 (μία δυάδα και μηδέν μονάδες) ενώ ο αριθμός πέντε είναι 101 (μία τετράδα, καμία δυάδα, και μία μονάδα).

Η δικαιολόγηση της ερμηνείας των συμβόλων είναι προς το παρόν φιλοσοφικό ζήτημα. Μπορεί να είναι εντελώς εξωγενής και ανεξάρτητη από το ίδιο το υπολογιστικό σύστημα. Μπορεί να εντοπίζεται στη *σχέση αντιστοιχίας* μεταξύ των μεμονωμένων συμβόλων και των εννοιών που αναπαριστώνται από τα σύμβολα (εξωτερική αναφορά). Μπορεί ακόμα να συγκροτείται από το συνολικό *πλέγμα σχέσεων* μεταξύ των συμβόλων, οι οποίες θα πρέπει να αντανakλούν αντίστοιχες σχέσεις μεταξύ των αναπαριστώμενων εννοιών (εσωτερική αναφορά). Η λειτουργία του συστήματος δεν προσδιορίζεται ούτε εξαρτάται από τη φύση και την αιτιακή προέλευση της ερμηνείας. Δηλαδή, δεν έχει σημασία για τη λειτουργία ούτε τι είναι αυτό που προσδίδει σημασιακό περιεχόμενο σε ένα σύμβολο ούτε ποια είναι η (εξελικτική, αναπτυξιακή ή μαθησιακή) διαδικασία μέσω της οποίας ένα σύμβολο αποκτά αυτό το σημασιακό περιεχόμενο.

2.2 Συμβολικές καταστάσεις

Λέγοντας ότι ένα υπολογιστικό σύστημα βρίσκεται σε μια «κατάσταση» εννοούμε ότι υπάρχει μια σύνθετη εξάρτηση ανάμεσα στη μέχρι τώρα πορεία και τη μελλοντική λειτουργία. Η εξάρτηση αυτή δεν είναι απλή αλληλουχία, δηλαδή η κατάσταση δεν αντιστοιχεί σε κάποιο συγκεκριμένο βήμα μιας προκαθορισμένης σειράς, αλλά μπορεί να συμπεριλάβει επεξεργασία εξωτερικών ερεθισμάτων ή αλληλεπίδραση με άλλες καταστάσεις. Ένα απλό παράδειγμα υπολογιστικής κατάστασης είναι ο τρόπος ρύθμισης των ψηφιακών ρολογιών που έχουμε στα σπίτια. Συνήθως υπάρχει ένα κουμπί αύξησης (+) κι ένα κουμπί μείωσης (-). Το τι αυξάνεται ή μειώνεται (λ.χ. η ώρα ή τα λεπτά) εξαρτάται από κάτι άλλο: Σε κάποιες συσκευές κρατάμε πατημένο ένα κουμπί για τη ρύθμιση της ώρας και ένα άλλο για τη ρύθμιση των λεπτών (ή για τη ρύθμιση της τρέχουσας ώρας και του ζυπνητηριού). Σε άλλες συσκευές πατάμε μια φορά ένα κουμπί για να ρυθμίσουμε την ώρα και δεύτερη φορά για να ρυθμίσουμε τα λεπτά. Σε κάθε περίπτωση, η ανταπόκριση στο κουμπί-ερέθισμα αύξησης ή μείωσης εξαρτάται από την υπολογιστική κατάσταση στην οποία έχουμε προηγουμένως θέσει το σύστημα με κάποιον άλλο τρόπο. Σε ένα γενικό υπολογιστικό σύστημα δεν υπάρχει περιορισμός στο πλήθος των πιθανών καταστάσεων ούτε στις σχέσεις μεταξύ καταστάσεων και ανταπόκρισης σε ερεθίσματα.

Η τρέχουσα κατάσταση ενός νοητικού συστήματος προσδιορίζει τον τρόπο ανταπόκρισης σε ερεθίσματα, για παράδειγμα πώς αντιδρούμε αν χτυπήσει τηλέφωνο σε ώρα μαθήματος ή σε ώρα ξεκούρασης. Η κατάσταση όμως δεν έχει να κάνει με το περιβάλλον αλλά με το ίδιο το σύστημα. Έτσι, ένα καλύτερο παράδειγμα είναι η διάθεση: αλλιώς ανταποκρινόμαστε σε ένα ειρωνικό σχόλιο όταν είμαστε σε εύθυμη διάθεση και αλλιώς όταν είμαστε εκνευρισμένοι. Στην υπολογιστική προσέγγιση του νου, ιδιαίτερο ενδιαφέρον έχει προσελκύσει μια οικογένεια νοητικών καταστάσεων που προσδιορίζεται από σχέση με συγκεκριμένο περιεχόμενο. Όπως αναφέρθηκε στην ενότητα 1.4 περί αναπαραστάσεων, μια τέτοια κατάσταση είναι η επιθυμία. Ας σκεφτούμε πώς μπορούμε να ορίσουμε λειτουργικά την έννοια της επιθυμίας ώστε να προκύπτουν οι κατάλληλοι συντακτικοί υπολογισμοί. Όπως είδαμε, το «θέλω» είναι μια νοητική κατάσταση που αναφέρεται σε κάποια αναπαρά-

σταση, δηλαδή στην πρόταση για την οποία ισχύει, π.χ. «θέλω να φάω», «θέλω ένα κόκκινο τραπέζι», ή «θέλω να γίνω πλούσιος». Μπορούμε λοιπόν να πούμε ότι για κάποιες προτάσεις π κατασκευάζουμε προτάσεις της μορφής «θέλω π », οι οποίες προδίδουν τη στάση (επιθυμίας) προς τις προτάσεις π και γι' αυτό ονομάζονται *προτασιακές στάσεις*. Πρόκειται για δήλωση μιας συγκεκριμένης *νοητικής σχέσης* του υποκειμένου προς την πρόταση π , άρα οι προτασιακές στάσεις είναι *καταστάσεις* του νου. Το ίδιο ισχύει, για παράδειγμα, και στην περίπτωση που κάποιος «πιστεύει π », δηλαδή αποτελεί το υποκείμενο μιας νοητικής σχέσης τύπου πεποίθησης ως προς την πρόταση π .

Ένας λειτουργικός ορισμός της προτασιακής στάσης «θέλω π » μπορεί να δοθεί απλουστευτικά ακολουθώντας μια πορεία ως εξής:

$$\text{θέλω } \pi \equiv [\text{πιστεύω ότι } \rho \Rightarrow \pi] \& [\text{κάνω } \rho].$$

Δηλαδή το να θέλω κάτι ισοδυναμεί με το να πράττω εκείνα τα οποία πιστεύω ότι θα το επιφέρουν.⁶ Για να μπορεί να σταθεί ένας τέτοιος ορισμός χρειάζεται και περιφερειακή υποστήριξη από πλήθος διευκρινίσεων, και την εξασφάλιση ότι οι συνέπειες του ρ περιορίζονται στο π ή είναι κατά τα άλλα ισοδύναμες με τις συνέπειες του όχι- ρ κλπ. Σε κάθε περίπτωση καταλήγουμε σε έναν ορισμό όπου για να εξηγήσουμε μια νοητική έννοια χρησιμοποιούμε μια άλλη νοητική έννοια. Στη συγκεκριμένη περίπτωση για τον ορισμό του «θέλω» εισάγουμε το «πιστεύω», δηλαδή εισάγουμε μια νέα προτασιακή στάση, μια νέα νοητική αδιευκρίνιστη έννοια, για να περιγράψουμε την αρχική. Η νέα αυτή νοητική έννοια θα απαιτήσει με τη σειρά της άλλες νοητικές έννοιες, ή ακόμα και την προηγούμενη (κυκλικά συν-οριζόμενη), για να προσδιοριστεί, κ.ο.κ. Με άλλα λόγια, υπάρχει μια κατηγορία νοητικών διεργασιών και καταστάσεων, οι προτασιακές στάσεις, οι οποίες δεν είναι δυνατόν να αναχθούν σε μη νοητικούς όρους και γι' αυτό αποτελούν κομβικό σημείο αναφοράς στη μελέτη του νου. Έτσι, τις ορίζουμε «λειτουργικά», μέσα από το πλέγμα των σχέσεων μεταξύ τους.

Τον ορισμό του «θέλω» θα μπορούσαμε να τον συμβολίσουμε με την έκφραση

$$\Theta(x) \equiv \Pi(y \Rightarrow x) \wedge K(y).^7$$

Γραμμένη έτσι, η έκφραση αυτή καθιστά σαφή την έννοια του συμβολισμού σε δύο σημαντικά επίπεδα:

- Πρώτον, για οποιοδήποτε προτάσεις-αναπαραστάσεις που έχουν τις κατάλληλες (συντακτικές) ιδιότητες ώστε να «ταιριάζουν» στις θέσεις x και y , η έκφραση αυτή ισχύει αυτόματα με τον ίδιο τρόπο και κατά την ίδια έννοια. Το αν ταιριάζουν δεν είναι θέμα περιεχομένου (ερμηνείας), δηλαδή δεν εξαρτάται από το τι σημαίνουν, τι

⁶Αγνοώντας προς το παρόν τις διάφορες επιπλοκές που μπορεί να συνδέονται με τις συγκεκριμένες πράξεις, καθώς και πολλές άλλες «ενοχλητικές λεπτομέρειες», ώστε να επικεντρώσουμε στην ουσία του παραδείγματος. Στη φιλοσοφία του νου ένας τέτοιος ορισμός θα περιλάμβανε οπωσδήποτε και τη δήλωση *ceteris paribus*, δηλαδή υπό την προϋπόθεση ότι δεν συντρέχουν άλλοι λόγοι ή ιδιαίτερες περιστάσεις.

⁷Με το σύμβολο \wedge αναπαριστούμε τη σύζευξη, δηλαδή το «και» – δηλαδή ότι οι παραστάσεις εκατέρωθεν του συμβόλου πρέπει να ισχύουν συγχρόνως και οι δύο.

αναπαριστούν, τα x και y . Θα μπορούσε να οριστεί ένας κατάλογος x -συμβατών παραστάσεων και ένας άλλος y -συμβατών παραστάσεων, οι οποίοι μπορεί και να έχουν κοινά στοιχεία.

- Δεύτερον, η ισχύς της έκφρασης δεν εξαρτάται από τι ερμηνεύουμε ως σημασιακό περιεχόμενο των Θ , Π και K . Όσον αφορά το σύστημα στο οποίο αναπαρίσταται η έκφραση αυτή, πρόκειται για μια αληθή σχέση μεταξύ συμβόλων. Αν μας δοθούν οι κατάλογοι των παραστάσεων που είναι συμβατές με τα x και y ώστε να διατηρείται η αλήθεια της έκφρασης τότε ίσως παρατηρήσουμε ότι η έκφραση αυτή μπορεί να αναπαριστά μια απλοποιημένη περιγραφή της επιθυμίας, ότι δηλαδή Θ =«θέλω», Π =«πιστεύω» και K =«κάνω». Η ερμηνεία αυτή είναι εξωτερική όσον αφορά στην έκφραση, όπως και η ερμηνεία των \times και \circ ως δυαδικών ψηφίων δεν ήταν εγγενής ιδιότητα του συστήματος καταχωρητών.

Η έκφραση αυτή λοιπόν δίνει μια απλουστευμένη ιδέα του συμβολισμού: Δυο σύμβολα (x και y) συμμετέχουν σε μια σχέση βάσει των ιδιοτήτων τους και η σχέση αυτή ορίζει μια νοητική κατάσταση. Η νοητική κατάσταση έχει αιτιακή ισχύ διότι μπορεί να προκαλέσει άλλες νοητικές καταστάσεις ή συμπεριφορές οι οποίες εμπλέκουν τα ίδια σύμβολα χωρίς αναφορά στο σημασιακό τους περιεχόμενο. Ο νους, υπό το πρίσμα αυτό, είναι ένας συντακτικός υπολογιστής ο οποίος επεξεργάζεται (ή απλώς μεταχειρίζεται) σύμβολα μεταβαίνοντας από μια νοητική κατάσταση σε άλλη, όπως και ένας ηλεκτρονικός υπολογιστής επεξεργάζεται στοιχεία συμβολικής πληροφορίας σύμφωνα με τον προγραμματισμό του και τις διαθέσιμες καταστάσεις του. Τελικά ο ηλεκτρονικός υπολογιστής εκτελεί έργα χρήσιμα για μας επειδή ισχύει μια συνεπής αντιστοιχισμός μεταξύ των καταστάσεών του και ερμηνειών που μας αφορούν (π.χ. αριθμοί, γράμματα, πράξεις, σχέδια κλπ.). Αντίστοιχα, ο νους εκτελεί νοητικές διεργασίες που έχουν νόημα στο συγκεκριμένο περιβάλλον του οργανισμού διότι βρίσκονται σε μια συνεπή σχέση με τις κατάλληλες ιδιότητες του περιβάλλοντος, χωρίς οι σημασίες καθαυτές να παίζουν ρόλο στην εσωτερική επεξεργασία.

Έτσι, μια ακόμα παρατήρηση για το σύστημα «πρόσθεσης μονοψηφίων δυαδικών αριθμών» είναι ο διαχωρισμός των επιπέδων (βλ. Ενότητα 1.5). Η φυσική αιτιοκρατία που κινεί το σύστημα πρέπει να αναζητηθεί στις φυσικές ιδιότητες του υποβάθρου (μηχανικού, ηλεκτρικού κλπ.). Όμως η λειτουργία του συστήματος περιγράφεται στο επίπεδο του αλγορίθμου, βάσει των τεσσάρων κανόνων, διότι το φυσικό υπόβαθρο είναι έτσι διευθετημένο ώστε οι κανόνες αυτοί να έχουν αιτιακή ισχύ. Πάνω κι απ' το αλγοριθμικό επίπεδο, η σημασία που εμείς δίνουμε στη λειτουργία του συστήματος (η πρόσθεση) είναι κάτι που δεν βρίσκεται στο σύστημα ούτε αποτελεί ιδιότητά του, αλλά δίνεται από τον παρατηρητή του συστήματος. Αυτή είναι ουσιαστικά και η θέση της κλασικής (υπολογιστικής, συμβολικής) γνωσιακής επιστήμης για τη λειτουργία του νου: το φυσικό υπόβαθρο είναι άσχετο με τις νοητικές λειτουργίες, το αλγοριθμικό επίπεδο νοητικής λειτουργίας είναι συμβολικό και υπολογιστικό με τη συντακτική έννοια, και η σημασία των συμβόλων του νου είναι αναδύομενη ιδιότητα που αποδίδουμε ερμηνευτικά στο νου, παρατηρώντας τη λειτουργία του μέσα στο περιβάλλον με το οποίο αλληλεπιδρά.

2.3 Παραγωγικότητα και συνδυαστικότητα

Μια βασική και ευρέως αποδεκτή παρατήρηση για το νου είναι ότι χαρακτηρίζεται από *συστηματικότητες*, δηλαδή ομοιότητες μεταξύ οντοτήτων και κανονικότητες φαινομένων που δεν μπορεί να είναι τυχαίες. Η παρατηρούμενη συστηματικότητα πρέπει να αποδοθεί θεωρητικά σε χαρακτηριστικά της δομής και λειτουργίας του νου. Η συμβολική προσέγγιση αποδίδει τη συστηματικότητα στην αφηρημένη φύση των συμβολικών υπολογισμών, και ειδικότερα σε δύο συγκεκριμένες ιδιότητες των συντακτικών επεξεργαστών, την παραγωγικότητα και τη συνδυαστικότητα.⁸

Με τον όρο *παραγωγικότητα* αναφερόμαστε στην ιδιότητα του νου να εφαρμόζει αφηρημένους *κανόνες* για να μεταβαίνει από ένα σύμβολο ή αλληλουχία συμβόλων σε μια άλλη. Η παραγωγικότητα είναι η υποθετική ιδιότητα του νου που επιτρέπει απεριόριστο αριθμό παραστάσεων από περιορισμένο αριθμό *ατομικών*⁹ συμβόλων. Η υπόθεση της παραγωγικότητας χρησιμοποιείται από τους θεωρητικούς γλωσσολόγους για να εξηγήσει την ικανότητα παραγωγής απεριόριστου αριθμού προτάσεων από ένα περιορισμένο σύνολο λημμάτων (λέξεων, ή καλύτερα *μορφημάτων*). Η ιδιότητα αυτή έχει νόημα όταν συνυπάρχει με την ιδιότητα σύνθεσης αναπαραστάσεων για το σχηματισμό σύνθετων αναπαραστάσεων από απλούστερες. Έτσι θεωρούμε ότι υπάρχει ένα περιορισμένο σύνολο ατομικών συμβόλων και ένα απεριόριστο σύνολο μη ατομικών αναπαραστάσεων, οι οποίες παράγονται από τις ατομικές με τους επιτρεπτούς συνδυασμούς τους.

Η έννοια του κανόνα είναι ουσιώδης και αποτελεί συχνό σημείο αντιπαράθεσης μεταξύ συμβολικών και μη συμβολικών συστημάτων. Αναφέρεται στο μηχανισμό που είναι υπεύθυνος για την εμφάνιση μιας κανονικότητας. Για παράδειγμα, στο σύνολο λέξεων ενικού αριθμού {παιδί, λουρί, κουμπί, φλουρί} αντιστοιχεί ένα σύνολο λέξεων πληθυντικού αριθμού: {παιδιά, λουριά, κουμπιά, φλουριά}. Στην αντιστοιχία αυτή υπάρχει μια προφανής κανονικότητα, σαν να αντικαθίσταται ένα μέρος της λέξης ενικού από κάτι άλλο για να σχηματιστεί η λέξη πληθυντικού. Η κανονικότητα αυτή μπορεί να εκφραστεί με έναν αφηρημένο γενικό κανόνα της μορφής [$w_i \Rightarrow w_i\acute{\alpha}$], όπου με w συμβολίζουμε μια *μεταβλητή*, τη θέση της οποίας μπορεί να καταλάβει το θέμα οποιασδήποτε λέξης από το σύνολο. Εναλλακτικά, η παρατηρούμενη κανονικότητα μπορεί να θεωρηθεί αποτέλεσμα ενός συνόλου συγκεκριμένων συνειρμών: {παιδί \Rightarrow παιδιά, λουρί \Rightarrow λουριά, κουμπί \Rightarrow κουμπιά, φλουρί \Rightarrow φλουριά}. Οι υποστηρικτές της παραγωγικότητας θεωρούν ότι με τους επιμέρους συνειρμούς χάνουμε την ουσία της κανονικότητας, καθώς υποθέτουν ότι οι όμοιες καταλήξεις όλων αυτών των λέξεων κρύβουν μια σημαντική και ουσιαστική γενίκευση της γλώσσας. Το πλεονέκτημα της παραγωγικότητας είναι ότι δεν χρειάζεται επιμέρους γνώση για καθεμιά λέξη, οπότε η κανονικότητα του συστήματος είναι εγγυημένη και ανεξάρτητη από γνώση κάθε λέξης. Η υπόθεση της παραγωγικότητας περιλαμβάνει τη θέση ότι οι λέ-

⁸Η παρουσίαση που ακολουθεί βασίζεται εν μέρει στο Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71.

⁹Η λέξη «ατομικός» εδώ χρησιμοποιείται όπως στο άτομο της φυσικής, δηλαδή με την έννοια ότι είναι στοιχειώδης, όχι σύνθετη, και δεν μπορεί να αναλυθεί σε μικρότερα τμήματα.

ξεις αποτελούνται από επιμέρους τμήματα και ότι γενικοί κανόνες επιδρούν στα τμήματα αυτά παράγοντας νέες αλληλουχίες συμβόλων ανεξάρτητα από τη συγκεκριμένη λέξη. Ο κανόνας είναι γενικός και αφηρημένος με την έννοια ότι δεν εξαρτάται από τις συγκεκριμένες λέξεις στις οποίες εφαρμόζεται. Για την ακρίβεια, εξαρτάται από κάποιες ιδιότητες των λέξεων, τις λεγόμενες συντακτικές, χάρη στις οποίες κάθε λέξη μπορεί να αποτελεί κατάλληλο πεδίο εφαρμογής του κανόνα ή όχι. Οι ιδιότητες αυτές δεν έχουν να κάνουν με το περιεχόμενο της λέξης (τη σημασία της) αλλά μόνο με τη μορφή. Στο συγκεκριμένο παράδειγμα, οποιοδήποτε ουδέτερο σε -ί μπορεί να λάβει μέρος στον κανόνα. Μάλιστα, δεν χρειάζεται να είναι καν γνωστή λέξη: Μια ψευδολέξη όπως «φλουπί» μπορεί κάλλιστα να αποτελέσει αντικείμενο μετασχηματισμού από τον κανόνα (\Rightarrow φλουπιά) καθώς διαθέτει τις κατάλληλες συντακτικές ιδιότητες, παρότι δεν έχει καν σημασιακό περιεχόμενο.

Ένα κρίσιμο στοιχείο της παραγωγικότητας στο συγκεκριμένο παράδειγμα είναι η διατήρηση του θέματος και η διατήρηση της κατάληξης. Πρόκειται για την έκφραση της κανονικότητας σε ένα επίπεδο που επιτρέπει τη γενίκευση με συντακτικά κριτήρια. Η διατήρηση του θέματος σημαίνει ότι η παραγωγή του πληθυντικού δεν είναι μια αυθαίρετη απεικόνιση από ένα σύμβολο σε άλλο, αλλά η αντικατάσταση ενός μέρους του σύνθετου συμβόλου (της «κατάληξης») έτσι ώστε το άλλο μέρος του (το «θέμα») να παραμένει ακριβώς το ίδιο (όχι απλώς να είναι όμοιο). Αντίστοιχα, η διατήρηση της κατάληξης σημαίνει ότι για οποιοδήποτε συνδυαστικό σύμβολο που έχει τις κατάλληλες συντακτικές ιδιότητες ώστε να ενεργοποιήσει τον κανόνα, το αποτέλεσμα που παράγεται δεν είναι μια αυθαίρετη κάθε φορά αντιστοιχία, αλλά η αντικατάσταση του «ί» από το ίδιο πάντα σύμβολο «ιά».

Συνδυαστικότητα ή *συνθετικότητα* είναι η ιδιότητα των συμβόλων να σχηματίζουν (ή να αποτελούνται από) αλληλουχίες συμβόλων έτσι ώστε οι (συντακτικές και σημασιακές) ιδιότητες του συνδυασμού να μπορούν να υπολογιστούν από τις ιδιότητες των επιμέρους στοιχείων. Για παράδειγμα, αν ένας νους μπορεί να περιλάβει και να χρησιμοποιήσει την αναπαράσταση που αντιστοιχεί στην πρόταση «ο Γιώργος αγαπά τη Μαρία» τότε μπορεί επίσης να περιλάβει και να χρησιμοποιήσει εξίσου την αναπαράσταση που αντιστοιχεί στην πρόταση «η Μαρία αγαπά το Γιώργο», χωρίς να χρειάζεται να γνωρίζει κάτι ιδιαίτερο για τη Μαρία, το Γιώργο ή τη σχέση μεταξύ τους. Εξυπακούεται ότι η γνώση των λέξεων και των συντακτικών περιορισμών τους επαρκεί για όλους τους επιτρεπτούς συνδυασμούς. Ομοίως, ένας νους που καταλαβαίνει ότι $(A \wedge B) \Rightarrow A$ μπορεί να συμπεράνει από την πρόταση «σήμερα είναι Δευτέρα και βρέχει» ότι «σήμερα είναι Δευτέρα», ανεξάρτητα από το περιεχόμενο των προτάσεων και ανεξάρτητα ακόμα από το αν έχει προαποθηκευμένες κάποιες γνώσεις για τις Δευτέρες ή αν ξέρει τι σημαίνει Δευτέρα. Η συνδυαστικότητα επιτρέπει την εξαγωγή γενικευμένων συμπερασμάτων χάρη στις σχέσεις μέρους-όλου που διακρίνουν τις σύνθετες αναπαραστάσεις από τα μέρη τους. Έτσι ένας συνδυαστικός νους μπορεί να συμπεράνει κάτι για το Γιώργο (με την έννοια ότι μπορεί να χειριστεί συντακτικά τις αντίστοιχες αναπαραστάσεις) από την πρόταση «ο Γιώργος αγαπά το φλουπί» χωρίς να χρειάζεται να γνωρίζει τι είναι το [φλουπί]. Σύμφωνα με τη συμβολική προσέγγιση, αυτό επιτυγχάνεται επειδή οι ατομικές αναπαραστάσεις [Γιώργος] και [αγαπά] συμμετέχουν

συνθετικά στη συνδυαστική αναπαράσταση της πρότασης. Με άλλα λόγια, οι ιδιότητες της πρότασης μπορούν να υπολογιστούν από τις ιδιότητες των μερών της και τις σχέσεις μεταξύ τους. Έτσι, αν κάποιος μπορεί να χειριστεί την αναπαράσταση «ο Γιώργος αγαπά τη Μαίρη», τότε θα πρέπει να διαθέτει τη μοναδιαία αναπαράσταση του [αγαπά] και να μπορεί να τη χειριστεί σε οποιαδήποτε πρόταση.

Συνεπώς, η ιδιότητα της παραγωγικότητας αφορά άμεσα στα σύμβολα και επιτρέπει στο νου να παράγει σύνθετες αναπαραστάσεις με συστηματικό τρόπο, ενώ η ιδιότητα της συνδυαστικότητας έχει να κάνει με δομές και διατάξεις συμβόλων και με τη δυνατότητα διατήρησης των νοημάτων μέσα στους συνδυασμούς τους, εφόσον το περιεχόμενο μιας σύνθετης αναπαράστασης προσδιορίζεται συστηματικά με βάση το περιεχόμενο και τη διάταξη των επιμέρους ατομικών συμβόλων που περιέχει.

Πιο γενικά από θέματα και καταλήξεις λέξεων, ένα βασικό πλεονέκτημα της παραγωγικότητας στην επεξεργασία των συμβόλων είναι η διατήρηση της τιμής αλήθειας, δηλαδή η γενικευμένη ορθολογική συμπεριφορά. Αν το σύστημα επεξεργασίας είναι έτσι καθορισμένο ώστε να ακολουθεί τους κανόνες της λογικής, τότε η συντακτική εφαρμογή των κανόνων διατηρεί πάντα την αλήθεια ανεξάρτητα από το εκάστοτε σημασιακό περιεχόμενο των συμβόλων που υφίστανται την επεξεργασία. Όπως ένα κομπιουτεράκι είναι κατασκευασμένο έτσι ώστε να διατηρεί τις αριθμητικές σχέσεις επεξεργαζόμενο μη αριθμητικά σύμβολα στο ηλεκτρονικό του κύκλωμα, έτσι και ένα φορμαλιστικό σύστημα νόησης, αν γνωρίζει ότι $(A \wedge B)$ μπορεί πάντοτε να συμπεραίνει ότι A , παράγοντας το ορθό αποτέλεσμα χωρίς να χρειάζεται να γνωρίζει τι σημαίνουν τα A και B . Στην περίπτωση του συστήματος νόησης θα πρέπει βέβαια οι συντακτικοί κανόνες να είναι έτσι ρυθμισμένοι ώστε να ακολουθούν τόσο τη λογική όσο και το σύστημα των εννοιών που αναπαριστούν. Αυτό θεωρείται ότι είναι αποτέλεσμα εξελικτικών διεργασιών επιλογής και προσαρμογής και μπορεί εύλογα να θεωρείται δεδομένο.

2.4 Αρθρώματα και τοπικότητα

Ένα σημαντικό ζήτημα για την κατασκευή γενικευμένων νοητικών μοντέλων, το οποίο γίνεται ιδιαίτερα αισθητό σε συστήματα τεχνητής νοημοσύνης, έχει να κάνει με τη διαθεσιμότητα των κατάλληλων κάθε φορά συμβόλων για την απαραίτητη επεξεργασία. Όπως είναι προφανές, για κάθε απόφαση ή δραστηριότητά μας μπορεί να παίξουν ρόλο πληροφορίες που διαθέτουμε από οποιαδήποτε πηγή ή εμπειρία του παρελθόντος, κατά τρόπο κάθε φορά απρόβλεπτο. Για παράδειγμα, στην αποδοχή μιας πρόσκλησης σε γεύμα ή μιας εντολής διορισμού, μπορεί τη στιγμή της απόφασης να παίξουν ρόλο όχι μόνο οι ενδεχόμενες απολαβές και άλλα πιθανά κέρδη, αλλά η διάθεσή μας, η ανάμνηση γεγονότων του παρελθόντος από κάποιον που είχε το ίδιο όνομα ή το ίδιο επάγγελμα, μια αγαπητή ή αντιπαθητική συγγενική φιγούρα που μας είχε πρόσφατα συμβουλέψει σχετικά με το θέμα αυτό, ο σημερινός καιρός ή ένα επεισόδιο στο δρόμο το πρωί, κλπ. Αυτό δεν σημαίνει ότι το σύνολο των πληροφοριών που διαθέτουμε θα επηρεάσει την απόφαση, αλλά ότι ένα εν πολλοίς απρόβλεπτο υποσύνολο από αυτές θα παίζει κάποιο ρόλο. Το πρόβλημα για

ένα μοντέλο είναι να επιλεγούν οι σχετικές κάθε φορά πληροφορίες και να χρησιμοποιηθούν για την επίλυση του προβλήματος, χωρίς να είναι απαραίτητο να εξεταστούν όλες οι διαθέσιμες πληροφορίες. Αυτό ισχύει διότι αν εξεταστούν όλες, τότε κατά κάποιον τρόπο έχουν όλες συμμετάσχει, είτε τελικά χρησιμοποιήθηκαν για την απόφαση είτε όχι. Η αποδοχή ενός μοντέλου στο οποίο το σύνολο των υπάρχουσών πληροφοριών συμμετέχει σε οποιαδήποτε, απλή ή σύνθετη, νοητική λειτουργία, είναι τουλάχιστον ανεπιθύμητη, αν όχι εντελώς απαράδεκτη. Από την άλλη, η ανάθεση σε κάποιο «εκτελεστικό» σύστημα του νου να επιλέγει τις κατάλληλες κάθε φορά πληροφορίες απλώς μεταθέτει το πρόβλημα από το σύστημα σε ένα υποσύστημά του. Αυτό είναι το πρόβλημα της *τοπικότητας*.¹⁰

Μια προσέγγιση στο πρόβλημα της τοπικότητας είναι ο κατακερματισμός των νοητικών λειτουργιών σε αυτόνομες λειτουργικές μονάδες που τις ονομάζουμε *σπονδύλους* ή *αρθρώματα* (modules). Μια αυτόνομη λειτουργική μονάδα είναι ένα υποσύστημα του νου το οποίο επιτελεί μια συγκεκριμένη λειτουργία, διαθέτει είσοδο (εισροή ερεθισμάτων) και έξοδο (εκροή αποκρίσεων), περιλαμβάνει στο εσωτερικό του όλες τις πληροφορίες που είναι απαραίτητες για τη λειτουργία του και μόνο αυτές, και στις πληροφορίες αυτές δεν έχουν πρόσβαση άλλα συστήματα. Ίσως το κυριότερο χαρακτηριστικό μιας αυτόνομης λειτουργικής μονάδας είναι αυτή η πληροφοριακή απομόνωση (ενθυλάκωση, encapsulation), η επιλεκτική δηλαδή πρόσβαση στις δικές της πληροφορίες και μόνο σε αυτές. Η κατανομή των πληροφοριών του συστήματος είναι τέτοια που απομονώνει κάθε λειτουργική μονάδα από τις πληροφορίες για τη λειτουργία των άλλων μονάδων. Με τον τρόπο αυτό περιορίζεται το πρόβλημα της τοπικότητας για τις λειτουργίες εκείνες που μπορούν να πραγματοποιηθούν με βάση περιορισμένο και προκαθορισμένο σύνολο πληροφοριών. Για παράδειγμα, μια αυτόνομη λειτουργική μονάδα οπτικής αντίληψης αντικειμένων δέχεται οπτικά ερεθίσματα και αποκρίνεται με την ταυτότητα των αναγνωρισμένων αντικειμένων, διαθέτοντας τους κατάλληλους μηχανισμούς ταύτισης προτύπων, όλα τα απαραίτητα οπτικά πρότυπα (ή παραδείγματα) αντικειμένων, και όλες τις οπτικές εμπειρίες και σχετικές μνημονικές συνειρμικές συνδέσεις μεταξύ αυτών. Μια αυτόνομη λειτουργική μονάδα ακουστικής αντίληψης θα δέχεται ακουστικά ερεθίσματα και θα διαθέτει τις απαραίτητες λειτουργίες, ακουστικές πληροφορίες και μνήμες, αλλά δεν θα μπορεί να έχει πρόσβαση στις οπτικές πληροφορίες της πρώτης μονάδας.

Πολλά έχουν ειπωθεί για τον κατακερματισμό του νου σε αυτόνομες λειτουργικές μονάδες και για τις πιθανές σχέσεις τέτοιων νοητικών μονάδων με εγκεφαλικές περιοχές και εντοπισμένες λειτουργίες. Τόσο η λειτουργική αυτονομία όσο και η πληροφοριακή απομόνωση αποτελούν αμφιλεγόμενες αρετές και, ούτως ή άλλως, δε λύνουν το πρόβλημα της τοπικότητας, δεδομένου ότι κάποιες τουλάχιστον «ανώτερες» νοητικές λειτουργίες θα πρέπει να έχουν πρόσβαση στο πλήρες σύνολο πληροφοριών (όπως π.χ. οι λειτουργίες για την προαναφερθείσα απόφαση αποδοχής ή όχι μιας επαγγελματικής θέσης).

Το σχήμα αλληλεπίδρασης μεταξύ αυτόνομων λειτουργικών μονάδων με ή χωρίς έ-

¹⁰Ένα σχετιζόμενο ζήτημα έχει να κάνει με τον αναγκαστικά τοπικό χαρακτήρα της συντακτικής επεξεργασίας συμβόλων και συνεπώς την ανεπάρκειά της για «ολικές» νοητικές λειτουργίες.

να κεντρικό επιτελικό (executive) υποσύστημα παρατηρείται σε μια οικογένεια ιδιαίτερα επιτυχημένων νοητικών μοντέλων που ονομάζονται *συστήματα παραγωγής* (production systems). Ένα σύστημα παραγωγής αποτελείται από κανόνες που ονομάζονται *παραγωγές* (productions). Κάθε παραγωγή είναι μια αυτόνομη λειτουργική μονάδα που χαρακτηρίζεται από μια συνθήκη ενεργοποίησης και μια ενέργεια που εκτελείται όποτε ικανοποιείται η συνθήκη. Η επικοινωνία μεταξύ των παραγωγών γίνεται χάρη σε έναν κοινό χώρο «ανακοινώσεων», στον οποίο έχουν όλες οι μονάδες πρόσβαση για πρόσληψη ή καταγραφή πληροφοριών. Κάθε τέτοια μονάδα παρακολουθεί τον κοινό χώρο και εξετάζει διαρκώς αν πληρούνται οι συνθήκες ενεργοποίησής της. Όταν διαπιστώσει ότι πληρούνται, τότε ενεργοποιείται και εκτελεί την προκαθορισμένη ενέργειά της, λαμβάνοντας όποιες πληροφορίες της είναι απαραίτητες από τον κοινό χώρο και αποθέτοντας εν συνεχεία εκεί το αποτέλεσμα της επεξεργασίας της. Μετά το πέρας της εργασίας επανέρχεται στην πρότερη κατάσταση κατά την οποία ελέγχει διαρκώς αν ικανοποιείται η συνθήκη ενεργοποίησής της. Ένα τέτοιο σύστημα δεν χρειάζεται κεντρικό ελεγκτή ούτε σύστημα διαχείρισης και καταμερισμού εργασίας, αφού κάθε παραγωγή τίθεται αυτόματα σε λειτουργία από τη συνθήκη ενεργοποίησής της. Έτσι, δεν είναι απαραίτητη η διαθεσιμότητα εξειδικευμένων πληροφοριών σε κάποιο κεντρικό σύστημα για την επιλογή της κατάλληλης κάθε φορά επεξεργαστικής μονάδας. Περισσότερες λεπτομέρειες για ένα συγκεκριμένο σύστημα αναφέρονται στην ενότητα 14 (σελ. 118) όπου παρουσιάζεται η αρχιτεκτονική ACT-R, που είναι ένα σύστημα παραγωγής ως γενικό μοντέλο του νου.

2.5 Περίληψη

Τα συμβολικά υπολογιστικά συστήματα χαρακτηρίζονται από μια αλληλουχία καταστάσεων και από κανόνες μετάβασης από μια κατάσταση στην επόμενη. Οι κανόνες αυτοί είναι ευαίσθητοι στις αναπαραστασιακές ιδιότητες των καταστάσεων, δηλαδή στα συντακτικά χαρακτηριστικά των συμβόλων που εκφράζονται σε κάθε κατάσταση. Οι κανόνες μετάβασης διατηρούν έτσι την αιτιότητα στο επίπεδο των συμβόλων, δίνοντας στις αναπαραστάσεις αιτιακή ισχύ, ενώ συγχρόνως διατηρούν τη σημασιακή αξιοπιστία ως προς τον αναπαριστώμενο κόσμο και τη γενικευσιμότητα στο νοητικά σημαντικό επίπεδο. Το σημαντικότερο πλεονέκτημα τέτοιων συστημάτων είναι η έκφραση της συστηματικότητας των νοητικών φαινομένων μέσω της παραγωγικότητας των διεργασιών και της συνδυαστικότητας των αναπαραστάσεων.

Στη συμβολική προσέγγιση θεωρείται ότι η απλοϊκή ψυχολογία είναι κατά βάση σωστή, δηλαδή επιθυμούμε να διατηρήσουμε στις εξηγήσεις μας έννοιες όπως επιθυμία, πεποίθηση κλπ., και τους αποδίδουμε αιτιακές ιδιότητες θεωρώντας τις ειδικές νοητικές καταστάσεις. Διατηρούμε γενικά κάθε γενικευμένη νοητική έννοια που φαίνεται πως έχει επεξηγηματική ισχύ, αποδίδοντάς την σε αντίστοιχη συμβολική αναπαράσταση. Η παρομοίωση του ανθρώπινου νου με ηλεκτρονικό υπολογιστή δεν είναι σχήμα λόγου αλλά ουσιαστική θεωρητική θέση, ειδικά όσον αφορά στην εξωτερικότητα των ερμηνειών που προκύπτει από την υπόθεση της αποκλειστικά συντακτικής επεξεργασίας των αναπαραστάσεων-συμβόλων.

3 Συνδεδεισμός

Μια άλλη προσέγγιση στη μελέτη του νου χρησιμοποιεί μοντέλα βασισμένα στην αρχιτεκτονική του εγκεφάλου, με την έννοια ότι διατηρούνται ορισμένες λειτουργικές ιδιότητες των φυσικών νευρωνικών δικτύων οι οποίες θεωρούνται σημαντικές για τις υπολογιστικές ιδιότητες του νου. Τα μοντέλα αυτά ονομάζονται τεχνητά νευρωνικά δίκτυα, ή *συνδεδειστικά* (ή συνδεδειακά) μοντέλα. Παρότι τα μοντέλα αυτά εμπνέονται από το νευρικό σύστημα, η λειτουργία τους στα πλαίσια των γνωσιακών επιστημών δεν αποσκοπεί στην κατανόηση και περιγραφή της νευρωνικής λειτουργίας αλλά της νόησης. Έτσι οι αντιστοιχίες μεταξύ φυσικών και τεχνητών νευρωνικών δικτύων είναι σε ένα αφηρημένο επίπεδο γενίκευσης.

3.1 Θεμελιώδεις ιδιότητες βιολογικών νευρώνων

Όπως σε κάθε απόπειρα κατασκευής μοντέλου, έτσι και στη σχεδίαση τεχνητών νευρωνικών δικτύων έχει μεγάλη σημασία η αφαιρετική προσέγγιση του φυσικού δικτύου. Δηλαδή η διατήρηση των ιδιοτήτων του φυσικού δικτύου που θεωρούνται σημαντικές στο επίπεδο λειτουργίας που εξετάζουμε και η απόρριψη όλων των υπολοίπων «λεπτομερειών» που δεν θεωρούνται απαραίτητες και που όχι μόνο δυσκολεύουν την υλοποίηση του μοντέλου αλλά αποσπούν και την προσοχή από την ουσία. Στο επίπεδο της νόησης δεν συμπεριλαμβάνονται οι ηλεκτρικές και χημικές ιδιότητες των κυττάρων αλλά η λειτουργία ενός συνόλου κυττάρων σε δίκτυο. Οι βασικές ιδιότητες των νευρώνων που θεωρούνται σημαντικές σε επίπεδο δικτύου και συνήθως διατηρούνται σε τεχνητά νευρωνικά δίκτυα είναι οι εξής:¹¹

- *Ο νευρώνας απαρτιώνει πληροφορίες.* Ανεξάρτητα από το σχήμα, το μέγεθος και τη θέση του, ένας νευρώνας δέχεται μηνύματα από άλλους νευρώνες και παράγει δυναμικά ενεργείας όταν η συνολική εισερχόμενη δραστηριότητα υπερβεί κάποιο χαρακτηριστικό όριο (κατώφλι ή ουδό, threshold). Δηλαδή κάθε νευρώνας δέχεται σήματα και ενεργοποιείται για την αποστολή σημάτων προς άλλους νευρώνες.
- *Ο νευρώνας αναμεταδίδει τη δραστηριότητα των εισόδων του.* Η ποσότητα που χαρακτηρίζει κάθε νευρώνα είναι ο ρυθμός παραγωγής δυναμικών ενεργείας, ως δείκτης του βαθμού ενεργοποίησης. Κάθε νευρώνας μπορεί να δέχεται συνολικά ισχυρή ή ασθενή αναστολή (inhibition) ή διέγερση (excitation). Όσο ισχυρότερη είναι η συνολική διέγερση τόσο συχνότερα είναι τα δυναμικά ενεργείας που παράγονται.
- *Οι νευρικές δομές (δίκτυα) λειτουργούν κατά στάδια.* Η επεξεργασία των πληροφοριών από τα βιολογικά νευρωνικά δίκτυα δεν γίνεται από μια αδιαμόρφωτη μάζα νευρικών κυττάρων αλλά από μια αλληλουχία ομάδων κυττάρων, καθεμιά από τις οποίες θεωρούμε ότι επιτελεί κάποια συγκεκριμένη λειτουργία και στη συνέχεια προωθεί το αποτέλεσμα της επεξεργασίας της στην επόμενη (ή επόμενες).

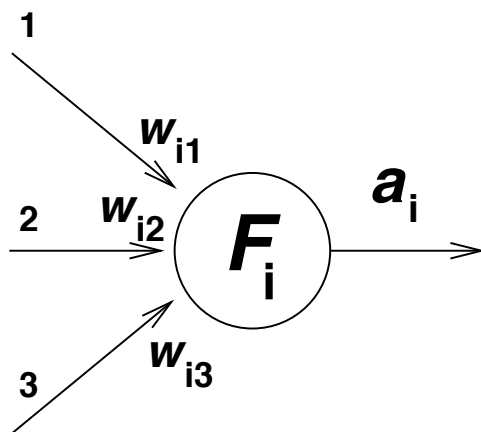
¹¹Από το McLeod, P., Plunkett, K., & Rolls, E. (1998). *Introduction to connectionist modeling of cognitive processes*. Oxford University Press.

- Η επίδραση ενός νευρώνα σε έναν άλλο είναι ανάλογη της ισχύος σύνδεσης μεταξύ τους. Δηλαδή για ένα υποθετικό ζευγάρι αλληλοσυνδεόμενων νευρώνων Α και Β δεν αρκεί να καθορίσουμε το βαθμό διέγερσης του Α για να προκύψει και η επίδραση της διέγερσης αυτής στο Β, αλλά χρειάζεται να γνωρίζουμε και την αποτελεσματικότητα της μεταξύ τους σύναψης. Οι διαφορές αυτές στο συναπτικό «βάρος» έχουν ως αποτέλεσμα το «ζύγισμα» της πληροφορίας που προέρχεται από άλλους νευρώνες πριν το συνυπολογισμό τους στο τελικό αποτέλεσμα, δηλαδή στο βαθμό διέγερσης του μετασυναπτικού νευρώνα.
- Η μάθηση επιτυγχάνεται με αλλαγή συναπτικών βαρών. Όταν υπάρχει μεταβολή στη λειτουργία ενός νευρωνικού δικτύου, αυτή κατά κανόνα οφείλεται σε μεταβολές στις συνάψεις οι οποίες μεταβάλλουν την αποτελεσματικότητά τους. Δηλαδή οι μεταβολές της λειτουργίας του δικτύου οφείλονται σε μεταβολές στις συνδέσεις μεταξύ των κυττάρων-μονάδων και όχι σε μεταβολές στη λειτουργία ή τον αριθμό των κυττάρων.

3.2 Μαθηματικοί νευρώνες και συνδεδετιστικά δίκτυα

Διατηρώντας τις παραπάνω βασικές αρχές για τη δημιουργία ενός συνδεδετιστικού μοντέλου, ορίζουμε κατ' αρχήν τις μονάδες επεξεργασίας, τις οποίες ονομάζουμε κόμβους (nodes), και τις συνδέσεις μεταξύ τους. Στο Σχήμα 1 απεικονίζεται σχηματικά ένας τέτοιος κόμβος, ή «μαθηματικός νευρώνας». Κάθε κόμβος περιλαμβάνει προσαγωγές συνδέσεις (προς αυτόν, afferent) και απαγωγές συνδέσεις (από αυτόν, efferent) και χαρακτηρίζεται από ένα βαθμό διέγερσης (ή επίπεδο δραστηριοποίησης, activation level). Ο βαθμός διέγερσης εξαρτάται από τον ίδιο τον κόμβο και από το σύνολο της εισερχόμενης διέγερσης.

Η σχέση μεταξύ της εισερχόμενης δραστηριότητας, που μεταφέρεται μέσω των προσαγωγών συνδέσεων, και του συνεπαγόμενου βαθμού διέγερσης του κόμβου ονομάζεται *συνάρτηση απαρτίωσης* (integration function, ή συνάρτηση ενεργοποίησης, activation function) και μπορεί να πάρει διάφορες μορφές. Για παράδειγμα, μπορεί να είναι μια



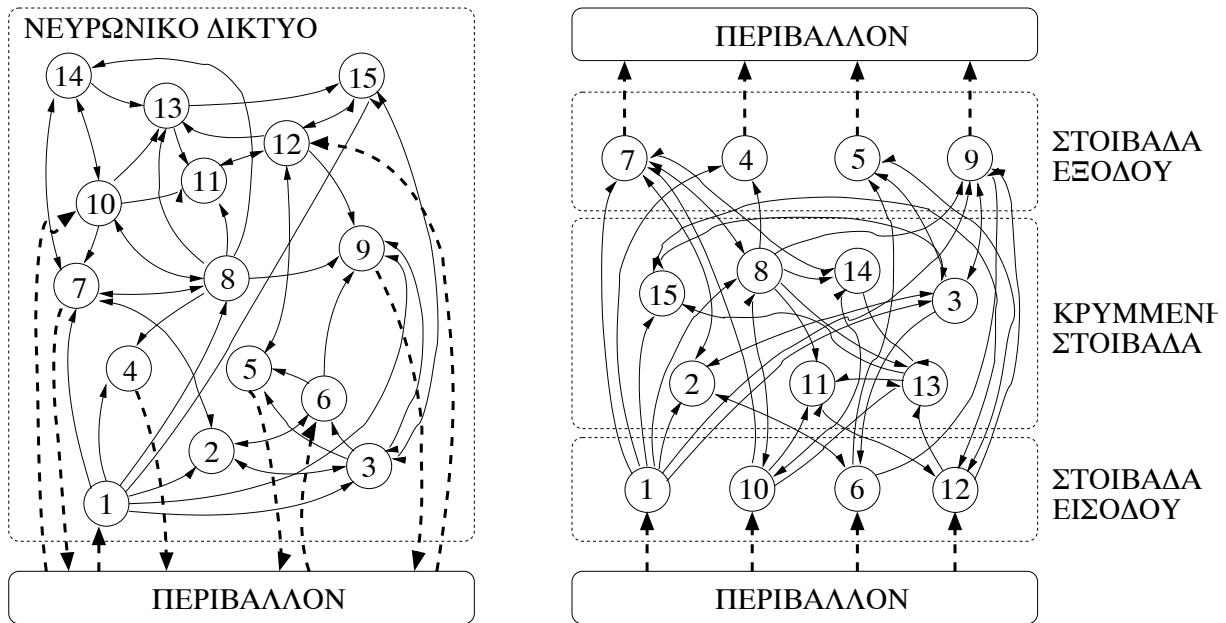
Σχήμα 1: Διάγραμμα ενός μαθηματικού νευρώνα και των συνδέσεών του. Ο απεικονιζόμενος κόμβος έχει αύξοντα αριθμό i και χαρακτηρίζεται από τη συνάρτηση απαρτίωσης F_i . Διακρίνονται οι προσαγωγές συνάψεις από άλλους κόμβους, με αύξοντες αριθμούς 1, 2 και 3, οι οποίες χαρακτηρίζονται από συναπτικά βάρη w_{i1} , w_{i2} και w_{i3} αντίστοιχα. Αν συμβολίσουμε με α_j το βαθμό διέγερσης κάθε κόμβου j , τότε η δραστηριότητα του κόμβου i θα είναι, συνεπώς, $\alpha_i = F(\sum_{j=1}^3 \alpha_j w_{ij})$, δηλαδή $\alpha_i = F(\alpha_1 w_{i1} + \alpha_2 w_{i2} + \alpha_3 w_{i3})$.

γραμμική σχέση, δηλαδή ο βαθμός διέγερσης του κόμβου να είναι ευθέως ανάλογος της συνολικής εισερχόμενης διέγερσης. Μπορεί όμως να είναι και μη γραμμική σχέση, όπως στην περίπτωση που ο κόμβος χαρακτηρίζεται από έναν ουδό ενεργοποίησης, έτσι ώστε για συνολική εισερχόμενη διέγερση μικρότερη μιας τιμής (κατωφλίου) ο τελικός βαθμός ενεργοποίησης του κόμβου να είναι μηδέν, ενώ για οποιαδήποτε εισερχόμενη διέγερση μεγαλύτερη της τιμής κατωφλίου ο τελικός βαθμός διέγερσης του κόμβου να είναι μέγιστος (βηματική συνάρτηση, με δυαδική εκροή). Η πιο συνηθισμένη συνάρτηση απαρίθωσης που χρησιμοποιείται στην πράξη είναι η λογιστική συνάρτηση, η οποία συνδυάζει χαρακτηριστικά της γραμμικής και της βηματικής συνάρτησης, καθώς η διέγερση εκροής αυξάνεται όσο αυξάνεται η συνολική εισροή, υπάρχει όμως μια μεσαία περιοχή απότομης αύξησης ενώ στα άκρα προσεγγίζεται αργά και σταδιακά το μηδέν (για μικρή εισροή) και το μέγιστο (για μεγάλη εισροή).

Οι κόμβοι ενός συνδεδεμένου μοντέλου επικοινωνούν μεταξύ τους μέσω διαβαθμισμένων συνδέσεων. Κάθε σύνδεση μεταξύ κόμβων χαρακτηρίζεται από μια αριθμητική τιμή που ονομάζεται βάρος της σύνδεσης. Η τιμή αυτή καθορίζει το βαθμό επιρροής του ενός κόμβου στον άλλο μέσω αυτής της σύνδεσης: αν είναι μικρή, τότε μικρό μέρος της διέγερσης περνά από τον ένα κόμβο στον άλλο, ενώ αν είναι μεγάλη περνά μεγάλο μέρος της διέγερσης. Αν το βάρος μιας σύνδεσης είναι αρνητικό τότε πρόκειται για ανασταλτική σύνδεση, που τείνει δηλαδή να μειώσει και όχι να αυξήσει τη διέγερση του «μετασυναπτικού» κόμβου ανάλογα με τη διέγερση του «προσυναπτικού» κόμβου.

Για τα συνδεδεμένα δίκτυα χρησιμοποιούνται συνήθως μαθηματικοί συμβολισμοί ώστε να διευκολύνεται η ανάλυση της λειτουργίας τους και ο προγραμματισμός των μοντέλων σε ηλεκτρονικούς υπολογιστές. Κάθε κόμβος λαμβάνει έναν αύξοντα αριθμό. Ο βαθμός διέγερσής του, που είναι και η μόνη ιδιότητα του κόμβου που μας ενδιαφέρει για το μοντέλο μας, συμβολίζεται με ένα γράμμα, π.χ. το α , με δείκτη ο οποίος αντιστοιχεί στον αύξοντα αριθμό του κόμβου. Δηλαδή ο βαθμός διέγερσης του πρώτου κόμβου είναι α_1 , του δεύτερου α_2 κ.ο.κ. Η αρίθμηση των κόμβων είναι αυθαίρετη, αρκεί να είναι συνεπής. Για να αναφερθούμε γενικά στη διέγερση οποιουδήποτε κόμβου χρησιμοποιούμε ένα γράμμα-δείκτη, π.χ. το i , δηλαδή γράφουμε α_i , όπου το i μπορεί να πάρει οποιαδήποτε τιμή από 1 μέχρι τον αριθμό των κόμβων N . Αντίστοιχα, με το συμβολισμό $\alpha_i, i = 1 \dots N$ αναφερόμαστε σε όλους τους κόμβους (σε κάθε α_i για i από ένα μέχρι N). Οι συνδέσεις συμβολίζονται με τα βάρη τους, χρησιμοποιώντας το γράμμα w . Αν ένας κόμβος δεν συνδέεται με κάποιον άλλο τότε θέτουμε το βάρος της σύνδεσης ίσο με μηδέν και πετυχαίνουμε το ίδιο αποτέλεσμα (δηλαδή να μην υπάρχει επίδραση του ενός κόμβου στον άλλο) χωρίς να διαταράζουμε την ομοιομορφία του συμβολισμού (για ευκολία στη μαθηματική ανάλυση). Η σύνδεση από τον κόμβο i στον κόμβο j συμβολίζεται w_{ji} . Αν υπάρχει και σύνδεση αντίστροφα, από τον κόμβο j προς τον κόμβο i , αυτή συμβολίζεται αντίστοιχα w_{ij} .

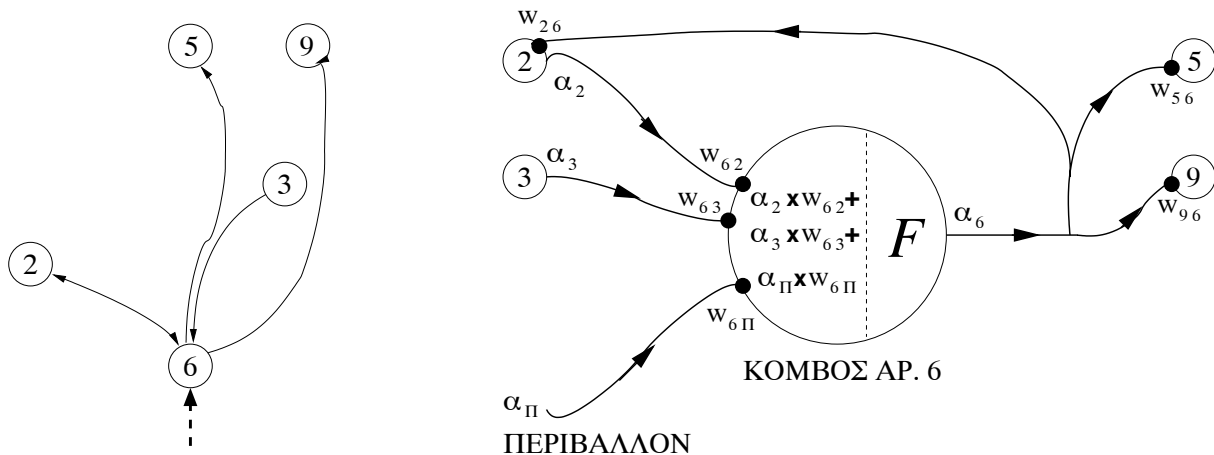
Η διέγερση που μεταδίδεται από τον κόμβο i προς τον κόμβο j μέσω της σύνδεσης μεταξύ τους (με βάρος w_{ji}) είναι ανάλογη της διέγερσης του «αρχικού» κόμβου i και του βάρους σύνδεσης, ισούται δηλαδή με το γινόμενο $\alpha_i \times w_{ji}$, που το γράφουμε απλά $\alpha_i w_{ji}$.



Σχήμα 2: Διάγραμμα ενός δικτύου 15 νευρώνων. Απεικονίζονται οι συνδέσεις μεταξύ των κόμβων («νευρώνων») καθώς και αυτές από και προς το περιβάλλον («ερέθισμα» και «απόκριση», αντίστοιχα). Τα βέλη στα άκρα των συνδέσεων απεικονίζουν τη φορά της ροής πληροφορίας (ή διάδοσης δραστηριότητας), η οποία μπορεί να είναι μονόδρομη ή αμφίδρομη. Το ίδιο ακριβώς δίκτυο απεικονίζεται αριστερά, χωρίς κάποια συγκεκριμένη διάταξη, και δεξιά, με τους κόμβους εισροής και εκροής να έχουν ομαδοποιηθεί σε στιβάδες. Η διάταξη δεξιά καθιστά προφανή την ύπαρξη στιβάδας κρυφών κόμβων. Η φορά της διάδοσης δραστηριότητας δεν επηρεάζεται από τη διάταξη στην απεικόνιση και είναι ανεξάρτητη από τη σύνδεση ή όχι ενός κόμβου με το «περιβάλλον».

Η συνολική διέγερση που δέχεται ο κόμβος j περιλαμβάνει, εκτός αυτής από τον κόμβο i , και τη διέγερση από κάθε ένα άλλο κόμβο, από 1 δηλαδή μέχρι N . Το άθροισμα που προκύπτει, $\alpha_1 w_{j1} + \alpha_2 w_{j2} + \alpha_3 w_{j3} + \dots + \alpha_N w_{jN}$, συμβολίζεται $\sum_{i=1}^N \alpha_i w_{ji}$, δηλαδή το άθροισμα όλων των γινομένων εισερχόμενης διέγερσης επί το αντίστοιχο βάρος σύνδεσης για κάθε κόμβο από 1 μέχρι N . Το σύνολο της διέγερσης που φτάνει σε έναν κόμβο μέσω των προσαγωγών συνδέσεων ισούται με αυτό το άθροισμα. Η συνεπαγόμενη διέγερση του κόμβου j που προκύπτει από αυτή τη διέγερση εξαρτάται από τη συνάρτηση απαρίθωσης του κόμβου j , την οποία συμβολίζουμε με F . Δηλαδή έχουμε τελικά $\alpha_j = F(\sum_{i=1}^N \alpha_i w_{ji})$.

Οι κόμβοι ενός συνδεδεμένου δικτύου μπορούν αυθαίρετα να καταταχθούν σε περισσότερα του ενός λειτουργικά επίπεδα, στιβάδες ή στρώματα (layers), ακολουθώντας τη βιολογική αρχή της επεξεργασίας κατά στάδια (βλ. και σχήμα 2, δεξιά). Έτσι, μπορούμε συνήθως να διακρίνουμε τους κόμβους εισροής και τους κόμβους εκροής σε ένα τεχνητό νευρωνικό δίκτυο. Οι κόμβοι εισροής, δηλαδή η «είσοδος» του δικτύου, είναι εκείνοι στους οποίους εφαρμόζεται κάποιο «ερέθισμα», δηλαδή κάποια διέγερση απευθείας από το εξωτερικό του μοντέλου περιβάλλον. Για παράδειγμα, σε ένα μοντέλο οπτικής επεξεργασίας



Σχήμα 3: Διευκρινιστική απεικόνιση για το συμβολισμό των συνδέσεων του σχήματος 2. Αριστερά απεικονίζεται το μέρος του δικτύου που αφορά στον κόμβο με αριθμό 6. Δεξιά διακρίνονται λεπτομερώς οι σημαντικές ποσότητες (διεγέρσεις, βάρη συνδέσεων, αθροίσεις κλπ.) και οι αντίστοιχες συνδέσεις διάδοσης διέγερσης για τον κόμβο αυτό.

μια ομάδα κόμβων θα μπορούσε να αντιστοιχεί στο μάτι και να δέχεται διέγερση που συμβολίζει το φωτεινό ερέθισμα. Αντίστοιχα, κόμβοι εκροής, δηλαδή «έξοδος» του δικτύου, είναι οι κόμβοι εκείνοι από τους οποίους επιλέγουμε να λαμβάνουμε το αποτέλεσμα της επεξεργασίας του δικτύου. Συνήθως πρόκειται για μια ομάδα κόμβων διαφορετική από αυτή των κόμβων εισόδου, στην οποία έμμεσα ή άμεσα καταλήγει η διέγερση. Πολλές φορές μπορεί να υπάρχουν και ενδιάμεσοι κόμβοι οι οποίοι δεν δέχονται εξωτερικά ερεθίσματα ούτε αποτελούν μέρος του αποτελέσματος της επεξεργασίας του δικτύου, αλλά παρεμβάλλονται με τις διασυνδέσεις τους μεταξύ εισροής και εκροής. Επειδή οι κόμβοι αυτοί δεν επικοινωνούν απευθείας με το περιβάλλον, ονομάζονται κρυμμένοι ή κρυφοί (hidden). Ο διαχωρισμός των κόμβων ενός δικτύου σε λειτουργικές υποομάδες δεν περιορίζει την κατεύθυνση της ροής των πληροφοριών. Δεν αποκλείεται δηλαδή το ενδεχόμενο «ανάδρομων» συνδέσεων από τους κόμβους εκροής προς τους κόμβους εισροής.

3.3 Προσδιορισμός αναπαραστάσεων και ρύθμιση των συνδέσεων

Έστω ένα σύνολο κόμβων συνδεδεμένων μεταξύ τους κατά τα ανωτέρω και έστω ένα σύνολο ερεθισμάτων από το περιβάλλον. Πώς επιτυγχάνεται η λειτουργία του δικτύου έτσι ώστε να παρατηρείται η επιθυμητή επεξεργασία; Το ερώτημα αυτό είναι κρίσιμης σημασίας, διότι από την απάντηση εξαρτάται το κατά πόσο μπορούμε να κατασκευάσουμε συνδυαστικά μοντέλα για συγκεκριμένες νοητικές λειτουργίες. Πώς μπορούμε να ρυθμίσουμε τα βάρη των συνδέσεων έτσι ώστε οι γενικές αρχές λειτουργίας που περιγράφηκαν παραπάνω να υλοποιούνται για συγκεκριμένα συμπεριφορικά δεδομένα και να εξηγούν έτσι τα αντίστοιχα νοητικά φαινόμενα; Συγκριτικά, για την κατασκευή συμβολικών μοντέλων, ο

ερευνητής διεξάγει μια θεωρητική ανάλυση των υπό διερεύνηση λειτουργιών, ορίζει ρητά τις απαιτούμενες αναπαραστάσεις και προγραμματίζει ένα σύνολο κανόνων έτσι ώστε να επιτελείται η νοητική λειτουργία-στόχος. Για την κατασκευή συνδεδετιστικών μοντέλων, η αντίστοιχη προσέγγιση προϋποθέτει (α) προσδιορισμό των στοιχείων του προβλήματος με όρους ενεργοποίησης κόμβων και (β) ρύθμιση των συνδέσεων μεταξύ των κόμβων έτσι ώστε να προκύπτει η επιθυμητή συμπεριφορά.

Το πρώτο σκέλος της προσέγγισης είναι και το πιο απαιτητικό, καθώς ο ερευνητής πρέπει να αναλύσει θεωρητικά το νοητικό φαινόμενο και να κάνει συγκεκριμένες υποθέσεις και προτάσεις αναφορικά με τις κρίσιμες ιδιότητες. Να εντοπίσει, δηλαδή, από τα χαρακτηριστικά των δεδομένων (ερεθισμάτων και αποκρίσεων) τις πληροφορίες εκείνες, τις συστηματικότητες και τις αποκλίσεις, που θεωρεί ότι δίνουν στο νοητικό φαινόμενο τον ουσιώδη χαρακτήρα του. Αφού επιλεγούν οι θεωρητικά θεμελιώδεις ιδιότητες, που θα πρέπει να συμπεριληφθούν στο μοντέλο, ο ερευνητής στη συνέχεια πρέπει να εκφράσει τις ιδιότητες αυτές ως ενεργοποιήσεις κόμβων, με τέτοιο τρόπο ώστε να διατηρούνται οι κρίσιμες σχέσεις. Δηλαδή, αν μια διάσταση του προβλήματος θεωρείται σημαντική, θα πρέπει η κωδικοποίηση των δεδομένων στους κόμβους να παρουσιάζει ομοιότητες όταν τα στοιχεία του προβλήματος είναι όμοια ως προς αυτή τη διάσταση—και, αντίθετα, η κωδικοποίηση των δεδομένων να είναι διαφορετική όταν τα στοιχεία διαφέρουν σε αυτή τη διάσταση. Η έννοια της ομοιότητας στο συνδεδετισμό είναι σχετικά απλό να οριστεί, καθώς όλα τα στοιχεία είναι ποσοτικά (ενεργοποιήσεις κόμβων), συνεπώς η ομοιότητα μεταξύ δύο «αναπαραστάσεων» (δηλαδή προτύπων ενεργοποίησης σε ένα σύνολο κόμβων) μπορεί να υπολογιστεί αριθμητικά. Για παράδειγμα, μπορεί κανείς να υπολογίσει με απλή αριθμητική αφαίρεση την ποσοτική διαφορά των δύο αναπαραστάσεων ή να τις θεωρήσει ως δύο διανύσματα και να υπολογίσει τη γωνία μεταξύ τους ή τη διαφορά μήκους τους.

Το ζητούμενο του βήματος αυτού είναι να προσδιοριστούν οι «αναπαραστάσεις» του δικτύου, δηλαδή σε τι ακριβώς αντιστοιχεί κάθε κόμβος και κάθε επίπεδο ενεργοποίησης στη σιβάδα εισροής και τη σιβάδα εκροής. Μόλις προσδιοριστεί αυτή η αντιστοιχία (ή «κωδικοποίηση»), το δίκτυο μπορεί να χρησιμοποιηθεί ως μοντέλο, δεχόμενο την κατάλληλη κωδικοποίηση ερεθισμάτων ως εισροή, και αποδίδοντας εκροές που μπορούν να ερμηνευθούν ως αποκρίσεις με την αντίστοιχη αποκωδικοποίηση.

Το δεύτερο βήμα της κατασκευής ενός συνδεδετιστικού μοντέλου είναι η ρύθμιση των βαρών των συνδέσεων. Για το βήμα αυτό υπάρχουν διάφορες τεχνικές «εκπαίδευσης» συνδεδετιστικών μοντέλων. Όλες βασίζονται στη σταδιακή μεταβολή των βαρών των συνδέσεων με κάποια ποσοτικά κριτήρια. Εφαρμόζοντας στο δίκτυο μια εισροή, υπολογίζουμε την ανταπόκριση του δικτύου, δηλαδή τη συνολική κατάσταση ενεργοποίησης όλων των κόμβων και την εκροή. Όταν τα βάρη των συνδέσεων δεν είναι σωστά ρυθμισμένα, η ανταπόκριση αυτή θα διαφέρει από την επιθυμητή. Ξεκινώντας από μια αρχική (συχνά τυχαία) ρύθμιση του δικτύου, το ζητούμενο είναι να υπολογιστεί η απαραίτητη μεταβολή για κάθε βάρος σύνδεσης, ώστε να βελτιωθεί λίγο η συμπεριφορά του δικτύου στο μέλλον. Επιφέροντας μια μικρή μεταβολή, η επίδοση του δικτύου επανελέγχεται στη συνέχεια με νέο ερέθισμα.

μέχρις ότου προσεγγιστεί η επιθυμητή συμπεριφορά ή μηδενιστεί ο ρυθμός βελτίωσης.

Ένας απλός και χρήσιμος τρόπος ρύθμισης των βαρών των συνδέσεων βασίζεται στην ενίσχυση των αποτελεσματικών συνδέσεων και την αποδυνάμωση των άλλων. Όπως και στις συνάψεις των βιολογικών νευρωνικών δικτύων, μια σύνδεση μεταξύ δύο κόμβων θεωρείται αποτελεσματική όταν μεταφέρει διέγερση, και άρα όταν η εισερχόμενη σε ένα κόμβο διέγερση συμπίπτει με διέγερση του ίδιου του κόμβου. Ο βαθμός σύμπτωσης στη δραστηριότητα δύο κόμβων μπορεί να αναπαρασταθεί από το γινόμενο των δραστηριοτήτων τους, $\alpha_i \alpha_j$, το οποίο είναι τόσο πιο θετικό όσο πιο διεγερμένοι ή ανεσταλμένοι είναι οι δύο κόμβοι συγχρόνως (και αντίστοιχα αρνητικό όταν διέγερση του ενός συμπίπτει με αναστολή του άλλου). Για να μεταβάλλουμε το βάρος της σύνδεσης από το i στο j ανάλογα με αυτό το γινόμενο, αρκεί να ορίσουμε τη μεταβολή $\Delta w_{ji} = \epsilon \alpha_i \alpha_j$, όπου με Δ συμβολίζεται η έννοια της μεταβολής ενώ με ϵ συμβολίζεται μια αυθαίρετη ποσότητα που καθορίζει το ρυθμό μεταβολής (όσο μεγαλύτερο το ϵ τόσο περισσότερο μεταβάλλεται το w_{ji}). Με τον τρόπο αυτό καταλήγουμε, μετά από πολλά βήματα, σ' ένα δίκτυο με πολύ ενισχυμένες συνδέσεις μεταξύ των κόμβων που δραστηριοποιούνται ταυτόχρονα. Με παρόμοιο τρόπο μεταβολής και κατάλληλη συνδεσμολογία των κόμβων μπορούμε να πετύχουμε τοπογραφική αναπαράσταση των ερεθισμάτων από το δίκτυο.

Ένας άλλος κανόνας που μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός τεχνητού νευρωνικού δικτύου απαιτεί την ύπαρξη ενός «σήματος διδασκαλίας», και γι' αυτό χαρακτηρίζεται ως *εποπτευόμενη εκπαίδευση* (ή με επίβλεψη, supervised). Χρειάζεται δηλαδή να γνωρίζουμε εκ των προτέρων ποια είναι η σωστή απόκριση που απαιτείται στη σιβάδα εκροής του δικτύου για κάθε πρότυπο ενεργοποίησης που εφαρμόζεται στη σιβάδα εισροής. Όταν έχουμε αυτήν την πληροφορία για κάθε ερέθισμα τότε μπορούμε να υπολογίσουμε το *σφάλμα* (error) κάθε κόμβου, πόσο δηλαδή απέχει η ενεργοποίησή του από το επιθυμητό. Έστω ότι το αποδεκτό αποτέλεσμα για τον κόμβο j της σιβάδας εξόδου είναι t_j ενώ στην πραγματικότητα ο βαθμός διέγερσης του κόμβου αυτού είναι α_j . Η διαφορά μεταξύ των δύο αυτών τιμών, $t_j - \alpha_j$ είναι το σφάλμα του κόμβου. Το σφάλμα αυτό προέρχεται από τις συνδέσεις όλων των άλλων κόμβων προς αυτόν. Κάθε διεγερμένος κόμβος που συνεισέφερε διέγερση στον κόμβο j είναι εν μέρει «υπεύθυνος» για το σφάλμα, και μάλιστα όσο πιο διεγερμένος τόσο πιο υπεύθυνος. Αυτό ισχύει επειδή η συνεισφορά κάθε κόμβου i είναι $\alpha_i w_{ji}$. Άρα πρέπει η μεταβολή του βάρους σύνδεσης από το i στο j να είναι τόσο μεγαλύτερη όσο μεγαλύτερη είναι η διέγερση του i . Τελικά λοιπόν καταλήγουμε στον απλό κανόνα «Δέλτα», συμπεριλαμβάνοντας και μια παράμετρο ρυθμού μεταβολής ϵ όπως και πριν: $\Delta w_{ji} = \epsilon (t_j - \alpha_j) \alpha_i$. Με τον τρόπο αυτό μεταβολής καταλήγουμε (μετά από πολλά βήματα-επαναλήψεις «διδασκαλίας») σ' ένα δίκτυο που για κάθε εισροή πετυχαίνει να προσεγγίσει το σωστό «σχήμα» διέγερσης ως εκροή.

Στην περίπτωση που παρεμβάλλονται ενδιάμεσες σιβάδες κόμβων μεταξύ της εισροής και της εκροής ο κανόνας Δέλτα γίνεται λίγο πιο περίπλοκος όσον αφορά στη μεταβολή των βαρών σύνδεσης προς τους ενδιάμεσους κόμβους. Αυτό συμβαίνει διότι ενώ μπορεί να γνωρίζουμε το σωστό τελικό αποτέλεσμα, την αποδεκτή δηλαδή δραστηριοποίηση των κόμ-

βων εκροής του δικτύου, δεν είναι δυνατό να γνωρίζουμε την επιθυμητή δραστηριοποίηση για τους ενδιάμεσους κόμβους. Έτσι, πρέπει να την υπολογίσουμε προχωρώντας αντίστροφα στο δίκτυο από την έξοδο βάσει των βαρών σύνδεσης, ώστε να βρούμε τη συνεισφορά κάθε ενδιάμεσης σύνδεσης στη διεγερση των ενδιάμεσων κόμβων και να μεταβάλλουμε αντίστοιχα τα βάρη των προηγούμενων συνδέσεων. Η μέθοδος αυτή ονομάζεται *αντίστροφη διάδοση σφάλματος* (backpropagation of error), διότι το σφάλμα εκροής διαδίδεται αντίστροφα, μέσω των βαρών σύνδεσης, προς την είσοδο.

Υπάρχουν και άλλοι κανόνες μεταβολής βαρών σύνδεσης, για την ακρίβεια ολόκληρες οικογένειες διαφορετικών μεθόδων και τεχνικών που έχουν διαφορετικές ιδιότητες και χρησιμοποιούνται σε διαφορετικές περιπτώσεις. Δεν έχουν βέβαια όλες τα ίδια πλεονεκτήματα ή χαρακτηριστικά, π.χ. οι εποπτευόμενες μέθοδοι, που απαιτούν γνώση του ορθού αποτελέσματος, είναι ακατάλληλες όταν το μοντέλο αφορά σε βιολογική ή ψυχολογική λειτουργία που διαμορφώνεται χωρίς αρχική γνώση του τελικού στόχου ή της κατάλληλης στρατηγικής. Τότε εφαρμόζονται τεχνικές ρύθμισης των βαρών χωρίς εποπτεία.

3.4 Είδη και ιδιότητες συνδεδετιστικών μοντέλων

Τεχνητά νευρωνικά δίκτυα χρησιμοποιούνται ως μοντέλα σε όλο το φάσμα των νευροεπιστημών, από τη νευροφυσιολογία μέχρι τη γνωστική ψυχολογία και τα μοντέλα νόησης της γνωσιακής επιστήμης. Σε κάθε κλάδο βέβαια τα χαρακτηριστικά των μοντέλων διαφέρουν ανάλογα με το πεδίο έρευνας. Αντίστοιχα διαφέρουν και οι απόψεις για το τι θεωρείται σημαντικό και τι όχι. Στη γνωσιακή επιστήμη δεν μας ενδιαφέρει η λειτουργία του κάθε νευρώνα ξεχωριστά. Δεν θα ήταν δυνατό άλλωστε να κατασκευάσουμε ρεαλιστικά μοντέλα όσον αφορά στον αριθμό των νευρώνων διότι οι νοητικά ενδιαφέρουσες λειτουργίες επιτελούνται από περίπλοκα δίκτυα δικτύων αποτελούμενα από εκατομμύρια νευρώνες. Συνεπώς τα συνδεδετιστικά μοντέλα στη γνωσιακή επιστήμη δεν αντιστοιχούν, ούτε και υποστηρίζει κανείς ότι πρέπει να αντιστοιχούν, σε βιολογικά νευρωνικά δίκτυα σε λεπτομερειακό επίπεδο. Οι λόγοι που χρησιμοποιούνται έχουν να κάνουν με τις ιδιότητές τους, οι οποίες θεωρούμε ότι αντικατοπτρίζουν τις ιδιότητες των βιολογικών νευρωνικών δικτύων ακόμα και στην περίπτωση που δεν υπάρχει αντιστοιχία μεταξύ των μονάδων.

Με βάση το είδος της λειτουργίας τους τα συνδεδετιστικά μοντέλα διακρίνονται σε διάφορες κατηγορίες, όπως τα συνειρμικά, τα οποία μαθαίνουν μια αυθαίρετη αντιστοιχία μεταξύ ερεθισμάτων (π.χ. σχημάτων) στην είσοδο και αποκρίσεων (π.χ. άλλων σχημάτων) στην έξοδό τους, τα αυτοσυνειρμικά, τα οποία μαθαίνουν να αποθηκεύουν ορισμένα «γνωστά» ερεθίσματα αυτούσια και να μπορούν να τα αναπαράγουν εν είδει μνήμης, τα δίκτυα κατηγοριοποίησης, τα οποία μαθαίνουν να ταξινομούν ένα πλήθος ερεθισμάτων εισροής σε λίγες κατηγορίες εκροής, κ.ά. Τα τελευταία ονομάζονται συχνά και δίκτυα «αναγνώρισης» διότι μπορούν να αναγνωρίσουν μέλη των γνωστών κατηγοριών.

Ταύτιση μνήμης και επεξεργασίας: Σε ένα συνδεδετιστικό μοντέλο δεν υπάρχει διάκριση μεταξύ επεξεργαστή και μνήμης. Η μνήμη του συστήματος είναι το σύνολο των βαρών των συνδέσεων μεταξύ των κόμβων, αφού στα βάρη αυτά οφείλονται οι εκάστοτε ιδιότητες του

συστήματος, αυτά μεταβάλλονται κατά τις διεργασίες μάθησης και σε αυτά είναι «αποθηκευμένη» η μαθησιακή ιστορία του συστήματος. Όμως τα βάρη των συνδέσεων δεν είναι μόνο μνήμη αφού κάθε στιγμή η απόκριση του συστήματος σε κάποιο ερέθισμα εξαρτάται από τη διέλευση της ενεργοποίησης των κόμβων μέσα από τις μεταξύ τους συνδέσεις, ενώ η διαμόρφωση της ενεργοποίησης κάθε κόμβου είναι αποτέλεσμα των διελεύσεων αυτών ανάλογα με το βάρος κάθε σύνδεσης. Δηλαδή, κατά τη λειτουργία του συστήματος, η κατάστασή του διαμορφώνεται διαμέσου των συνδέσεων μεταξύ των κόμβων όπου και βρίσκεται η «μνήμη» του. Η διαδικασία της μάθησης έχει ως αποτέλεσμα τη μεταβολή της μελλοντικής λειτουργίας όχι έμμεσα (επειδή κάποιο πρόγραμμα θα ανατρέξει να ανασύρει τη σχετική πληροφορία και να τη χειριστεί ανάλογα) αλλά άμεσα, επειδή η ίδια η λειτουργική διαμόρφωση του συστήματος μεταβάλλεται. Μ' άλλα λόγια, σε ένα συνδεδειστικό μοντέλο, μνήμη είναι η ιδιότητα του συστήματος να μεταβάλλει τη λειτουργία του ως αποτέλεσμα της λειτουργίας του, σε μια ενιαία και αδιάσπαστη θεώρηση. Η ιδιότητα αυτή βρίσκεται σε αντίθεση με τη συνήθη σχέση επεξεργασίας και μνήμης σε ένα συμβολικό υπολογιστικό σύστημα, όπου η μνήμη είναι ανεξάρτητη και άσχετη από τη λειτουργία, ενώ και η μνήμη λειτουργιών (των προγραμμάτων και των παραμέτρων τους) μπορεί να είναι ανεξάρτητη από τη μνήμη «δεδομένων».

Σε ένα συνδεδειστικό μοντέλο ο διαχωρισμός μεταξύ μακρόχρονης και εργαζόμενης μνήμης μπορεί να είναι εντελώς φυσικός, αν θεωρήσουμε ως εργαζόμενη μνήμη την τρέχουσα κατάσταση ενεργοποίησης του συστήματος και ως μακρόχρονη μνήμη το σχέδιο συνδέσεων μεταξύ των κόμβων. Έτσι η εργαζόμενη μνήμη αντανακλά κάθε στιγμή την τρέχουσα λειτουργία και τις αναπαραστασιακές ποσότητες που λαμβάνουν μέρος σε αυτή, ενώ η μακρόχρονη μνήμη αντιστοιχεί στη δυναμικότητα του συστήματος για τις διάφορες λειτουργίες του, τρέχουσες και μελλοντικές. Αντίθετα, σε ένα συμβολικό υπολογιστικό σύστημα ο διαχωρισμός μεταξύ εργαζόμενης και μακρόχρονης μνήμης είναι εξωγενής και λειτουργικά αυθαίρετος, αφού περιορίζεται στον εξωτερικό προσδιορισμό μιας αποθηκευτικής μονάδας ως «βραχύχρονης» και μιας άλλης ως «μακρόχρονης».

Τοπική και κατανεμημένη αναπαράσταση: Σε ένα συνδεδειστικό δίκτυο κάθε κόμβος μπορεί, λόγω της σχεδίασης του συστήματος, να αντιστοιχεί σε (δηλαδή να αναπαριστά) κάτι συγκεκριμένο. Για παράδειγμα μπορεί η ενεργοποίησή του να αντιστοιχεί στην έκφραση μιας συγκεκριμένης έννοιας, ενός χαρακτηριστικού, ή μιας νοητικής ποσότητας που ενδιαφέρει τον κατασκευαστή του μοντέλου. Στην περίπτωση αυτή έχουμε τη λεγόμενη τοπική (local) ή εντοπιστική (localist) αναπαράσταση. Αν ένας τέτοιος κόμβος καταστραφεί τότε χάνεται εντελώς από το σύστημα η δυνατότητα αναπαράστασης της έννοιας που αντιστοιχούσε στον κόμβο αυτόν.

Συχνά, σε σύγχρονα συνδεδειστικά μοντέλα, χρησιμοποιείται *κατανεμημένη αναπαράσταση* (distributed representation), δηλαδή κάθε κόμβος συμμετέχει στην αναπαράσταση πολλών εννοιών ή ποσοτήτων και κάθε έννοια ή ποσότητα αναπαρίσταται μέσω του σχεδίου ενεργοποίησης πολλών κόμβων. Οι πληροφορίες μπορούν να αναπαρίστανται σε ολόκληρα δίκτυα, αποθηκευμένες στις συνδέσεις, με τρόπο που στοιχεία για κάθε πληροφορία να

βρίσκονται σε πολλές συνδέσεις και κάθε σύνδεση να συμμετέχει στην αναπαράσταση πολλών πληροφοριών. Τελικά η συνολική συνδεσμολογία του δικτύου κωδικοποιεί το σύνολο των αποθηκευμένων πληροφοριών, έχουμε δηλαδή *κατανεμημένη μνήμη*. Μια ενδιάμεση περίπτωση είναι όταν κάθε κόμβος αντιστοιχεί σε ένα επιμέρους χαρακτηριστικό (τοπική αναπαράσταση σε «κατώτερο» επίπεδο) και έτσι συμμετέχει ενεργά σε όλες τις έννοιες που διαθέτουν το χαρακτηριστικό αυτό (κατανεμημένη αναπαράσταση σε «ανώτερο» επίπεδο).

Ως συνέπεια της παράλληλης κατανεμημένης επεξεργασίας και μνήμης, στα συνδεδεμένα δίκτυα παρατηρούνται οι εξής ιδιότητες:

- *Ευαισθησία στην ομοιότητα*, δηλαδή για παρόμοια ερεθίσματα παράγονται παρόμοια αποτελέσματα, ακόμα και αν πρόκειται για ερεθίσματα που δεν έχουν παρουσιαστεί στο παρελθόν. Χάρη στην ιδιότητα αυτή ένα συνδεδετιστικό μοντέλο μπορεί να κατηγοριοποιήσει σωστά άγνωστα ερεθίσματα και να παρουσιάσει τα κλασικά φαινόμενα προτύπων της γνωστικής ψυχολογίας, αλλά και να παράγει αποδεκτές αποκρίσεις σε ασαφείς ή αόριστες εισροές. Η λειτουργία βάσει ομοιότητας απορρέει αυτόματα από την κατανεμημένη επεξεργασία και είναι στατιστική ιδιότητα των μοντέλων.
- *Εξάρτηση από το περιεχόμενο*, που σε μεγάλο βαθμό είναι η άλλη όψη της ευαισθησίας στην ομοιότητα. Σε αντίθεση με τη συντακτική επεξεργασία των συμβολικών μοντέλων, όπου το σημασιακό περιεχόμενο δεν έπαιζε ρόλο στο χειρισμό του συμβόλου, σε ένα συνδεδετιστικό μοντέλο οι αναπαραστάσεις ενεργοποιούνται και συμμετέχουν σε «υπολογισμούς» βάσει του περιεχομένου τους, διότι δεν υπάρχει διάκριση μεταξύ συντακτικών και σημασιακών ιδιοτήτων, καθώς το μόνο που καθορίζει κάθε «αναπαράσταση» είναι οι ποσοτικές ενεργοποιήσεις των κόμβων του δικτύου. Όμοιες αναπαραστάσεις αντιστοιχούν σε όμοιες σημασίες και προκαλούν όμοια αποτελέσματα στη λειτουργία του μοντέλου. Έτσι η γενίκευση είναι απόρροια της κανονικότητας και ισχύει στο βαθμό που υποστηρίζεται από στατιστική συστηματικότητα. Το σύστημα μπορεί να παρουσιάζει κανονικότητα χωρίς να περιλαμβάνει ρητούς κανόνες και ταυτόχρονα μπορεί να κωδικοποιεί τις «εξαιρέσεις» των «κανόνων» χωρίς ειδικές λειτουργικές υποθέσεις.
- *Αναισθησία στο θόρυβο και σε τοπικές ή διάχυτες βλάβες*, δηλαδή (α) ανοχή σε σημαντική αλλοίωση του ερεθίσματος χωρίς σημαντική επιδείνωση της λειτουργίας του δικτύου, και (β) ανοχή σε μερική απώλεια συνδέσεων χωρίς σημαντικές επιπτώσεις στη συνολική λειτουργία. Οι ιδιότητες αυτές απορρέουν από την κατανεμημένη επεξεργασία: εφόσον για την παραγωγή του αποτελέσματος λαμβάνεται υπόψη η συνεισφορά πλήθους κόμβων, αν εκλείψουν ορισμένοι ή αν αλλοιωθούν κάποια από τα χαρακτηριστικά εισροής υπερσχύει το πλήθος των υπολοίπων που λειτουργούν κανονικά.
- *Ευπλασιότητα* (ή πλαστικότητα, plasticity) και *αυτοοργάνωση* (self-organization), δηλαδή δυνατότητα μεταβολής της λειτουργίας του δικτύου, τέτοιας ώστε το αποτέλεσμα να αντικατοπτρίζει κάποια εξωτερική δομή του περιβάλλοντος με βάση τις στατιστικές ιδιότητες των ερεθισμάτων που το δίκτυο επεξεργάστηκε από τις εισόδους του. Η μεταβολή της λειτουργίας επέρχεται μέσω ρύθμισης των βαρών των συνδέσεων, η οποία μπορεί να γίνεται διαρκώς κατά τη λειτουργία του συστήματος.

- *Τεράστια στατιστική ισχύς* για την κωδικοποίηση περιβαλλοντικών ερεθισμάτων, δηλαδή μεγάλη χωρητικότητα που επαρκεί για την απεικόνιση τόσο των γενικών χαρακτηριστικών (ώστε να λειτουργεί βάσει ομοιότητας) όσο και των ειδικών χαρακτηριστικών και λεπτομερειών (ώστε να μη γίνονται λάθη στις περιπτώσεις γνωστών εξαιρέσεων από τον κανόνα της ομοιότητας).

Προσθιοδρομικά (feedforward) και *αναδρομικά* (ή επαναδρομικά, recurrent) μοντέλα: Συνηθίζεται, όπως προαναφέρθηκε, οι κόμβοι στα συνδεδεστικά μοντέλα να ομαδοποιούνται κατά στάδια, και να ορίζεται μια κατευθυντικότητα. Ένα στάδιο θεωρείται ότι λαμβάνει ερεθίσματα από το περιβάλλον, και ονομάζεται στάδιο εισροής, ενώ ένα άλλο στάδιο θεωρείται ότι παράγει αποκρίσεις, και ονομάζεται στάδιο εκροής. Συχνά, για λόγους ευκολίας υπολογισμών, οι συνδέσεις μεταξύ των κόμβων ακολουθούν αυστηρά μία κατεύθυνση, από την εισροή προς την εκροή. Στην περίπτωση αυτή κάθε κόμβος μπορεί να συνδέεται προς (δηλαδή να τροφοδοτεί) μόνο κόμβους σε επόμενα στάδια, και συνηθέστερα μόνο κόμβους στο αμέσως επόμενο στάδιο. Στην περίπτωση αυτή έχουμε να κάνουμε με *προσθιοδρομικό* δίκτυο ή δίκτυο *προσωδιάδοσης* (feedforward). Στα προσθιοδρομικά μοντέλα, η παρουσίαση ενός ερεθίσματος οδηγεί στην άμεση παραγωγή της αντίστοιχης απόκρισης διασχίζοντας το δίκτυο κατά την ορθή φορά (από την εισροή προς την εκροή) χωρίς δυνατότητα σταδιακής παραγωγής της απόκρισης ή αμφιταλαντεύσεων. Το δίκτυο παράγει υποχρεωτικά μία και μόνο απόκριση, πάντα την ίδια για το ίδιο ερέθισμα. Η «διδασκαλία» (ρύθμιση) ενός τέτοιου μοντέλου είναι σχετικά απλή, όμως το εύρος των υπολογιστικών του δυνατοτήτων είναι περιορισμένο.

Όταν σε ένα μοντέλο υπάρχει τουλάχιστον μία σύνδεση από κάποιον κόμβο προς κόμβο προηγούμενου ή του ίδιου σταδίου (ή προς τον εαυτό του) τότε δημιουργείται στο σύστημα ένας βρόχος ενεργοποίησης, με την έννοια ότι ένα μέρος τουλάχιστον του δικτύου μπορεί να είναι αυτοτροφοδοτούμενο. Στην περίπτωση αυτή το σύστημα ονομάζεται «αναδρομικό» ή «επαναδρομικό» (recurrent ή recursive) και η σταθερότητά του δεν είναι εγγυημένη, διότι εξαρτάται από τις σχέσεις μεταξύ ανασταλτικών και διεγερτικών συνδέσεων. Η παρουσίαση ενός ερεθίσματος μπορεί να οδηγήσει σε διαφορετικές αποκρίσεις ανάλογα με την τρέχουσα κατάσταση του συστήματος κατά τη στιγμή της παρουσίασης. Ο υπολογισμός της απόκρισης μπορεί να είναι περίπλοκος και μεγάλης διάρκειας, διότι δεν αρκεί μια διαδρομή από την εισροή προς την εκροή αλλά πρέπει να γίνουν πολλές επαναλήψεις του βρόχου ανατροφοδότησης μέχρις ότου σταθεροποιηθεί η απόκριση του συστήματος (αν τελικά συμβεί αυτό). Η διδασκαλία ενός τέτοιου μοντέλου μπορεί να είναι αρκετά δύσκολη, υποκείμενη σε αστάθειες, το εύρος όμως των υπολογιστικών δυνατοτήτων τέτοιων μοντέλων είναι ιδιαίτερα αυξημένο και ανώτερο πολλών κλασικών υπολογιστικών μηχανών.

3.5 Περίληψη

Τα συνδεδεστικά υπολογιστικά συστήματα χαρακτηρίζονται από ένα σύνολο κόμβων με συνδέσεις μεταξύ τους. Η λειτουργία τους περιγράφεται πλήρως από έναν πίνακα ενεργοποιήσεων (των κόμβων) και έναν πίνακα βαρών (των συνδέσεων). Πρόκειται δηλαδή για

ποσοτικά συστήματα, που λειτουργούν με αριθμητικές πράξεις πάνω στο περιεχόμενο των αναπαραστάσεων, όπου ως αναπαραστάσεις εννοούνται τα σχέδια ενεργοποίησης των κόμβων. Παρότι οι κόμβοι είναι διακριτές μονάδες επεξεργασίας, οι αναπαραστάσεις είναι διαβαθμισμένες (καθώς εκφράζονται ποσοτικά, από το βαθμό ενεργοποίησης των κόμβων) και κατανεμημένες (καθώς εξαρτώνται από την ενεργοποίηση πολλών κόμβων). Η αιτιακή ισχύς των αναπαραστάσεων οφείλεται στην άμεση συμμετοχή τους στις αριθμητικές πράξεις που καθορίζουν το αποτέλεσμα των υπολογισμών, δηλαδή στην εκροή του δικτύου. Το σημαντικότερο πλεονέκτημα αυτών των συστημάτων είναι ότι εξαιτίας της διαβάθμισης και της διασποράς των αναπαραστάσεων εμφανίζεται φυσική ευαισθησία στην ομοιότητα, ενώ και η έννοια της ομοιότητας μπορεί να οριστεί ποσοτικά. Έτσι, η κανονικότητα της συμπεριφοράς δεν αποδίδεται σε θεωρητική παραγωγικότητα, καθώς το σύστημα δεν αναπαριστά αφηρημένους κανόνες, αλλά σε συστηματικές σχέσεις μεταξύ εισροών και εκροών.

Στο συνδεδετισμό θεωρείται ότι βάση των νοητικών φαινομένων είναι τα στατιστικά χαρακτηριστικά των προβλημάτων και των λύσεών τους. Αποδίδουμε δηλαδή μέγιστη επεξηγηματική ισχύ στα χαρακτηριστικά του προβλήματος και στην ικανότητα των νοητικών συστημάτων να εσωτερικεύουν και να εκμεταλλεύονται αυτά τα χαρακτηριστικά. Παράλληλα, υποβαθμίζεται ο ρόλος των εγγενών αρχιτεκτονικών και των προσχεδιασμένων περιορισμών, χωρίς όμως να εκμηδενίζεται. Η προσομοίωση του ανθρώπινου νου με όρους εγκεφαλικής λειτουργίας νοείται ταυτόχρονα σε δύο επίπεδα: Αφενός, ο εγκέφαλος αποτελεί το φυσικό σύστημα υλοποίησης του νου. Αφετέρου, με βάση τις ίδιες αρχές λειτουργίας, τα νοητικά φαινόμενα περιγράφονται ως αποτέλεσμα αλληλεπίδρασης επεξεργαστικών μονάδων και επεξεργασίας αναπαραστασιακού περιεχομένου. Έτσι, ενώ η απάντηση του συμβολισμού στο ερώτημα «τι μηχανή είναι ο νους;» είναι «συντακτικός επεξεργαστής συμβόλων» (σαν ηλεκτρονικός υπολογιστής, ή μηχανή Turing), η απάντηση του συνδεδετισμού στο ίδιο ερώτημα είναι «δίκτυο αλληλοσυνδεδεμένων μονάδων που αλληλεπιδρούν» (σαν νευρωνικό δίκτυο, ή μηχανή στατιστικής γενίκευσης).

Ένα σημείο που πρέπει να τονιστεί ιδιαίτερα είναι η σχέση του συνδεδετισμού με φαινόμενα μάθησης. Είναι εύκολο να παρανοήσει κανείς τη διαδικασία ρύθμισης των βαρών, η οποία αποκαλείται φάση «εκπαίδευσης» του δικτύου, και να θεωρήσει ότι τα συνδεδετιστικά δίκτυα είναι μοντέλα μάθησης. Παρότι μπορούν να είναι μοντέλα μάθησης, και μάλιστα με μεγάλη επιτυχία, αυτό δεν είναι θέμα αλγορίθμου ρύθμισης αλλά εξαρτάται από το στόχο και τη σχεδίαση της προσομοίωσης. Μοντέλο μάθησης έχουμε μόνο όταν ο ερευνητής σχεδιάσει την αλληλουχία παρουσίασης ερεθισμάτων στο δίκτυο έτσι ώστε αυτή να αντιστοιχεί σε κάποια αναπτυξιακή ή μαθησιακή διαδικασία, και όταν μελετά την πορεία της επίδοσης του δικτύου κατά τη διάρκεια της εκπαίδευσης. Διαφορετικά η φάση «εκπαίδευσης» δεν είναι παρά ο αλγοριθμικός τρόπος με τον οποίο ρυθμίζεται το δίκτυο ώστε να αποκτήσει τη μορφή που απαιτείται για τις επακόλουθες προσομοιώσεις. Η παρανόηση της σχέσης του συνδεδετισμού με τη μάθηση έχει και αντίκτυπο στις άλλες προσεγγίσεις, καθώς μπορεί να υποθέσει κανείς—λανθασμένα—ότι ο συνδεδετισμός έχει κάποιο προνόμιο στην προσομοίωση μαθησιακών διαδικασιών. Κάτι τέτοιο όμως δεν ευσταθεί διότι όλες

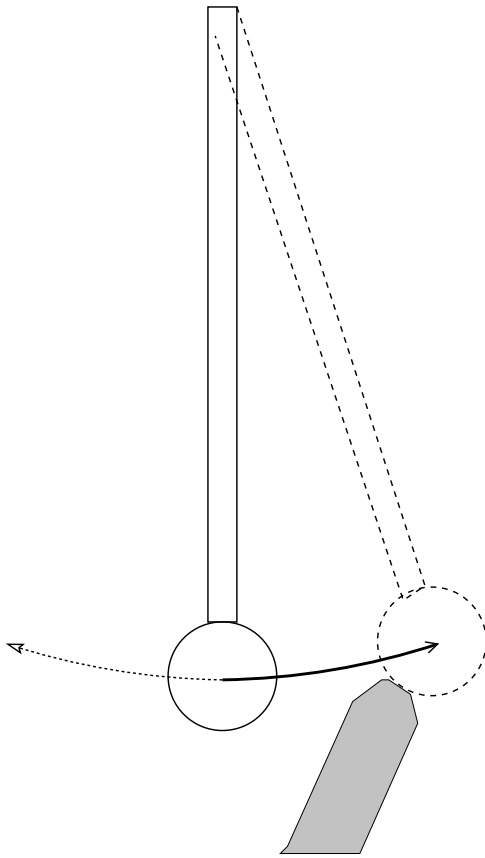
οι προσεγγίσεις έχουν τους δικούς τους μηχανισμούς μάθησης, δηλαδή μεταβολής των στοιχείων τους σύμφωνα με την «εμπειρία» τους. Η μάθηση παίρνει διαφορετική μορφή σε κάθε είδος συστήματος, φυσικά, αλλά δεν θα ήταν δυνατό να μιλάμε για προσεγγίσεις στη νόηση χωρίς ικανότητα μάθησης. Εκείνο που διακρίνει το συνδετισμό είναι οι συγκεκριμένοι τρόποι με τους οποίους τα δίκτυα μπορούν να ενσωματώσουν πληροφοριακές δομές του περιβάλλοντος μέσα από διαδικασίες σταδιακής αυτοδιόρθωσης και αυτοοργάνωσης.

4 Δυναμικά συστήματα

Από τα συμβολικά συστήματα απουσιάζει εντελώς η έννοια του χρόνου και της σταδιακής εξέλιξης των καταστάσεων. Συνεπώς δεν υπάρχει δυνατότητα έκφρασης διαρκειών, καθυστερήσεων, ή χρονικών στιγμών μετάβασης από μία κατάσταση σε μια άλλη, ούτε δυνατότητα περιγραφής ενδιάμεσων ή ασαφών καταστάσεων. Δεν έχει λοιπόν κανένα νόημα να αναφερόμαστε σε χρόνο λειτουργίας ή χρόνο απόκρισης μέσα στα πλαίσια των ιδιοτήτων των συστημάτων. Υπάρχει μόνο μία αλληλουχία διαδοχικών καταστάσεων, ώστε για δύο οποιεσδήποτε καταστάσεις να είναι σαφές ποια από τις δύο προηγείται. Αντιστοιχίζοντας ένα σταθερό χρονικό διάστημα σε κάθε εναλλαγή μεταξύ διαδοχικών καταστάσεων (κάθε «βήμα») μπορούμε να πετύχουμε μια εξωγενή προσέγγιση στη διάσταση του χρόνου—εξωγενή, επειδή ο χρόνος δεν εκφράζεται εσωτερικά μέσα στο ίδιο το σύστημα. Το τέχνασμα αυτό χρησιμοποιείται προκειμένου να εκφράσουμε υπολογιστικά χαρακτηριστικά σε σχέση με παρατηρήσιμα χρονικά μεγέθη από την πειραματική ψυχολογία (π.χ. χρόνο απόκρισης). Δεν μπορεί όμως ποτέ σε ένα συμβολικό υπολογιστικό σύστημα να αναπαρασταθεί εγγενώς ούτε η συνέχεια ούτε η ομοιότητα, τόσο στο χρόνο όσο και στις καταστάσεις. Αντίθετα, σε ένα φυσικό σύστημα που εξελίσσεται διαρκώς στο χρόνο, η ποσοτική συνέχεια στις τέσσερις διαστάσεις είναι θεμελιώδης ιδιότητα του συστήματος.

4.1 Τροχιές στο χώρο των φάσεων

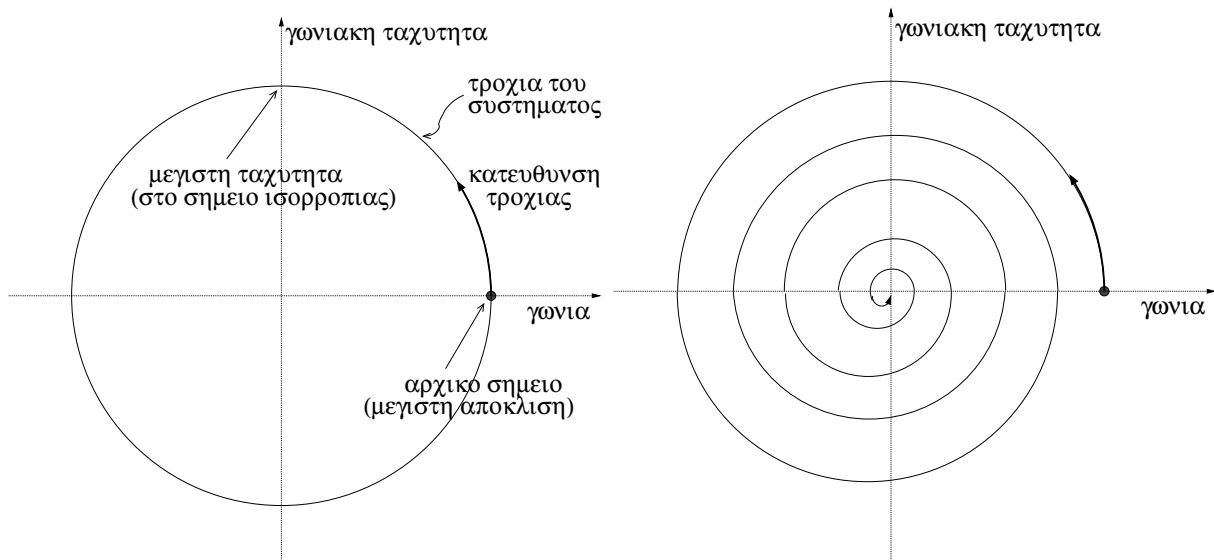
Η πιο σημαντική έννοια για την κατανόηση των δυναμικών συστημάτων, και κατ' επέκταση της δυναμικής προσέγγισης στο νου, είναι η έννοια της τροχιάς που διαγράφει η κατάσταση του συστήματος με την πάροδο του χρόνου. Για να αντιληφθούμε την έννοια αυτή είναι σκόπιμο να εξετάσουμε πρώτα ένα απλό και οικείο φυσικό σύστημα. Στο Σχήμα 4 φαίνεται ένα απλό εκκρεμές, το οποίο αποτελείται από μία ράβδο με ένα βάρος στην άκρη της. Το σύστημα αυτό μπορεί να περιγραφεί πλήρως οποιαδήποτε στιγμή αν γνωρίζουμε δύο πράγματα: τη θέση του και την ταχύτητα της κίνησής του (ακριβέστερα, τη γωνία απόκλισης και τη γωνιακή ταχύτητα, αλλά δεν έχει τόση σημασία για το σκοπό μας). Το εκκρεμές μπορεί να βρίσκεται στην ίδια θέση με διαφορετικές ταχύτητες, καθώς άλλοτε κινείται προς τα δεξιά και άλλοτε προς τα αριστερά. Επίσης μπορεί να έχει την ίδια ταχύτητα σε διαφορετικές θέσεις. Άρα μας χρειάζονται και οι δύο αυτές πληροφορίες, κάθε στιγμή. Δεν υπάρχει καμία άλλη πληροφορία που να προσδιορίζει το σύστημα, καθώς με



Σχήμα 4: Σχηματικό διάγραμμα ενός απλού δυναμικού συστήματος το οποίο αποτελείται από μια ράβδο με ένα βάρος στο κάτω μέρος της. Το σύστημα αρχικά παραμένει σταθερά ακίνητο στη θέση ισορροπίας (συνεχείς γραμμές). Διαταράσσουμε την ισορροπία του συστήματος με ένα ξένο σώμα (ανοιχτό γκρι στο σχήμα), όπως φαίνεται στο σχήμα (διακεκομμένες γραμμές), προκαλώντας μια απόκλιση στο βάρος. Αποσύροντας την εξωτερική παρέμβαση, το σύστημα μένει ελεύθερο να εξελιχθεί στο χρόνο υπό την επίδραση της βαρύτητας. Το αποτέλεσμα θα είναι μια επαναλαμβανόμενη ταλάντωση, δεξιά-αριστερά. Το βάρος θα μηδενίζει την ταχύτητά του στο δεξιότερο και το αριστερότερο σημείο της ταλάντωσης, ενώ θα έχει μέγιστη ταχύτητα στο αρχικό σημείο ισορροπίας, καθώς φτάνει κατερχόμενο ως εκεί και εν συνεχεία ανέρχεται (ενάντια στη βαρύτητα) προς την αντίθετη πλευρά. Αν δεν υπάρχουν τριβές στο σύστημα τότε η ταλάντωση θα διατηρεί το πλάτος της σε κάθε επαναφορά και θα συνεχιστεί για πάντα. Αν υπάρχουν απώλειες τότε η ταλάντωση θα είναι όλο και μικρότερη, μέχρις ότου το σύστημα κάποτε σταματήσει εντελώς και παραμείνει ακίνητο στην αρχική θέση ισορροπίας.

αυτές τις δύο γνωρίζουμε τα πάντα. Άρα, πρόκειται για ένα σύστημα δύο *διαστάσεων*.

Οι δύο διαστάσεις του συστήματος ονομάζονται «φάσεις» και ο χώρος που προσδιορίζεται από τις διαστάσεις αυτές ονομάζεται *χώρος φάσεων* (phase space) ή *χώρος καταστάσεων* (state space). Στο Σχήμα 5 απεικονίζουμε τις δυο διαστάσεις του συστήματος πάνω στο χαρτί σημειώνοντας δύο γραμμές ως άξονες. Το σημείο όπου συναντιούνται οι άξονες αντιστοιχεί σε μηδενική γωνία και μηδενική ταχύτητα, άρα αντιστοιχεί στην ισορροπία του συστήματος. Η διαταραχή του συστήματος που σημειώνεται στο Σχήμα 4 ισοδυναμεί με μια μετακίνηση της θέσης προς τα δεξιά. Όσο κρατάμε το εκκρεμές ακίνητο στην αποκλίνουσα θέση, αυτό σημειώνεται με ένα σημείο στο χώρο φάσεων που αντιστοιχεί σε δεξιά απόκλιση και μηδενική ταχύτητα, όπως το σημείο του Σχήματος 5 (αριστερά) με την ένδειξη «αρχικό σημείο». Μόλις αποσύρουμε τη διαταραχή, το εκκρεμές θα αρχίσει να κινείται προς τα αριστερά υπό την επίδραση της βαρύτητας, αυξάνοντας την ταχύτητά του προς τα αριστερά και ταυτόχρονα μειώνοντας την απόκλισή του από τη θέση ισορροπίας. Η αύξηση της ταχύτητας ισοδυναμεί στο χώρο φάσεων με μετακίνηση του σημείου προς τα πάνω, ενώ η μείωση της απόκλισης με μετακίνηση προς τα αριστερά. Η μετακίνηση



Σχήμα 5: Διάγραμμα φάσεων για το εκκρεμές του Σχήματος 4. Στον οριζόντιο άξονα απεικονίζεται η απόκλιση του συστήματος από τη θέση ισορροπίας, άρα όσο πιο δεξιά βρίσκεται κάποιο σημείο τόσο πιο δεξιά αποκλίνει το εκκρεμές. Στον κατακόρυφο άξονα απεικονίζεται η ταχύτητα του συστήματος, ώστε όσο πιο ψηλά βρίσκεται κάποιο σημείο τόσο πιο γρήγορη κίνηση προς τα αριστερά του αντιστοιχεί. *Αριστερά*, η περίπτωση που στο εκκρεμές δεν υπάρχουν απώλειες (τριβές κλπ.). Το σύστημα διαγράφει κυκλική τροχιά, καθώς επανέρχεται περιοδικά στην αρχική κατάσταση. Οι μέγιστες αποκλίσεις (δεξιότερο και αριστερότερο σημείο του κύκλου) αντιστοιχούν σε μηδενική ταχύτητα, καθώς πέφτουν πάνω στον οριζόντιο άξονα. Οι μέγιστες ταχύτητες (υψηλότερο και χαμηλότερο σημείο του κύκλου) αντιστοιχούν σε μηδενική απόκλιση, δηλαδή στο σημείο ισορροπίας. *Δεξιά*, η περίπτωση που στο εκκρεμές υπάρχουν απώλειες. Το σύστημα διαγράφει σπειροειδή τροχιά πλησιάζοντας στο κέντρο σε κάθε ταλάντωση, καθώς χάνεται ενέργεια κι έτσι μειώνεται τόσο η απόκλιση όσο και η μέγιστη ταχύτητα. Κάποτε το σύστημα φτάνει στο κέντρο του χώρου και σταματά (απόκλιση μηδέν και ταχύτητα μηδέν).

αυτή, σημείο προς σημείο, σημειώνεται με έντονο μαύρο καμπύλο βέλος, το οποίο ξεκινά από το «αρχικό σημείο» και φέρει την ένδειξη «κατεύθυνση τροχιάς». Αφήνοντας το εκκρεμές να εξελιχθεί ελεύθερο μέσα στο χρόνο, το σημείο που του αντιστοιχεί σε κάθε στιγμή διαγράφει έναν κύκλο. Το βασικό είναι να γίνει κατανοητό ότι (α) ο χώρος των φάσεων απεικονίζει όλες τις δυνατές καταστάσεις του συστήματος, καθώς κάθε σημείο αντιστοιχεί σε μια δυνατή κατάσταση, και (β) η κατάσταση του συστήματος κάθε στιγμή αντιστοιχεί σε ένα συγκεκριμένο σημείο του χώρου των φάσεων, άρα η χρονική εξέλιξη της κατάστασης του συστήματος αντιστοιχεί σε μια γραμμή, η οποία σχηματίζεται από τα διαδοχικά σημεία, δηλαδή τις διαδοχικές καταστάσεις. Η γραμμή αυτή ονομάζεται *τροχιά* του συστήματος και απεικονίζει πλήρως την εξέλιξη της κατάστασής του.

Η τροχιά του συστήματος δεν εξαρτάται από εξωτερικές παρεμβάσεις, καθώς αποτελεί μια απεικόνιση της φυσικής του εξέλιξης. Το σύστημα είναι πλήρως *ντετερμινιστικό*,

δηλαδή καθορισμένο αιτιοκρατικά. Αν γνωρίζουμε το αρχικό σημείο της τροχιάς και τις ιδιότητες του συστήματος (λ.χ. αν υπάρχουν απώλειες ή όχι) τότε αυτομάτως και εξαρχής είναι δεδομένο ποια θα είναι η τροχιά. Από οποιοδήποτε πιθανό αρχικό σημείο, υπάρχει και είναι προκαθορισμένη η τροχιά που θα ακολουθήσει το σύστημα ξεκινώντας από την αντίστοιχη κατάσταση. Είναι σαν ο χώρος των φάσεων να είναι γεμάτος αόρατες γραμμές, οι οποίες σχηματίζουν τις δυνατές τροχιές του συστήματος, καθώς αυτές είναι πλήρως προκαθορισμένες από τις ιδιότητες του συστήματος. Ο χώρος των φάσεων με τις δυνατές τροχιές αποτελεί μια πλήρη περιγραφή του συστήματος. Προσέξτε ότι ο χρόνος δεν απεικονίζεται στο χώρο των φάσεων, διότι δεν είναι ούτε αιτία ούτε παράμετρος στο σύστημα. Η μεταβολή του συστήματος με την πάροδο του χρόνου κάνει το σύστημα να εμπεριέχει τη χρονικότητα μέσα στην εξέλιξή του. Κάθε σημείο της τροχιάς αντιστοιχεί σε μια χρονική στιγμή, όμως η ίδια η τροχιά δεν απεικονίζει το χρόνο ούτε και το ρυθμό με τον οποίο τη διατρέχει η κατάσταση του συστήματος.

Κάποια σημεία του χώρου των φάσεων έχουν ιδιαίτερα χαρακτηριστικά. Στο παράδειγμα με το εκκρεμές, ένα τέτοιο «ιδιαιτερο» σημείο είναι το κέντρο του χώρου, όπου συναντιούνται οι άξονες. Το σημείο αυτό είναι ιδιαίτερο διότι είναι σταθερό σημείο. Αυτό σημαίνει πως δεν φεύγει καμία τροχιά από αυτό. Αν η αρχική κατάσταση του συστήματος είναι αυτό το σημείο τότε το σύστημα θα παραμείνει εκεί για πάντα (με άλλα λόγια, αν δεν μετακινήσουμε το εκκρεμές από τη θέση ισορροπίας αυτό δεν πρόκειται να ξεκινήσει να κινηθεί από μόνο του). Στην περίπτωση που το σύστημα έχει απώλειες, το σημείο αυτό έχει και ένα ακόμα χαρακτηριστικό: όλες οι κοντινές τροχιές καταλήγουν σε αυτό. Δηλαδή, από οποιοδήποτε αρχικό σημείο κι αν ξεκινήσει το σύστημα η τροχιά του θα το οδηγήσει κάποτε στο κέντρο (με άλλα λόγια, όσο μεγάλη ταλάντωση και αν έχει στην αρχή, το εκκρεμές κάποτε θα σταματήσει). Τα σημεία στα οποία καταλήγουν οι τροχιές της περιοχής τους ονομάζονται *ελκυστές* (attractors) διότι «τραβούν» επάνω τους την εξέλιξη της κατάστασης. Στο απλό σύστημα του εκκρεμούς ο μοναδικός ελκυστής τραβά όλες τις τροχιές του χώρου. Σε πιο σύνθετα συστήματα ένας ελκυστής δεν έχει τόσο καθολική επίδραση, αλλά σε αυτόν καταλήγουν μόνο οι τροχιές που περνούν από σχετικά κοντινή απόσταση και βρίσκονται μέσα σε μια περιοχή που ονομάζεται «λεκάνη έλξης» (attractor basin).

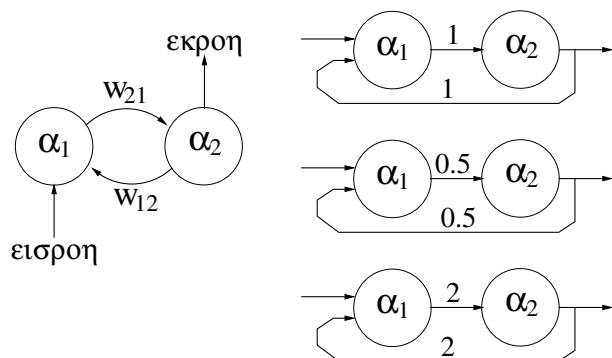
4.2 Συνεχής και διακριτός χρόνος

Ας αναλογιστούμε στη συνέχεια ένα αναδρομικό συνδεδετιστικό δίκτυο δύο κόμβων, όπως στο Σχήμα 6, αριστερά. Στο δίκτυο αυτό υπάρχουν δύο ποσότητες που μεταβάλλονται, οι οποίες προσδιορίζουν την κατάσταση του συστήματος κάθε στιγμή: είναι οι ενεργοποιήσεις των κόμβων, α_1 και α_2 . Συνεπώς πρόκειται για ένα σύστημα δύο διαστάσεων, και ο χώρος των φάσεων θα ορίζεται από μία διάσταση α_1 και μία διάσταση α_2 , όπως στο Σχήμα 7. Η κατάσταση του συστήματος θα είναι ένα σημείο στο χώρο αυτό, ενώ η χρονική του εξέλιξη θα εξαρτάται από την τρέχουσα κατάσταση και τις ιδιότητες του συστήματος. Οι ιδιότητες του δικτύου από τις οποίες εξαρτάται η εξέλιξή του είναι τα βάρη των συνδέσεων και η συνάρτηση απαρίθωσης των κόμβων. Ας υποθέσουμε, για το παράδειγμα αυτό,

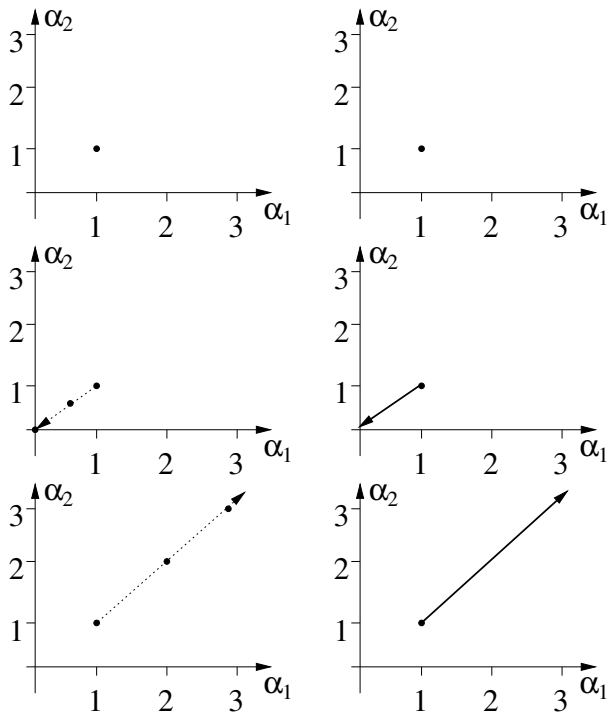
πως η συνάρτηση απαρτίωσης είναι ταυτότητα, δηλαδή η εκροή κάθε κόμβου ισούται με τη συνολική εισροή του. Ας υποθέσουμε επίσης ότι θέτουμε αρχική διέγερση ίση με τη μονάδα και στους δύο κόμβους. Το Σχήμα 7 δείχνει τη χρονική εξέλιξη του δικτύου για τρεις περιπτώσεις: όταν και τα δύο βάρη των συνδέσεων είναι 1, όταν είναι 0.5 και όταν είναι 2. Στην αριστερή στήλη η ανανέωση των ενεργοποιήσεων γίνεται σε διακριτά βήματα ή «καρέ» (frames) ενώ στη δεξιά στήλη ο χρόνος είναι συνεχής.

Η συνέχεια στο χρόνο δεν είναι απαραίτητη προϋπόθεση για ένα δυναμικό σύστημα. Ο χρόνος μπορεί να είναι διακριτός, ανανεώνοντας την κατάσταση του συστήματος σε μια αλληλουχία βημάτων. Στην περίπτωση αυτή η «τροχιά» είναι η ακολουθία των μεμονωμένων σημείων, γι' αυτό και στο Σχήμα 7, αριστερά, τα σημεία ενώνονται μόνο ενδεικτικά με διάστικτη γραμμή. Εκείνο που έχει σημασία είναι ότι ο χώρος των φάσεων είναι συνεχής και ότι το σύστημα διαθέτει την απαιτούμενη δυναμική ώστε να εξελίσσεται αυτόνομα στο χρόνο. Στην περίπτωση του συγκεκριμένου δικτύου οι προϋποθέσεις αυτές πληρούνται διότι οι κόμβοι αλληλοτροφοδοτούνται και οι ενεργοποιήσεις τους μπορούν να πάρουν οποιαδήποτε τιμή μέσα σε κάποιο διάστημα ελάχιστης ως μέγιστης ενεργοποίησης (είναι πραγματικοί αριθμοί). Σε ένα σύστημα διακριτού χρόνου η νέα κατάσταση που προκύπτει σε κάθε επόμενο βήμα μπορεί να υπολογιστεί με μια εξίσωση διαφοράς, απευθείας από την τρέχουσα κατάσταση, όπως για παράδειγμα στην ανανέωση των ενεργοποιήσεων μιας στιβάδας κόμβων σε ένα συνδεδειστικό δίκτυο που παρουσιάστηκε στο προηγούμενο κεφάλαιο (βλ. Σχήμα 1, σελ. 26).

Όταν ένα σύστημα εξελίσσεται σε συνεχή χρόνο τότε οι εξισώσεις υπολογισμού της κατάστασης δεν μπορούν να δώσουν απευθείας μια νέα τιμή ενεργοποίησης διότι η ενεργοποίηση μεταβάλλεται συνεχώς. Αντί λοιπόν για εξισώσεις διαφοράς στην περίπτωση αυτή έχουμε τις λεγόμενες *διαφορικές εξισώσεις*, στις οποίες υπολογίζεται ο ρυθμός μεταβολής της ενεργοποίησης και όχι απευθείας η νέα τιμή της. Αν γνωρίζουμε την τρέχουσα τιμή και το ρυθμό αύξησης (ή μείωσής) της, τότε μπορούμε να εκτιμήσουμε πόση θα είναι η τιμή της ενεργοποίησης μετά από κάποιο απειροελάχιστο χρονικό διάστημα. Οι διαφορικές εξισώσεις εκφράζουν τη φύση του συστήματος καθώς αυτές είναι που προσδιορίζουν τη μεταβολή της κατάστασής του κάθε στιγμή μέσα στη ροή του συνεχούς χρόνου, άρα και μακροπρόθεσμα την τροχιά εξέλιξης στο χώρο των φάσεων.



Σχήμα 6: Αναδρομικό συνδεδειστικό δίκτυο δύο κόμβων. *Αριστερά*, η γενική διάταξη με έναν κόμβο εισροής και έναν κόμβο εκροής. Με α συμβολίζεται η ενεργοποίηση των κόμβων και με w τα βάρη των συνδέσεων. *Δεξιά*, τρεις ειδικές περιπτώσεις για το ίδιο δίκτυο, με βάρη συνδέσεων ίσα με 1.0, 0.5 και 2.0 (από πάνω προς τα κάτω).



Σχήμα 7: Χώρος φάσεων του αναδρομικού συνδεδετιστικού δικτύου δύο κόμβων που φαίνεται στο Σχήμα 6. Οι τρεις σειρές αντιστοιχούν στις τρεις εκδοχές του συστήματος στη δεξιά στήλη του Σχήματος 6 (πάνω για βάρη σύνδεσης ίσα με 1, στη μέση 0.5 και κάτω 2). Σε κάθε περίπτωση η αρχική κατάσταση του συστήματος τίθεται $\alpha_1 = \alpha_2 = 1$. Στην αριστερή στήλη η κατάσταση του συστήματος ανανεώνεται σε διακριτά βήματα, υπολογίζοντας τις νέες ενεργοποιήσεις των κόμβων διατρέχοντας πλήρως το σύστημα από την εισροή ως την εκροή διαμέσου των συνδέσεων και της συνάρτησης απαρίτωσης. Στη δεξιά στήλη η κατάσταση του συστήματος εξελίσσεται σε συνεχή χρόνο, καθώς η ενεργοποίηση «ρέει» διαρκώς διαμέσου των συνδέσεων και η ενεργοποίηση κάθε κόμβου υπολογίζεται βάσει διαφορικής εξίσωσης που συνδέει την τρέχουσα εισροή όχι με την ενεργοποίηση αλλά με το ρυθμό μεταβολής της ενεργοποίησης.

4.3 Χαρακτηριστικά της δυναμικής προσέγγισης

Η δυναμική προσέγγιση στη μελέτη του νου ξεκινά με την υπόθεση ότι ο νους μαζί με το περιβάλλον του είναι ένα φυσικό δυναμικό σύστημα, δηλαδή ένα κλειστό σύστημα ποσοτήτων και σχέσεων μεταξύ τους το οποίο μπορεί να μεταβάλλεται στο χρόνο κι έτσι για κάθε χρονική στιγμή χαρακτηρίζεται από την κατάστασή του. Το αντικείμενο της δυναμικής υπόθεσης δεν είναι ο εγκέφαλος αλλά ο νους και συνεπώς όχι οι νευροφυσιολογικές αλλά οι νοητικές διεργασίες. Η διευκρίνιση αυτή είναι απαραίτητη διότι ο εγκέφαλος (το νευρικό σύστημα γενικότερα) και ο χώρος που τον περιβάλλει είναι προφανώς ένα φυσικό σύστημα και συνεπώς η παρατήρηση ότι πρόκειται για ένα δυναμικό σύστημα είναι τετριμμένη. Προφανώς οι φυσικές (χημικές, ηλεκτρικές κλπ.) καταστάσεις των στοιχείων του νευρικού συστήματος, σε αλληλεπίδραση με τις περιβαλλοντικές παραμέτρους, μεταβάλλονται ως φυσικά μεγέθη συνεχώς συναρτήσει του χρόνου· για κάθε χρονική στιγμή μπορεί να προσδιοριστεί (έστω θεωρητικά) μια αριθμητική τιμή για κάθε ιδιότητα και παράμετρο· οπότε το σύνολο των ποσοτήτων αυτών απαρίτζει κάθε στιγμή την τρέχουσα κατάσταση του συστήματος. Αυτό δεν είναι η δυναμική προσέγγιση στη μελέτη του νου.

Εστιάζοντας στο επίπεδο της νοητικής περιγραφής, όπως είναι αυτονόητο για τη γνωσιακή επιστήμη, η δυναμική προσέγγιση θεωρεί ότι ο νους είναι ένα σύστημα που δια-

γράφει τροχιές σε ένα πολυδιάστατο χώρο φάσεων. Οι τροχιές αυτές προσδιορίζονται από τη σύσταση του συστήματος, δηλαδή από τις εσωτερικές του παραμέτρους. Το πρόβλημα εδώ είναι να προσδιοριστούν οι διαστάσεις του συστήματος, δηλαδή οι άξονες του χώρου των φάσεων, και οι διαφορικές εξισώσεις που καθορίζουν τη χρονική εξέλιξη του συστήματος, δηλαδή τις νοητικές τροχιές. Στην πράξη αυτό είναι πολύ δύσκολο, οπότε συχνά προσεγγίζουμε το πρόβλημα μέσω του συνδετισμού, όπως στο παράδειγμα της προηγούμενης ενότητας, συνδυάζοντας τις δύο προσεγγίσεις. Έτσι, αξιοποιούμε γνωστές τεχνικές του συνδετισμού και χρησιμοποιούμε αναδρομικά δίκτυα κόμβων. Οι διαστάσεις του συστήματος ταυτίζονται με τις ενεργοποιήσεις των κόμβων, ενώ οι συναρτήσεις εξέλιξης του συστήματος ταυτίζονται με τις συναρτήσεις απαρτίωσης. Οι «εισορές» του συστήματος ορίζουν ένα αρχικό σημείο στο χώρο των φάσεων, ενώ στη συνέχεια το σύστημα εξελίσσεται αυτόνομα, είτε σε διακριτό είτε σε συνεχή χρόνο.

Λέγοντας ότι το δυναμικό σύστημα του νου είναι *κλειστό* σύστημα εννοούμε ότι ένα σύνολο ποσοτήτων αλληλεπιδρούν μεταξύ τους χωρίς να επηρεάζονται από μεταβολές σε άλλες ποσότητες εκτός του συνόλου. Η ικανοποίηση της συνθήκης του κλειστού συστήματος μπορεί να γίνει με δύο τρόπους: Ο ένας είναι να θεωρήσουμε ότι όλοι οι περιβαλλοντικοί παράγοντες που αλληλεπιδρούν με το νου είναι μέσα στο σύστημα, δηλαδή αποτελούν μέρος του δυναμικού συστήματος, του νου. Η προσέγγιση αυτή είναι χρήσιμη στο βαθμό που μας θυμίζει ότι ο νους δεν είναι μια ανεξάρτητη συσκευή, αποκομμένη από το περιβάλλον και τις συνθήκες λειτουργίας της, κάτι που θα δούμε εκτενέστερα στο επόμενο κεφάλαιο. Ο δεύτερος τρόπος είναι να θεωρήσουμε ότι ορισμένοι περιβαλλοντικοί παράγοντες παραμένουν σταθεροί ή ότι τουλάχιστον δεν επηρεάζονται από τη λειτουργία του νου παρότι μπορούν να την επηρεάζουν. Έτσι αυτοί οι περιβαλλοντικοί παράγοντες μπορούν να θεωρηθούν *παράμετροι* για το δυναμικό σύστημα του νου, εφόσον η τιμή τους δεν επηρεάζεται από τη λειτουργία του συστήματος. Στην περίπτωση αυτή το δυναμικό σύστημα είναι παραμετρικό αλλά όσον αφορά στην αλληλεπίδραση παραμένει κλειστό και ισχύουν οι ίδιοι κανόνες ανάλυσης και (μαθηματικής) περιγραφής.

Τελικά ένα δυναμικό σύστημα αποτελείται από το σύνολο των ποσοτήτων που αλληλεπιδρούν και από τις σχέσεις αλληλεπίδρασης μεταξύ τους, οι οποίες καθορίζουν και τη μεταβολή της κατάστασης του συστήματος συναρτήσει του χρόνου. Οι ποσότητες αυτές θα πρέπει, αν είναι αναπαραστασιακές, να αντιστοιχούν σε νοητικές γενικεύσεις που μας ενδιαφέρουν. Η αντιστοιχία αυτή μπορεί να είναι ένα-προς-ένα, δηλαδή κάθε συνεχής ποσότητα του δυναμικού συστήματος να αντιστοιχεί σε μια νοητική οντότητα. Η αντιστοιχία μπορεί να είναι και σε μορφή κατανεμημένης αναπαράστασης, δηλαδή περισσότερες από μία ποσότητες του δυναμικού συστήματος να «κωδικοποιούν» μία ή περισσότερες νοητικά ουσιώδεις γενικεύσεις. Αν οι ποσότητες που συναπαρτίζουν το δυναμικό σύστημα δεν είναι αναπαραστασιακές τότε η ερμηνεία των καταστάσεων του συστήματος είναι έμμεση (και δυσκολότερη). Τότε η έμφαση δίνεται στην ακρίβεια της περιγραφής των σχέσεων των ποσοτήτων του συστήματος με εξωτερικές παραμέτρους και στις χρονικές μεταξύ τους εξαρτήσεις. Η δυναμική προσέγγιση θεωρείται μη αναπαραστασιακή και το χαρακτηριστικό

αυτό θεωρείται σημαντικό προτέρημα από τους υπέρμαχους των δυναμικών μοντέλων.

Τι σημαίνει ότι μια προσέγγιση είναι «μη αναπαραστασιακή»; Σίγουρα δεν σημαίνει ότι οι καταστάσεις του συστήματος δεν αντιστοιχούν σε έννοιες ή ιδιότητες του περιβάλλοντος, διότι τότε δεν θα ήταν δυνατή η προσαρμοστική λειτουργία του νου σε αρμονία με το περιβάλλον. Για να το κατανοήσουμε καλύτερα πρέπει να θυμηθούμε ότι η έννοια της αναπαράστασης εντάχθηκε στη μελέτη του νου ως φορέας αιτιότητας, για να εξηγήσουμε τις διεργασίες με τις οποίες ο νους καταλήγει στη συστηματική συμπεριφορά που παρατηρούμε. Υποθέσαμε ότι για να δικαιολογηθεί μια παρατηρήσιμη συμπεριφορά πρέπει να υπάρχει κάποια οντότητα στο νου που να τη δικαιολογεί και να αποτελεί αιτία της, καθώς πλήθος συμπεριφορών δεν μπορούν να εξηγηθούν με αναφορά σε υπάρχοντα ερεθίσματα και τρέχουσες καταστάσεις. Άρα το κύριο χαρακτηριστικό των αναπαραστασιακών προσεγγίσεων δεν είναι η αντιστοιχία νοητικών και εξωτερικών καταστάσεων αλλά η αιτιακή ισχύς των νοητικών οντοτήτων. Αν σε κάποια προσέγγιση δεν υπάρχουν διακριτές νοητικές οντότητες στις οποίες να μπορεί να αποδοθεί αιτιότητα μέσα στο σύστημα τότε η προσέγγιση αυτή δεν είναι ουσιαστικά αναπαραστασιακή.

Στη δυναμική προσέγγιση η χρονική εξέλιξη του συστήματος προσδιορίζεται από το χώρο των φάσεων και τις προκαθορισμένες τροχιές μέσα σε αυτόν. Δεν υπάρχουν διακριτές οντότητες που να οδηγούν σε διακριτές καταστάσεις ή να αποτελούν σαφή αιτία διακριτών συμπεριφορών. Η τροχιά του συστήματος διέρχεται από περιοχές του χώρου των φάσεων εξαιτίας της δυναμικής του συστήματος, όπως ένα εκκρεμές που αφήνεται να ταλαντωθεί ελεύθερα και προσεγγίζει περιοδικά διαφορετικές περιοχές του δικού του χώρου φάσεων. Ορισμένες περιοχές του χώρου φάσεων μπορεί να θεωρηθεί, είτε εξαιτίας σχεδίασης είτε για διευκόλυνση της ερμηνείας, ότι αντιστοιχούν σε συγκεκριμένες έννοιες ή εξωτερικές καταστάσεις. Οπότε υπό μία έννοια η διέλευση μιας τροχιάς μέσα από αυτές τις περιοχές μπορεί να θεωρηθεί ένα είδος δυναμικής «αναπαράστασης», χωρίς όμως αιτιακή ισχύ. Οι ελκυστές του χώρου των φάσεων μπορούν ενίοτε να ερμηνευθούν αναπαραστασιακά, αρκεί να είναι σαφές ότι δεν αποτελούν αιτίες της εξέλιξης του συστήματος αλλά σταθερές ιδιότητες του χώρου των φάσεων και ως εκ τούτου έχουν περιγραφική και όχι αιτιακή αναπαραστατικότητα.

4.4 Σύγκριση συμβολικών και δυναμικών συστημάτων

Τι σημαίνει να είναι ο νους δυναμικό σύστημα; Με δεδομένο ότι ο εγκέφαλος (ή καλύτερα, το νευρικό σύστημα και το περιβάλλον του) είναι σίγουρα ένα φυσικό δυναμικό σύστημα, και ότι πολλές νοητικές λειτουργίες περιγράφονται συμβολικά με τρόπο που διατηρεί ορισμένα σημαντικά και ενδιαφέροντα χαρακτηριστικά τους (παραγωγικότητα, συνδυαστικότητα), αξίζει να διακρίνουμε την ουσιαστική διαφορά μεταξύ των θεωρήσεων. Συνοπτικά, κάθε θέση δέχεται τη δικιά της περιγραφή ως θεμελιώδη και την άλλη ως παρασιτική, σε σχέση με το επίπεδο γενίκευσης και περιγραφής που θεωρείται σημαντικό.

Μέσο υλοποίησης: Ένας ηλεκτρονικός υπολογιστής, όπως τον εννοεί ένας προγραμματιστής ή ένας απλός χρήστης, είναι ένα συμβολικό σύστημα που λειτουργεί με αλγόριθμους και περιγράφεται με λογικές πράξεις. Το σύστημα αυτό βρίσκεται σε μια καλά καθορισμένη κατάσταση για κάθε παλμό του «ρολογιού» του. Η σχέση των παλμών αυτών με τον πραγματικό φυσικό χρόνο είναι η λεγόμενη συχνότητα χρονισμού του επεξεργαστή. Όσον αφορά στη λειτουργία του επεξεργαστή και των συστημάτων αποθήκευσης, δε νοείται άλλη χρονική παράμετρος έξω από τους παλμούς χρονισμού και δεν νοείται έννοια χρόνου «μεταξύ» των παλμών ή «μεταξύ» των διαδοχικών καταστάσεων που αυτοί καθορίζουν.

Το σύστημα αυτό υλοποιείται από ηλεκτρονικά στοιχεία που λειτουργούν σύμφωνα με τους νόμους της φυσικής (για τους ημιαγωγούς, τις κινήσεις των ηλεκτρονίων κλπ.). Προφανώς, κάθε στιγμή το φυσικό σύστημα των ηλεκτρονικών στοιχείων που απαρτίζουν τον ηλεκτρονικό υπολογιστή βρίσκεται σε μια καθορισμένη και μετρήσιμη κατάσταση. Άρα για οποιαδήποτε χρονική στιγμή θα μπορούσε κάποιος να καταγράψει ποσοτικά την κατάσταση όλων των ηλεκτρονικών στοιχείων του υπολογιστή και να ορίσει έτσι τη συνολική κατάσταση του συστήματος. Το φυσικό σύστημα των ηλεκτρονικών στοιχείων είναι ένα σύστημα που εξελίσσεται στο χρόνο, δηλαδή ένα δυναμικό σύστημα, το οποίο αλληλεπιδρά με παραμέτρους του περιβάλλοντος (τμήματα εισροής-εκροής του υπολογιστή).

Χάρη στην έξυπνη σχεδίαση των ηλεκτρονικών στοιχείων και την ύπαρξη του χρονοδότη, ένα φυσικό δυναμικό σύστημα μπορεί να υλοποιεί ένα συμβολικό υπολογιστικό σύστημα – μάλιστα αν δεν υπήρχαν περιορισμοί μνήμης θα επρόκειτο για μια τέλεια μηχανή Turing ικανή για τον υπολογισμό οποιασδήποτε υπολογίσιμης συνάρτησης. Η σχέση αυτή μεταξύ φυσικού-δυναμικού και συμβολικού-υπολογιστικού συστήματος μπορεί να ονομαστεί σχέση υλοποίησης και αντιστοιχεί στη θέση της κλασικής γνωσιακής επιστήμης. Κανείς δεν αρνείται ότι το νευρικό σύστημα είναι ένα φυσικό δυναμικό σύστημα (μαζί με το περιβάλλον του), όμως ο συμβολισμός υποστηρίζει ότι οι νοητικά ουσιώδεις λειτουργίες του περιγράφονται καλύτερα σε ένα (συμβολικό) αλγοριθμικό επίπεδο «πάνω» από το επίπεδο της (δυναμικής) φυσικής περιγραφής.

Η θεώρηση του «μέσου υλοποίησης» είναι μια θέση του καθαρού συμβολισμού, σύμφωνα με την οποία η δυναμική προσέγγιση περιγράφει κάτι που βρίσκεται πιο κάτω από το νοητικό επίπεδο. Υπό αυτό το πρίσμα γίνεται αποδεκτή η δυναμική φύση του υποστρώματος αλλά απορρίπτεται η δυναμική προσέγγιση στην εξήγηση των λειτουργιών σε νοητικό επίπεδο. Λέει ο συμβολιστής: καλή η δυναμική προσέγγιση αλλά αυτό που περιγράφει δεν είναι αυτό που μας ενδιαφέρει.

Διαφορά σχεδίασης: Στο ηλιακό σύστημα οι πλανήτες και ο ήλιος αλληλεπιδρούν με βαρυτικές δυνάμεις. Για τον υπολογισμό των θέσεων και των ταχυτήτων των πλανητών οι φυσικοί χρησιμοποιούν την έννοια του κέντρου μάζας, χάρη στην οποία ένας ολόκληρος πλανήτης συμβολίζεται με ένα σημείο, και οι αποστάσεις μεταξύ πλανητών αντικαθίστανται από τις αποστάσεις μεταξύ των σημείων αυτών, δηλαδή των κέντρων μάζας τους. Έτσι αφενός απλοποιούνται πολύ οι υπολογισμοί και αφετέρου είναι πιο οικεία η ερμηνεία τους (ένας πλανήτης-ένα σημείο). Σε ένα τραπέζι μπιλιάρδου όμως, αν αντικαταστήσουμε

τις μπάλες από τα κέντρα μάζας τους τότε θα χάσουμε σχεδόν όλες τις ιδιότητες που έχουν σημασία για το παιχνίδι (συγκρούσεις, σπόντες και φάλτσα). Η ίδια ακριβώς απλοποίηση είναι στο σύστημα αυτό αναποτελεσματική, διότι οι υπολογισμοί είναι ανακριβείς και η ερμηνεία τους ελλιπής. Στην πρώτη περίπτωση η απλοποίηση δικαιολογείται από τις ιδιότητες του συστήματος (σχέση μεταξύ διαστάσεων και αποστάσεων πλανητών) ενώ στη δεύτερη όχι. Το πόσες και ποιες από τις φυσικές ιδιότητες του συστήματος στόχου πρέπει να συμπεριλάβουμε στο μοντέλο μας, λοιπόν, προκειμένου να κατανοήσουμε τη λειτουργία του και να κάνουμε ακριβείς προβλέψεις, εξαρτάται από τη φύση του συστήματος.

Στην περίπτωση ενός ηλεκτρονικού υπολογιστή, η αντικατάσταση των συνεχών καταστάσεων του φυσικού δυναμικού συστήματος από διακριτές συμβολικές καταστάσεις δικαιολογείται από την ύπαρξη του χρονισμού και από τις οριακές συνθήκες λειτουργίας των ηλεκτρονικών στοιχείων, τα οποία είναι σχεδιασμένα για να προσεγγίζουν γρήγορα ένα άνω ή κάτω όριο (δυναμική λειτουργία). Άρα για τον ηλεκτρονικό υπολογιστή είναι ενδεδειγμένο να απομακρυνόμαστε από την περίπλοκη δυναμική περιγραφή και να εστιάζουμε σε μια αφηρημένη συμβολική περιγραφή, καθώς με την τελευταία πετυχαίνουμε μεγαλύτερη απλότητα και κατανόηση. Στην περίπτωση του εγκεφάλου δεν είναι προφανές ότι υπάρχει τέτοια αντιστοιχία, διότι τα νευρικά κύτταρα δε φαίνεται να συντονίζονται από κάποιο χρονοδότη ενώ η δραστηριότητά τους δεν είναι δυαδική, αν υποθέσουμε ότι ο ρυθμός παραγωγής δυναμικών ενεργείας είναι η σημαντική παράμετρος που εκφράζει το βαθμό ενεργοποίησής τους και τη λειτουργική συνεισφορά τους στο σύστημα. Άρα, σύμφωνα με τους υποστηρικτές της δυναμικής προσέγγισης, δεν ευσταθεί το επιχείρημα των συμβολιστών περί υλοποίησης, εκτός κι αν αποδειχθεί ειδικά για το νευρικό σύστημα ότι το υλοποιούμενο νοητικό σύστημα είναι πράγματι συμβολικό.

Με άλλα λόγια, ο συμβολισμός τονίζει την αφαιρετικότητα της νοητικής περιγραφής, δίνοντας έμφαση στην απομάκρυνση από το φυσικό σύστημα του εγκεφάλου και του τρόπου λειτουργίας του, ελπίζοντας ότι αυτό οδηγεί σε επεξηγηματική απλότητα και ισχύ. Αντίθετα, ο δυναμισμός θεωρεί ότι η αφαιρετικότητα αφαιρεί στοιχεία ζωτικής σημασίας για την κατανόηση της λειτουργίας του νου και επιχειρεί νοητικές περιγραφές που να διατηρούν τη δυναμική φύση του συστήματος, όπως για παράδειγμα τη σταδιακή προσέγγιση περιοχών καταστάσεων, την παλινδρόμηση, και τη διέλευση σε κοντινές ή μακρινές αποστάσεις από διαφορετικές λεκάνες έλξης.

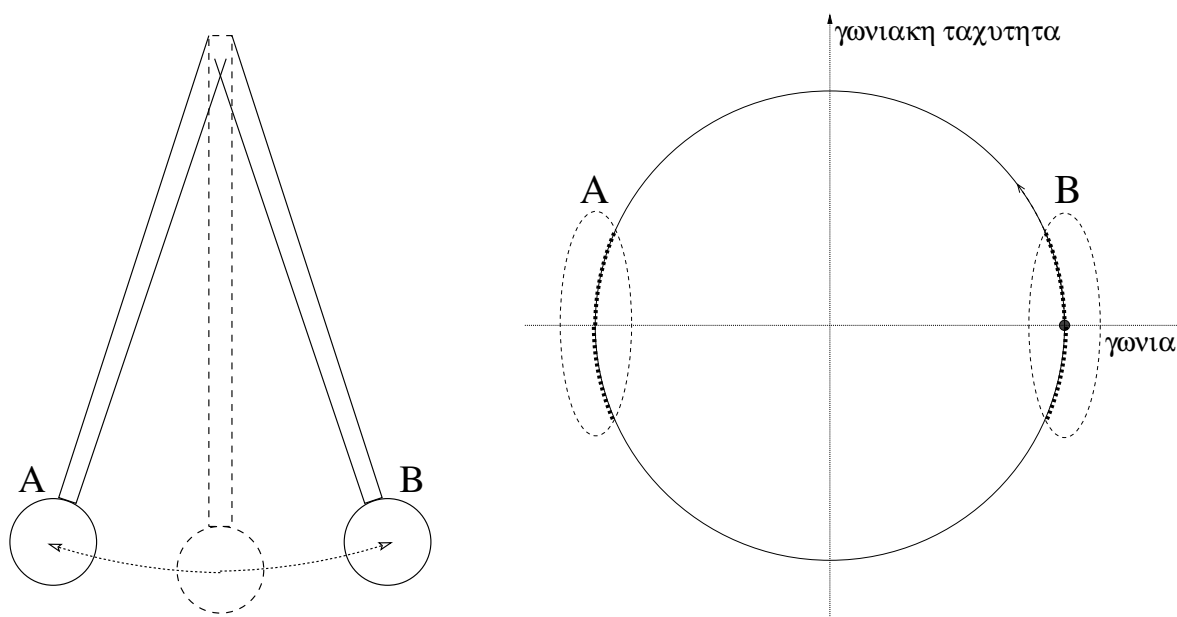
Η θεώρηση της «διαφοράς σχεδίασης» εκφράζει μια ενδιάμεση θέση, σύμφωνα με την οποία ο συμβολισμός και ο δυναμισμός έχουν διαφορετικά πεδία εφαρμογής στην περιγραφή νοητικών λειτουργιών. Κάθε προσέγγιση έχει ιδιαίτερα πλεονεκτήματα και μειονεκτήματα και επιλέγουμε εκείνη που προσφέρεται περισσότερο για το φαινόμενο που εστιάζουμε και μελετάμε.

Πράξη παρατήρησης: Όταν συνομιλούν δύο άνθρωποι υποθέτουμε ότι κάποιες ιδέες από το νου του ενός μετασχηματίζονται σε ηχητικά κύματα μέσω κάποιου πολύπλοκου συστήματος μετατροπής, και τα ηχητικά κύματα μετασχηματίζονται πάλι σε ιδέες στο νου του δεύτερου χάρη σε κάποιο αντίστοιχο, σχεδόν αντίστροφο, σύστημα μετατροπής. Η δια-

τύπωση αυτή είναι χαλαρά αναπαραστασιακή, ενώ το υποτιθέμενο σύστημα μετατροπής μπορεί να περιλαμβάνει νοητικά αλλά και φυσικά στάδια, π.χ. ιδέες → λέξεις → φθόγγοι → σχέδια άρθρωσης → κίνηση → ήχος ομιλίας κ.ο.κ. Η κατανόηση αυτού του φαινομένου διευκολύνεται αν κατηγοριοποιήσουμε καταστάσεις του φυσικού κόσμου (ηχητικά κύματα σε ακουστικές πραγματώσεις των φθόγγων) και του νοητικού κόσμου (αιτιακές νοητικές καταστάσεις σε λέξεις με φωνολογική συγκρότηση-σύνθεση και σημασιακό περιεχόμενο). Οι φυσικές ιδιότητες, τόσο του ήχου όσο και του νου, συνεχείς στο χώρο και στο χρόνο, αντιστοιχίζονται σε διακριτά σύμβολα που εκφράζονται σε διαδοχικές αλλά άχρονες καταστάσεις, με αποτέλεσμα την απλούστευση της περιγραφής. Στην περίπτωση αυτή η αντιστοιχία θεωρείται επιτυχημένη εφόσον διατηρεί τη συστηματικότητα στις συμπεριφορές, τόσο στις λέξεις, που μπορεί να εκφράσουν ιδέες, όσο και στους φθόγγους, που απαρτίζουν όλες τις δυνατές λέξεις.

Υπό την έννοια αυτή ο συμβολισμός αποτελεί μια πράξη παρατήρησης και αφαίρεσης, κατά την οποία οι συνεχείς φυσικές ιδιότητες του υπό παρατήρηση συστήματος ομαδοποιούνται σε διακριτά και ασυνεχή σύμβολα, ενώ η συνεχής εξέλιξη του συστήματος στο χρόνο μεταφράζεται σε αλληλουχία διακριτών καταστάσεων. Για να περιγράψουμε με ακρίβεια σε ένα συμβολικό σύστημα τη λειτουργία του φυσικού δυναμικού συστήματος πρέπει τα σύμβολα και οι σχέσεις μεταξύ τους (κανόνες μετάβασης του συμβολικού συστήματος) να ενσωματώνουν αφενός τις ιδιότητες του φυσικού συστήματος και αφετέρου τις συναρτήσεις αντιστοιχίας, δηλαδή τις επιπτώσεις της μεταφοράς από το συνεχές φυσικό στο διακριτό συμβολικό πεδίο. Όταν η κανονικότητα της λειτουργίας του φυσικού συστήματος το επιτρέπει, μια τέτοια μεταφορά μπορεί να απλοποιεί την περιγραφή των λειτουργιών του. Όμως η μεταφορά δεν παύει να είναι μια παραμόρφωση της εστίας και της φύσης των αιτιακών σχέσεων μεταξύ των καταστάσεων του συστήματος, ο δε αναπαραστασιακός χαρακτήρας της συμβολικής περιγραφής μπορεί να είναι παραπλανητικός.

Για παράδειγμα, στο Σχήμα 8 φαίνεται το εκκρεμές (αριστερά) και ο αντίστοιχος χώρος των φάσεων με την τροχιά της κατάστασης του συστήματος (δεξιά). Φανταστείτε ότι κοιτάζουμε το εκκρεμές που ταλαντώνεται, μέσα από μια κακής ποιότητας κάμερα εξαιτίας της οποίας ό,τι κινείται φαίνεται θαμπό ή δεν φαίνεται καθόλου. Τότε θα βλέπουμε καθαρά το εκκρεμές μόνο όταν αυτό μειώνει πολύ την ταχύτητά του και ακινητοποιείται στιγμιαία στα ακραία σημεία της ταλάντωσης. Αν παρατηρούμε και αναφέρουμε την ορατή θέση του εκκρεμούς, το αποτέλεσμα θα είναι $A \dots B \dots A \dots$ κ.ο.κ. Δηλαδή θα καταγράψουμε μια αλληλουχία διακριτών καταστάσεων, ως συμβολοσειρά $ABAB \dots$, την οποία θα μπορούμε να «εξηγήσουμε» με τους συμβολικούς κανόνες $K1:A \rightarrow B$ και $K2:B \rightarrow A$. Η συμβολοσειρά προκύπτει από την παρατήρηση ενός συνεχούς δυναμικού συστήματος, του οποίου η πραγματική εξέλιξη δίνεται από την εξίσωση κίνησης του εκκρεμούς. Στη συγκεκριμένη περίπτωση, η συμβολική προσέγγιση παράγει μια επαρκή περιγραφή δύο επιλεγμένων όψεων του συστήματος, απλούστερη από την υποκείμενη δυναμική περιγραφή. Χάνει όμως τις χρονικές σχέσεις, τις λεπτομέρειες της στιγμιαίας θέσης, καθώς και τις μεταβολές στην ταχύτητα κίνησης του εκκρεμούς, τα οποία μπορεί για κάποιες εφαρμογές να είναι



Σχήμα 8: Δυναμικό σύστημα και συμβολική παρατήρηση. *Αριστερά*, το εκκρεμές του Σχήματος 4. *Δεξιά*, ο αντίστοιχος χώρος φάσεων, όπως στο Σχήμα 5. Με τις ενδείξεις A και B σημειώνονται περιοχές του φυσικού χώρου (αριστερά) και, αντίστοιχα, του χώρου των φάσεων (δεξιά), όπου το εκκρεμές βρίσκεται σε θέσεις μεγάλης απόκλισης και μικρής ταχύτητας.

σημαντικά. Η «σχέση παρατήρησης» μεταξύ του φυσικού-δυναμικού συστήματος και της συμβολικής περιγραφής του είναι μια σχέση απλοποιημένης αντιστοιχίας, η οποία μπορεί σε κάποιες περιπτώσεις ή για κάποια συστήματα να είναι χρήσιμη.

Στη δυναμική θεώρηση του νου το συνεχές φυσικό σύστημα δεν είναι το νευρικό σύστημα αλλά το νοητικό σύστημα που υλοποιείται απ' αυτό. Η συστηματικότητα στη λειτουργία του νου είναι εγγενές μη συμβολικό χαρακτηριστικό του. Ένα απεριοδικό δυναμικό σύστημα μπορεί να προσεγγίζει τις ίδιες περιοχές στο χώρο φάσεων χωρίς να επαναλαμβάνει ακριβώς την ίδια τροχιά αλλά περνώντας *σχεδόν περιοδικά* από πολύ όμοιες (κοντινές) καταστάσεις. Η τοποθέτηση της «πράξης παρατήρησης» είναι η άλλη όψη του επιχειρήματος του «μέσου υλοποίησης», από τη μεριά των δυναμιστών: Η θεμελιώδης διαφορά μεταξύ των δύο είναι η πρωταρχικότητα της κάθε αναπαράστασης. Κατά την πράξη παρατήρησης, το νοητικό σύστημα είναι δυναμικό, ενώ η συμβολική περιγραφή του προέρχεται από την παρατήρησή του, είναι δηλαδή δευτερεύουσα και παραμορφωτική. Αντίθετα, στη θεώρηση του μέσου υλοποίησης το νοητικό σύστημα είναι εγγενώς συμβολικό, άρα η συμβολική περιγραφή του πρωταρχική, ενώ η δυναμική περιγραφή είναι άσχετη με τη νοητική λειτουργία και αφορά μόνο στο φυσικό μέσο υλοποίησης. Με άλλα λόγια, η «πράξη παρατήρησης» είναι η απάντηση του δυναμισμού στο «μέσο υλοποίησης». Αποδέχεται τις ίδιες διαφορές αλλά τις ερμηνεύει με αντίθετο τρόπο. Λέει ο δυναμιστής: καλά τα σύμβολα, αλλά υπεραπλουστεύουν και συσκοτίζουν τη λειτουργία του συστήματος. Η

κανονικότητά τους είναι παραπλανητική, επιφανειακή και όχι αιτιακή.

Έτσι, στη δυναμική προσέγγιση οι αναπαραστάσεις, αν υπάρχουν, δεν είναι φορείς αιτιότητας αλλά απλώς κάποιες βολικές καταστάσεις του συστήματος. Η συστηματικότητα του νου οφείλεται σε παραπλήσιες τροχιές του νοητικού συστήματος κατά τη διαρκή μεταβολή του ή σε ειδικές καταστάσεις (ελκυστές). Σύμφωνα με το δυναμισμό, η απλοϊκή ψυχολογία δεν αποτελεί κατάλληλο επεξηγηματικό πλαίσιο για τις νοητικές διεργασίες και θα πρέπει να εγκαταλειφθεί. Τα δυναμικά συστήματα μοιάζουν να συγγενεύουν με τη θέση του εξαλειπτικού υλισμού, κατά την οποία οι νοητικές έννοιες θα αντικατασταθούν όταν με την πρόοδο της επιστήμης κατανοήσουμε καλύτερα τις νοητικές λειτουργίες.

5 Ενσώματη νόηση

Η δυναμική προσέγγιση στο νου αποτελεί μια απομάκρυνση από την αφαιρετικότητα της συμβολικής προσέγγισης, προς την κατεύθυνση του φυσικού δυναμικού συστήματος που υλοποιεί το νου. Απορρίπτοντας τη διακριτότητα των καταστάσεων, και εστιάζοντας στη συνολική τροχιά που διαγράφει ο νους, ο δυναμισμός αποτελεί μια θεμελιώδη ρήξη με την παραδοσιακή παρομοίωση του νου με ψηφιακό ηλεκτρονικό υπολογιστή. Παρόλα αυτά, διατηρεί σε μεγάλο βαθμό την αφαιρετικότητα όσον αφορά στο πλαίσιο μέσα στο οποίο λειτουργεί ο νους, καθώς δέχεται τη συμμετοχή του περιβάλλοντος ως μέρος του συστήματος παραμέτρων και ποσοτήτων που αλληλεπιδρούν. Η πιο πρόσφατη εξέλιξη της δυναμικής προσέγγισης θέτει το νου στο φυσικό πλαίσιο λειτουργίας και μέσο έκφρασής του, δηλαδή το σώμα και τις φυσικές αλληλεπιδράσεις του σώματος με το περιβάλλον. Ξεκινά με την υπόθεση ότι ο νους υπάρχει με βασικό σκοπό τη συμπεριφορά σε επίπεδο φυσικού σώματος, και δεν εμφανίζεται στη φύση ως αφηρημένο και αποκομμένο συλλογιστικό σύστημα. Αυτή είναι η υπόθεση της *ενσώματης νόησης* (embodied cognition).

5.1 Νοητικές αναπαραστάσεις του σώματος

Το σώμα ως νοητικό εξάρτημα εκφράζεται στο νου μέσα από δύο διαφορετικές εκφάνσεις:¹² την *εικόνα του σώματος* (body image) και το *σχήμα του σώματος*¹³ (body schema).

Η *εικόνα* του σώματος είναι αντιληπτική γνώση της τρέχουσας θέσης και κατάστασης των μερών του σώματος. Περιλαμβάνει, για παράδειγμα, την πληροφορία ότι έχουμε δύο χέρια (όχι ένα ή τρία) και την πληροφορία για το πού βρίσκεται το καθένα και σε ποια στάση ή κίνηση. Η *εικόνα* του σώματος είναι, λοιπόν, η γνώση που έχουμε για την ύπαρξη και τη στάση του σώματος. Ο Gallagher ορίζει την *εικόνα* του σώματος ως ένα «σύστημα

¹²Διάκριση βασισμένη στο Gallagher, S. (2005). *How the body shapes the mind*. Oxford University Press.

¹³Δεν εννοούμε εδώ τη λέξη «σχήμα» με τη σημασία της μορφής, αλλά με τη σημασία του αντίστοιχου όρου (schema) της γνωστικής ψυχολογίας, ο οποίος αναφέρεται σε γνωσιακή δομή οργάνωσης πληροφοριών που σχετίζονται με κάποιο θέμα και χρησιμοποιούνται στην κατανόηση και επεξεργασία των πληροφοριών.

αντιλήψεων, στάσεων και πεποιθήσεων» για το σώμα (σελ. 24). Περιλαμβάνει και τη γνώση της ιδιοκτησίας του σώματος, που ξεχωρίζει το σώμα ως φυσικό εαυτό από το περιβάλλον.

Το *σχήμα* του σώματος είναι η άρρητη γνώση των αισθησιοκινητικών δυνατοτήτων που εμπεριέχονται στην τρέχουσα θέση και κατάσταση του σώματος. Η γνώση αυτή συμπεριλαμβάνει το τι είναι εφικτό καθώς και ποια είναι τα προβλεπόμενα αποτελέσματα οποιασδήποτε δράσης. Για παράδειγμα, περιλαμβάνεται η γνώση της εμβέλειας: κάποια αντικείμενα γύρω μου τα φτάνω και μπορώ να απλώσω το χέρι να τα πιάσω, ενώ άλλα όχι. Κατά κανόνα δεν χρειάζεται να δοκιμάσουμε να τεντώσουμε το χέρι μας για να δούμε αν κάτι είναι μέσα στην ακτίνα δράσης ή όχι. Επίσης, περιλαμβάνεται η γνώση της προσέγγισης: για κάθε αντικείμενο γύρω μου είναι προφανές ποια θέση, άνοιγμα και γωνία προσέγγισης του χεριού είναι κατάλληλη για να το πιάσει. Αλλιώς πιάνω ένα φύλλο χαρτί, μια μπάλα, μια γεμάτη βαλίτσα, μια κούπα κλπ. Η άρρητη γνώση για την οποία μιλάμε δεν περιορίζεται, φυσικά, στα χέρια. Περιλαμβάνει τη γνώση, καθώς κινούμαστε, για το πώς πρέπει να τοποθετηθεί το πόδι για να περάσει ανάμεσα από κάποια εμπόδια ή πώς πρέπει να γείρει το σώμα προκειμένου να μη χτυπήσει σε μια κολώνα, μια καρέκλα ή κάτι που κρέμεται. Όλη αυτή η γνώση είναι ανά πάσα στιγμή διαθέσιμη ως κατανόηση των δυνατοτήτων του σώματος, έτοιμη να αξιοποιηθεί όταν το απαιτήσουν οι περιστάσεις.

Η προσδοκία των επιδράσεων της δράσης είναι σημαντικό χαρακτηριστικό του σχήματος του σώματος, καθώς εμπλέκεται στη διαρκή αλληλεπίδραση με το περιβάλλον κατά την αισθητηριακή εξερεύνηση και τη μετακίνηση. Για παράδειγμα, η θέση των ματιών και του κεφαλιού κάθε στιγμή, σε συνδυασμό με τη θέση και τη στάση του σώματος, προσδιορίζει ποιες όψεις των ορατών αντικειμένων είναι δυνητικά προσβάσιμες και ποιες όχι. Η αντίληψη ενός στοιχείου ως αντικειμένου περιλαμβάνει την προσδοκία ότι αν κινήσω το κεφάλι μου (ή τον κορμό) προς κάποια κατεύθυνση τότε αυτό το στοιχείο (α) θα κινηθεί σε αντίθετη κατεύθυνση στο οπτικό μου πεδίο και (β) θα «αποκαλύψει» μέρος του, στην πλευρά προς την οποία θα κινηθώ, το οποίο ήταν κρυμμένο στην «πίσω πλευρά» του στην προηγούμενη θέση. Η γνώση αυτή επιτρέπει τον ενεργητικό χειρισμό των οπτικών αισθητηρίων για την αντίληψη των αντικειμένων αλλά και των δυνατοτήτων πλοήγησης που παρέχει το περιβάλλον. Κάθε κίνηση που προγραμματίζεται περιλαμβάνει μέσα στον προγραμματισμό της και την προσδοκία για το απτικό και οπτικό αποτέλεσμά της. Όταν η προσδοκία δεν επιβεβαιώνεται, η έκπληξη που νιώθουμε αποδεικνύει ότι υπήρχε προσδοκία.

Η άρρητη εμπλοκή του σχήματος του σώματος στη νόηση, μέσω της δράσης, αποκτά ιδιαίτερο ενδιαφέρον σε περιπτώσεις ιδιαίτερα εξασκημένης (και εξαρτημένης) χρήσης εξωτερικών εξαρτημάτων. Για παράδειγμα, το ραβδί με το οποίο ο τυφλός εξερευνά ενεργά το χώρο, αντίστοιχα με την οπτική εξερεύνηση των βλεπόντων, δεν είναι απλώς ένα οργανοβοήθημα, αλλά μια προέκταση του εαυτού η οποία λειτουργεί ως μέρος του σώματος και ειδικότερα του σχήματος του σώματος. Σε μικρότερο βαθμό, το σφυρί του μαραγκού, η στέκα του παίκτη του μπυλιάρδου και το πληκτρολόγιο του δακτυλογράφου γίνονται νοητικές προεκτάσεις καθώς απορροφώνται στο σχήμα του σώματος και εντάσσονται στο σύστημα άρρητου ελέγχου της δράσης και προσδοκιών των αποτελεσμάτων τους.

Η μελέτη της ανάπτυξης του σχήματος του σώματος ως επεξηγηματικής οντότητας σε νοητικά μοντέλα αποτελεί μέρος της ενσώματης προσέγγισης του νου. Σε μια πολυσυζητημένη σειρά προσομοιώσεων, η Thelen και οι συνεργάτες της¹⁴ επιχείρησαν να περιγράψουν την αναπτυξιακή πορεία του σφάλματος «Α-όχι-Β» με ένα δυναμικό μοντέλο το οποίο περιλαμβάνει δύο αλληλεπιδρώντα πεδία κίνησης, το ένα μνημονικό και το άλλο για τον προγραμματισμό. Το σφάλμα «Α-όχι-Β» είναι ένα φαινόμενο που παρατηρείται σε βρέφη ηλικίας 7-12 μηνών, το οποίο πρωτοπεριγράφηκε από τον Piaget. Για την παρατήρησή του χρησιμοποιούνται δύο δοχεία με καπάκι και ένα αντικείμενο, συνήθως παιχνίδι, επιθυμητό στο βρέφος. Ο εξεταστής τοποθετεί το αντικείμενο μέσα στο ένα δοχείο (το Α) και το σκεπάζει, και μετά από μια σύντομη καθυστέρηση επιτρέπει στο βρέφος να το ανακτήσει. Η διαδικασία αυτή επαναλαμβάνεται μερικές φορές. Στη συνέχεια ο εξεταστής τοποθετεί το αντικείμενο στο άλλο δοχείο (το Β). Όταν το βρέφος αφεθεί ελεύθερο να αναζητήσει το αντικείμενο, αντί να κινηθεί προς το Β, κινείται πάλι προς το Α. Το φαινόμενο αυτό παρουσιάζεται παρότι το βρέφος γνωρίζει ότι το αντικείμενο βρίσκεται στο Β, κάτι που μπορεί να διαπιστωθεί με αντιληπτικές μελέτες. Η πιθανότητα εμφάνισης του φαινομένου εξαρτάται από την ηλικία του βρέφους, από το πόσο επιθυμητό είναι το αντικείμενο (ή αν είναι φαγώσιμο), από το πόσο διαφορετικά είναι μεταξύ τους τα καπάκια των δοχείων και πόσο ευδιάκριτα είναι σε σχέση με το υπόβαθρο, ακόμα και από το αν το βρέφος διατηρείται στην ίδια θέση και στάση μεταξύ των επαναλήψεων.

Η ενσώματη προσέγγιση στο πρόβλημα «Α-όχι-Β» δεν θεώρησε ότι υπάρχει δυσκολία νοητικής αναπαράστασης ή αντίληψης του αντικειμένου, αλλά θεώρησε ότι το βασικό πεδίο εμφάνισης του φαινομένου ταυτίζεται με το πεδίο δράσης στο οποίο προγραμματίζεται η κίνηση. Η εμφάνιση του αντικειμένου στο οπτικό πεδίο διεγείρει ένα αντιληπτικό πεδίο, το οποίο με τη σειρά του διεγείρει το πεδίο κινητικού προγραμματισμού. Παράλληλα, ένα μνημονικό πεδίο διατηρεί (ως παραμένουσα ενεργοποίηση) ένα συγκεντρωτικό ίχνος των προηγούμενων κινήσεων του βρέφους. Το αντιληπτικό, το μνημονικό και το προγραμματιστικό πεδίο αλληλεπιδρούν με ισχύ που μεταβάλλεται αργά κατά τη διάρκεια της ανάπτυξης. Η διατήρηση της διεγερσης κάθε πεδίου επίσης μεταβάλλεται κατά τη διάρκεια της ανάπτυξης. Το αποτέλεσμα είναι ότι το πεδίο προγραμματισμού δεν καταφέρνει να διατηρήσει επαρκώς ενεργή την κατεύθυνση που δίνεται από το αντιληπτικό πεδίο για μεγάλο χρονικό διάστημα, οπότε μπορεί να προγραμματίσει λανθασμένη κίνηση προς την κατεύθυνση που υποδεικνύει το μνημονικό πεδίο. Η επίδραση αυτή εξαρτάται από τη σχετική ισχύ των διεγέρσεων των πεδίων καθώς και από τη διατήρηση της σχέσης μνημονικού και προγραμματιστικού πεδίου (η οποία εκφράζει τη σταθερή στάση του σώματος του βρέφους). Σε μεγαλύτερες ηλικίες η κινητική διεγερση που οφείλεται στο αντιληπτικό πεδίο μπορεί να ανατροφοδοτηθεί και να διατηρηθεί για μεγαλύτερο χρονικό διάστημα, υπερνικώντας την επίδραση του μνημονικού πεδίου, και άρα εξαλείφοντας το φαινόμενο.

Η εξήγηση αυτή δεν είναι αναπαραστασιακή με την κλασική έννοια κατά την οποία

¹⁴Thelen, E., et al. (2001). The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, 24, 1, 1-86.

χρησιμοποιούμε την υπόθεση της αναπαράστασης του αντικειμένου ως ενιαία και αιτιακή. Αντίθετα, χρησιμοποιούμε διάφορα αλληλεπιδρώντα πεδία στα οποία εδράζονται διαφορετικού χαρακτήρα «αναπαραστάσεις» του αντικειμένου, καθεμιά από τις οποίες σχετίζεται με διαφορετικό τρόπο με το υποκείμενο της νόησης: άλλη είναι άμεσα οφειλόμενη στο οπτικό πεδίο, άλλη στο κινητικό ιστορικό και άλλη στον επικείμενο κινητικό προγραμματισμό. Η εξήγηση είναι γνωσιακή στο βαθμό που είμαστε διατεθειμένοι να θεωρήσουμε τον προγραμματισμό και το αποτέλεσμα της κίνησης ως μέρος της διαδικασίας της νόησης.

5.2 Αρχές της ενσώματης νόησης

Η προσέγγιση της ενσώματης νόησης μπορεί να συνοψιστεί σε έξι βασικές θέσεις,¹⁵ οι οποίες εκφράζουν τις διαφορετικές εκφάνσεις της βασικής ιδέας ότι ο νους υπάρχει και λειτουργεί εξαιτίας, βάσει και διαμέσου του σώματος:

1. *Η νόηση είναι τοποθετημένη* (ή «εγκατεστημένη», situated) μέσα σε ένα πλαίσιο, ένα περιβάλλον, και δεν λειτουργεί αποκομμένα και αυτόνομα σε κάποιο αφηρημένο υπολογιστικό χώρο. Αυτή είναι ίσως η βασικότερη θέση της ενσώματης νόησης, και έχει να κάνει με τη διαρκή ανταλλαγή πληροφοριών μεταξύ νου και περιβάλλοντος διαμέσου της αντίληψης και της δράσης. Οι νοητικές εισροές και εκροές δεν είναι δυο άκρα μιας απομονωμένης διαδικασίας αλλά στοιχεία διαρκούς αλληλεπίδρασης, καθώς ο νους και το περιβάλλον αλληλοεπηρεάζονται συνεχώς (βλ. παρακάτω, στην ενότητα για την οικολογική ψυχολογία). Η θέση αυτή αντιτίθεται στην κλασική προσέγγιση της νόησης που μοιάζει περισσότερο με αυτόματο πωλητή: ρίχνεις το κέρμα (εισορή), κάτι γίνεται μέσα, και κάποτε πέφτει το αναψυκτικό (εκροή). Δεν αντιτίθεται στην ιδέα ότι υπάρχουν και κάποιες νοητικές διεργασίες που μπορούν να εκτελεστούν χωρίς άμεση αναφορά στην τρέχουσα κατάσταση του περιβάλλοντος, όπως είναι ο προσχεδιασμός, η αναπόληση ή η περισυλλογή. Στην προσέγγιση αυτή υποβαθμίζεται εσκεμμένα η σημασία των αναλυτικών και αφαιρετικών νοητικών διεργασιών, όπως λ.χ. η ενασχόληση με τα μαθηματικά, ενώ (υπερ)τονίζεται ο ρόλος των άμεσων διεργασιών πρόσληψης και δράσης στην κατανόηση της ουσίας και της λειτουργίας του νου. Από όλες τις απόψεις, η θέση αυτή διαχωρίζει με αγεφύρωτο χάσμα την ενσώματη από τη συμβολική προσέγγιση, τείνοντας κατά το μέγιστο δυνατό προς την αντίθετη κατεύθυνση στις βασικές παραδοχές και επιλογές μελέτης.
2. *Η νόηση υπόκειται σε χρονική πίεση*. Η θέση αυτή, σε συνδυασμό με την προηγούμενη, τονίζει τη σημασία της άμεσης και διαρκούς ανταπόκρισης στη συνεχή ροή πληροφοριών και την αλληλεπίδραση με το περιβάλλον. Η ανταλλαγή των πληροφοριών είναι συνεχής και η νοητική λειτουργία επηρεάζει το περιβάλλον διαμέσου της δράσης ενώ παράλληλα το περιβάλλον μπορεί να μεταβάλλεται, τροποποιώντας τις νοητικές εισροές χωρίς να περιμένει ανταπόκριση σε προηγούμενες καταστάσεις.

¹⁵Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9 (4), 625–636.

Η αντιμετώπιση του νου ως εγκατεστημένου σε ένα πραγματικό περιβάλλον που εξελίσσεται στο χρόνο θέτει περιορισμούς στη νοητική λειτουργία, καθώς ο νους θα πρέπει να ανταποκρίνεται άμεσα. Η απαίτηση άμεσης και διαρκούς ανταπόκρισης έρχεται σε αντίθεση με τις αλγοριθμικές προσεγγίσεις του συμβολισμού όπου δεν λαμβάνεται υπόψη ο αριθμός των βημάτων ή η πολυπλοκότητα της επεξεργασίας, καθώς το ζητούμενο είναι η βέλτιστη λύση και όχι απλώς μια επαρκής ανταπόκριση. Η έμφαση στη χρονική πίεση διαμορφώνει διαφορετικές απαιτήσεις από τη νοητική λειτουργία, καθώς η καθυστερημένη ανταπόκριση δεν είναι αποδεκτή, δίνοντας προτεραιότητα σε σχετικά ικανοποιητικές εκροές, οι οποίες μπορεί να απέχουν από το βέλτιστο δυνατό αλλά είναι εφικτό να παρασχεθούν μέσα σε εύλογο χρονικό διάστημα και με βάση τους διαθέσιμους πόρους (γνώση και επεξεργαστική ισχύ).

3. *Η νόηση αποφορτίζεται αξιοποιώντας το περιβάλλον* («ξεφορτώνοντας» νοητικό έργο). Αυτή είναι μια ιδιαίτερη θέση της ενσώματης προσέγγισης κατά την οποία το ενσώματο νοητικό σύστημα χρησιμοποιεί το περιβάλλον ώστε να αποφύγει την εσωτερική αναπαράσταση ή συγκράτηση και επεξεργασία πληροφοριών. Η θέση αυτή μπορεί να πάρει διαφορετικές μορφές ανάλογα με τη φύση του νοητικού έργου που εξετάζεται. Σε αυτήν εμπίπτει η χρήση σημειώσεων, σχεδίων, διαγραμμάτων, ακόμα και το μέτρημα με τα δάχτυλα, καθώς σε όλες αυτές τις περιπτώσεις αφηρημένης διεργασίας αξιοποιείται μια φυσική δομή έξω από το νου ως «σημειωματάριο», η οποία εξυπηρετεί τη συγκράτηση και την επεξεργασία πληροφοριών. Σε νοητικά έργα χωρικού τύπου αξιοποιούνται τα ίδια τα σημεία του φυσικού χώρου ως αναφορά για τη θέση τους, όταν για παράδειγμα κοιτάζουμε ένα σχήμα ή μια θέση του χώρου προκειμένου να επεξεργαστούμε μια σχέση ή τοποθέτηση. Στην επίλυση οπτικοχωρικών προβλημάτων, π.χ. παζλ, σχέδια με κύβους, ή προσαρμογή γεωμετρικών σχημάτων, οι οφθαλμοκινήσεις μαρτυρούν ότι δεν εσωτερικεύουμε κάποια αναπαράσταση μορφής την οποία επεξεργαζόμαστε εσωτερικά (π.χ. μετακινώντας ή περιστρέφοντας την) αλλά αντίθετα αναφερόμαστε συνεχώς στο ορατό δείγμα προκειμένου να αξιολογήσουμε πιθανές λύσεις. Η νοητική αυτή στρατηγική έχει μελετηθεί εκτενώς στο διαδομένο βιντεοπαιχνίδι tetris, όπου φαίνεται ότι οι καλοί παίκτες περιστρέφουν και παρατηρούν τα κομμάτια που πέφτουν κάνοντας πολλές δοκιμές με ενεργητικό οπτικό τρόπο πάνω στην οθόνη και όχι με εσωτερική νοερή περιστροφή.
4. *Το περιβάλλον αποτελεί μέρος του νοητικού συστήματος*. Η θέση αυτή συναντά την αντίστοιχη ιδέα της δυναμικής προσέγγισης, όπου ένα σύστημα περιλαμβάνει εντός του όλες τις μεταβλητές και τις παραμέτρους έτσι ώστε να είναι «κλειστό» και να μπορεί να μελετηθεί αυτόνομα. Αποτελεί συνδυασμό και προέκταση των παραπάνω τριών θέσεων, καθώς από αυτές προκύπτει ότι ο νους δεν είναι δυνατό να αποσυνδεθεί από το περιβάλλον προκειμένου να μελετηθεί ξεχωριστά. Αφενός η διαρκής αλληλεπίδραση και αφετέρου η ενεργή αξιοποίηση του περιβάλλοντος ως νοητική στρατηγική, υποδηλώνουν μια τόσο στενή σύζευξη που τελικά προτείνεται

να μελετώνται ο νους και το περιβάλλον του ως ένα ενιαίο σύστημα «εκτεταμένου νου» (extended mind, ή «εκτεταμένη νόηση», extended cognition).¹⁶ Η θέση αυτή βρίσκει την πιο ακραία διατύπωσή της στην έννοια της «κατανεμημένης νόησης» (distributed cognition) κατά την οποία οι νοητικές διεργασίες δεν εντοπίζονται πουθενά, όχι μόνο στη λειτουργία του εγκεφάλου μαζί με το υπόλοιπο σώμα και το περιβάλλον του, αλλά ούτε καν απαραίτητα σε ένα μεμονωμένο άτομο, καθώς οι κοινωνικές αλληλεπιδράσεις οδηγούν σε τόσο ισχυρή σύζευξη διεργασιών και αναπαραστάσεων που τελικά ένας νους μπορεί να συγκροτείται από πολλούς δράστες και να κατανέμεται στα μέλη μιας κοινωνικής ομάδας και το κοινό τους περιβάλλον.

5. *Η νόηση γίνεται με σκοπό τη δράση.* Η θέση αυτή τονίζει μια όψη της νόησης που είναι συστηματικά παραμελημένη από τη συμβολική (και σε μεγάλο βαθμό και από τη συνδετιστική) προσέγγιση. Ο νους υπάρχει ως λειτουργία των νοημόνων οργανισμών επειδή εξυπηρετεί τη δράση τους μέσα στο φυσικό τους περιβάλλον και όχι για να στοχάζονται και να λύνουν αφηρημένα προβλήματα. Παρότι σήμερα ορισμένοι οργανισμοί έχουν επιτύχει να αξιοποιούν νοητικές λειτουργίες αρκετά αποκομμένες από το άμεσο φυσικό περιβάλλον τους, ο βασικός σκοπός της νόησης και ο λόγος της ύπαρξής της, σύμφωνα με την ενσώματη προσέγγιση, είναι ο προγραμματισμός, ο έλεγχος και η εκτέλεση πράξεων μέσα σε ένα πλαίσιο-περιβάλλον. Δηλαδή, δεν βλέπουμε για να κατασκευάσουμε νοητικές αναπαραστάσεις αντικειμένων, γενικά και αφηρημένα, αλλά για να πλοηγηθούμε στο χώρο, να χειριστούμε τα αντικείμενα που μας περιβάλλουν, να αλληλεπιδράσουμε με άλλους νοήμονες οργανισμούς κλπ. Η έμφαση στο σκοπό της νόησης μας οδηγεί να επαναπροσδιορίσουμε ορισμένα θεμελιώδη ζητήματα της γνωστικής ψυχολογίας σε νέα βάση, αναθεωρώντας τις θεωρητικές εισροές και εκροές των μοντέλων μας. Για παράδειγμα, αν σκοπός της νόησης είναι η δράση τότε το ζητούμενο δεν είναι ένα σύστημα «αναγνώρισης», όπου μια οπτική εισροή θα δίνει ως εκροή μια «αναπαραστάση μήλου», αλλά ένα σύστημα που θα παράγει τα κατάλληλα προγράμματα για την κίνηση που πιάνει το μήλο και το μεταχειρίζεται ως φαγώσιμο, διαφορετικά από ένα σφυρί ή ένα φίδι (όλα εξίσου «αντικείμενα» του περιβάλλοντος υπό τη στενή οπτική έννοια). Μια θεώρηση άμεσης σύζευξης μεταξύ νόησης και δράσης, η οποία βασίζεται στο σώμα, καταλήγει σε μια θεώρηση όπου η νόηση *είναι* δράση, καθώς συλλέγει ενεργά πληροφορίες και δεν ανταποκρίνεται παθητικά σε αυτές.
6. *Η χρονικά αποκομμένη νόηση βασίζεται στο σώμα.* Λέγοντας «χρονικά αποκομμένη» (off-line) εννοούμε τη νόηση που δεν γίνεται σε στενή σύζευξη με το περιβάλλον, εγκατεστημένη και υπό χρονική πίεση, όπως θέλουν οι θέσεις 1 και 2. Παρότι η προσέγγιση της ενσώματης νόησης δίνει έμφαση στις άμεσες και αλληλεπιδραστικές διεργασίες, δεν θα μπορούσε να αγνοήσει ότι υπάρχουν και νοητικές λειτουργίες που πραγματοποιούνται έξω από τον ρου της δράσης, με χαρακτηριστικά πιο αφαιρετικά

¹⁶Βλ. Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58, 1, 7–19.

και συλλογιστικά, όπως η νοερή απεικόνιση, η αναπόληση κ.ά. Για τις περιπτώσεις αυτές η ενσώματη προσέγγιση υποστηρίζει ότι βασίζονται στις αναπαραστάσεις του σώματος και των δυνατοτήτων που αυτό παρέχει, συμπεριλαμβάνοντας την εμπειρία στο χειρισμό του σώματος για τη λύση προβλημάτων. Η οπτική εξέταση αντικειμένων, δηλαδή η ενεργητική διεργασία της αντίληψης, αξιοποιείται αποσυζευγμένη για τη δημιουργία και το χειρισμό νοερών εικόνων. Ο κινητικός προγραμματισμός της ομιλίας στρατολογείται αποσυνδεδεμένος από τους αρθρωτές της φωνητικής οδού για τη συγράτηση λεκτικών πληροφοριών στην ενεργό (ή «εργαζόμενη») μνήμη. Το βίωμα μιας εμπειρίας ανασυντίθεται εν μέρει στην ανάμνησή της συμπαρασύροντας όψεις και κινήσεις της παρελθοντικής στιγμής. Επανειλημμένα και εδραιωμένα πρότυπα δράσης αποσυνδέονται σταδιακά από την εγκατεστημένη εφαρμογή τους και αποτελούν τη βάση της αφαιρετικής σκέψης και του συλλογισμού.

5.3 Εννοιολογική μεταφορά και αφαιρετικότητα

Η ενσώματη προσέγγιση στη γλώσσα συνίσταται στη θεωρία της *εννοιολογικής μεταφοράς*,¹⁷ σύμφωνα με την οποία οι έννοιες (και κατ' επέκταση οι λέξεις που χρησιμοποιούμε) βασίζονται, άμεσα ή έμμεσα, σε εμπειρίες με τις φυσικές διαστάσεις του χώρου διαμέσου της δράσης του σώματος. Αυτό σημαίνει πως οι έννοιες που αφορούν σε φυσικές διαστάσεις (μέγεθος, θέσεις, κίνηση) είναι πρωτογενείς ενώ οι αφηρημένες έννοιες είναι δευτερογενείς, εξαρτημένες από τις προηγούμενες μέσω της μεταφοράς. Ένα συνηθισμένο παράδειγμα είναι η έννοια του χρόνου, για την οποία κατά κανόνα χρησιμοποιούμε χωρικές εκφράσεις. Για παράδειγμα, θεωρώντας το χρόνο σα μια γραμμή, στην οποία οι χρονικές στιγμές θεωρούνται σημεία, νοείται κάποιο σημείο να βρίσκεται μπροστά ή πίσω από κάποιο άλλο («κοιτάζουμε μπροστά», «αφήνουμε πίσω το παρελθόν», «έχει το μέλλον μπροστά του», «πάμε πίσω στο χρόνο»). Θεωρώντας τη ροή του χρόνου ως κίνηση, έχουμε δύο εναλλακτικές επιλογές: Η μία είναι να κινείται το υποκείμενο πάνω στη γραμμή του χρόνου, ενώ η άλλη είναι να είναι σταθερό το υποκείμενο και να κινείται η γραμμή. Και οι δύο εναλλακτικές αξιοποιούνται στη σύλληψη και επικοινωνία των χρονικών εννοιών, για παράδειγμα «όταν έρθει η ώρα», «η προθεσμία πλησιάζει», «μας έφυγε η μέρα», «έρχεται η άνοιξη», αλλά και «βαδίζουμε προς την καταστροφή», «θα συναντήσουμε προβλήματα», «πλησιάζουμε στο τέλος», «έχουμε περάσει το χειμώνα», «φτάσαμε στο Γενάρη». . .

Η εννοιολογική μεταφορά χρησιμοποιείται κατά κόρον για αφηρημένες έννοιες. Για παράδειγμα η σκέψη μπορεί να θεωρηθεί κίνηση και οι ιδέες θέσεις στο χώρο. Έτσι έχουμε τις εκφράσεις «πού τρέχει ο νους σου;», «πώς έφτασες σε αυτό το συμπέρασμα;», «μην πηδήξετε κανένα βήμα στη λύση του προβλήματος», «κόλλησα», «ακολούθησε αυτή τη σκέψη», «το κυνηγάω αυτό το θέμα εδώ και καιρό», «με οδήγησε σε μια νέα ιδέα», «παρακολουθώ τι λες», «πρέπει να πάμε πίσω και να το ξαναπιάσουμε από την αρχή» κλπ. Για

¹⁷ Παρουσίαση βασισμένη στους Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh*. Basic Books, και Gibbs, R. W. (2006). *Embodiment and cognitive science*. Cambridge University Press.

τη σκέψη υπάρχουν και άλλες εννοιολογικές μεταφορές, μερικώς ασύμβατες μεταξύ τους, που τις χρησιμοποιούμε για να εκφράσουμε διαφορετικές εκφάνσεις για τις οποίες δεν προσφέρεται η κίνηση. Για παράδειγμα, η σκέψη είναι αντίληψη («ξεκάθαρη ιδέα», «έριξε φως στην υπόθεση», «έχει παρωπίδες», «δεν ακούει»), η σκέψη είναι χειρισμός αντικειμένων («ρίξε καμιά ιδέα», «ανταλλάξαμε ιδέες», «ας το σχηματοποιήσουμε καλύτερα αυτό», «αυτή η ιδέα έχει πολλές όψεις», «έπιασε το επιχείρημα και το ξέσκισε»), η σκέψη είναι απόκτηση ιδεών ως κατανάλωση τροφής («το κατάπιε αμάσητο», «έχει όρεξη για μάθηση», «ανούσια θεωρία», «δεν μπορώ να το χωνέψω»), κ.ά. Θα μπορούσε κανείς να αντιτείνει ότι πρόκειται για μεταφορική χρήση της γλώσσας και όχι των εννοιών, όμως εδώ οι Lakoff & Johnson έχουν στοιχειοθετήσει εξαντλητικά την παραγωγικότητα των μεταφορικών σχημάτων, καθώς δεν πρόκειται αποκλειστικά για γνωστές και προκαθορισμένες εκφράσεις αλλά για αναλογίες εννοιών από τις οποίες μπορούν να κατασκευαστούν νέες μεταφορικές εκφράσεις οι οποίες να είναι όχι απλώς κατανοητές αλλά εντελώς φυσικές. Για παράδειγμα, μπορούμε να πούμε, για μια περίπλοκη θεωρητική πρόταση: «πιάσε την ιδέα, κόψε τη σε μικρά κομμάτια, και προσπάθησε να στηρίξεις γερά το καθένα, προτού συνθέσεις το πλήρες σχήμα». Μια τέτοια φράση δεν είναι παγιωμένη στη γλώσσα και δεν θα έβγαζε νόημα αν δεν υποστηριζόταν από ένα πλούσιο και παραγωγικό πλαίσιο εννοιολογικής μεταφοράς. Άλλωστε, αν δεν υπήρχε η αναλογία στο εννοιολογικό πεδίο δεν θα μπορούσε να δικαιολογηθεί ούτε το γλωσσικό φαινόμενο της μεταφοράς σε παγιωμένες εκφράσεις.

Η εννοιολογική μεταφορά δίνει σαφή υπόσταση στην ενσώματη προσέγγιση αναφορικά με αφηρημένες νοητικές αναπαραστάσεις και διεργασίες, τεκμηριώνοντας σε εννοιολογικό και γλωσσικό επίπεδο ότι είναι όχι απλώς εφικτή αλλά πολύ συνηθισμένη η μεταφορά και αξιοποίηση νοητικών σχημάτων από ένα από και συγκεκριμένο πεδίο αντίληψης-δράσης σε άλλα πεδία αφαιρετικής νόησης. Κατ' επέκταση, η αποσυζευγμένη, αφαιρετική νόηση αντιμετωπίζεται από την ενσώματη προσέγγιση ως μια προέκταση των δομών και διεργασιών που έχουν προκύψει από τη δράση του σώματος και βασίζονται σε παραλλαγές και συνδυασμούς αισθησιοκινητικών προγραμμάτων. Σε εντελώς αντίθετη κατεύθυνση από την κλασική, συμβολική προσέγγιση, όπου η νόηση είναι θεμελιωδώς αφαιρετική, τυπική (συντακτική) και αποκομμένη από το σώμα και το περιβάλλον, η ενσώματη προσέγγιση θέλει τη νόηση θεμελιωδώς συγκεκριμένη, βασισμένη στο περιεχόμενο και το φορέα δράσης, σε στενή σύζευξη και αλληλεπίδραση με το περιβάλλον. Ενώ στη συμβολική προσέγγιση οι συγκεκριμένες διεργασίες είναι δείγματα-εφαρμογές των θεμελιωδών υποκείμενων τυπικών διεργασιών, στην ενσώματη προσέγγιση οι διεργασίες δεν είναι ποτέ εντελώς αφηρημένες αλλά μόνο ευέλικτοι ανασυνδυασμοί και περιορισμένες γενικεύσεις των θεμελιωδών αισθησιοκινητικών διεργασιών. Υπό αυτό το πρίσμα η νοητική λειτουργία δεν είναι ποτέ τυπική με τη μαθηματική έννοια, αλλά το πολύ πολύ να είναι επαρκώς γενικευμένη και ευέλικτη ώστε να προσεγγίζει τις εκροές μιας τυπικής λειτουργίας μέσα σε ένα ευρύ φάσμα δειγμάτων και περιστατικών. Για παράδειγμα, ο νους δεν εκτελεί ποτέ τυπική λογική σύζευξη με την αφηρημένη έννοια, αλλά μπορεί να γενικεύσει στη βάση πολλών συγκεκριμένων και ειδικών συζεύξεων που είτε έχει μάθει είτε είναι προκαθορισμέ-

νος να εκτελεί. Αντίστοιχα, στη γλωσσική συμπεριφορά, η παρατηρούμενη κανονικότητα δεν είναι αποτέλεσμα παραγωγικών κανόνων, αλλά γενίκευση από τα γνωστά δείγματα προς επαρκώς παρόμοιες περιστάσεις, στο πνεύμα και του συνδεδετισμού.

Πώς μπορεί μια αφαιρετική λειτουργία να βασιστεί σε μια συγκεκριμένη; Για παράδειγμα, στη νοερή απεικόνιση, αξιοποιούνται οι δυνατότητες επεξεργασίας οπτικών μορφών, αποκομμένες από την αισθητηριακή εισροή. Έτσι ο νους αξιοποιεί με ευέλικτο τρόπο ένα υπόστρωμα αισθητηριακής επεξεργασίας για συγγενικές λειτουργίες, κληροδοτώντας σε αυτές και τις αντίστοιχες ιδιότητες, τόσο τα πλεονεκτήματα και τις δυνατότητες όσο και τα μειονεκτήματα και τους περιορισμούς. Ο προγραμματισμός κινήσεων αναπνοής, μάσησης και κατάποσης, αποσυζευγμένος από το αρχικό πλαίσιο διατροφής, παρέχει ένα κινητικό υπόστρωμα για την εξέλιξη της προφορικής επικοινωνίας μέσω της ομιλίας. Η δυνατότητα ευέλικτης αξιοποίησης του κινητικού προγραμματισμού με νέους αυθαίρετους συνδυασμούς κινήσεων επιτρέπει τη σύσταση νέων κινητικών προτύπων ομιλίας με βάση προϋπάρχοντα συστατικά. Στη συνέχεια, σε επόμενο βήμα αφαίρεσης, ο ευέλικτος κινητικός προγραμματισμός της άρθρωσης, αποκομμένος από την κινητική εκτέλεση της ομιλίας, μπορεί να αξιοποιηθεί για τη συγκρότηση ενός βρόχου επανάληψης που διατηρεί ενεργά σχήματα φωνολογικού προγραμματισμού, δηλαδή την «εσωτερική ομιλία» που συγκρατεί πληροφορίες για μικρά χρονικά διαστήματα. Έτσι, ενώ είναι λάθος να επιχειρούμε ανάλυση της ομιλίας με αναγωγή σε διατροφικές λειτουργίες μάσησης και κατάποσης, λόγω της πλήρους αποσύζευξης, είναι εξίσου λάθος να επιχειρούμε ανάλυσή της με φορμαλιστικές και αφηρημένες δομές, διότι η προέλευσή της δεν είναι υπολογιστική αλλά αισθησιοκινητική. Η ενσώματη προσέγγιση προτείνει ότι ακόμα και οι πιο αφηρημένες μορφές σκέψης ανάγονται σε συγκεκριμένο αισθησιοκινητικό υπόστρωμα του αναπτυξιακού ή εξελικτικού παρελθόντος, και με βάση αυτό θα πρέπει να εξεταστούν.

5.4 Οικολογική ψυχολογία

Η ενσώματη προσέγγιση έχει τις ρίζες της στην οικολογική ψυχολογία του J. J. Gibson από τη δεκαετία του 1950. Η οικολογική ψυχολογία αντιτέθηκε ριζικά στο συμπεριφορισμό: ενώ η «γνωστική επανάσταση» διατήρησε το σχήμα ερέθισμα-απόκριση, παρεμβάλλοντας απλώς το νου και εστιάζοντας στις εσωτερικές του αναπαραστάσεις και διεργασίες, η οικολογική προσέγγιση απέρριψε την παθητική αντίληψη του νου ως επεξεργαστή εξωγενών πληροφοριών και θεώρησε το νου ως ενεργό συλλέκτη πληροφοριών, σε στενή σύζευξη και διαρκή αλληλεπίδραση με το περιβάλλον.¹⁸ Χωρίς να παραγνωρίζει το ρόλο οπτικών γνωρισμάτων για την αισθητηριακή επεξεργασία, ο Gibson συνέβαλε σε μια ρηξικελεύθη και ανατρεπτική προσέγγιση, επαναπροσδιορίζοντας το περιβάλλον και τα «αντικείμενα» που υπάρχουν σε αυτό με αναφορά στο υποκείμενο της αντίληψης. Απέρριψε την αντικειμενικότητα της ταυτοποίησης αντικειμένων ως ιδανικό στόχο της αντίληψης και εστίασε στη χρησιμότητα των πληροφοριών για το νου. Έτσι, ένα μήλο δεν είναι ένα κόκκινο

¹⁸Βασισμένο στο Gibson, J. J. (1979). *The ecological approach to visual perception*. Erlbaum.

αντικείμενο αλλά κάτι που τρώγεται, μια καρέκλα δεν είναι ένα έπιπλο με τέσσερα πόδια αλλά κάτι που προσφέρεται για κάθισμα κ.ο.κ. Χρησιμοποίησε ένα νεολογισμό για την προσέγγιση αυτή, την *επιδοχή* (affordance): ένα μήλο επιδέχεται φάγωμα, μια καρέκλα επιδέχεται κάθισμα, ένα χερούλι επιδέχεται πιάσιμο. Έτσι, σκοπός της οπτικής αντίληψης δεν είναι η κατηγοριοποίηση του αντικειμένου αλλά η άμεση αναγνώριση των επιδοχών που παρέχει κάθε στοιχείο του περιβάλλοντος. Λέγοντας «άμεση» εννοούμε ότι δεν παρεμβάλλεται κάποια αναπαράσταση της ταυτότητας του αντικειμένου, καθώς η αντίληψη των επιδοχών θεωρείται πρωτογενής.

Η υποκειμενοστραφής προσέγγιση της οικολογικής ψυχολογίας αφορά και στην ενεργητικότητα του υποκειμένου, καθώς η αντίληψη θεωρείται μέσο δειγματοληψίας του περιβάλλοντος και ελέγχου της δράσης. Το υποκείμενο χειρίζεται τη θέση και την κίνησή του προκειμένου να αποκομίσει πληροφορίες για τις επιδοχές του περιβάλλοντος μέσω των οπτικών υφών (textures) και της σχέσης κάλυψης (occlusion) και αποκάλυψης αντικειμένων καθώς μεταβάλλεται το οπτικό πεδίο. Η ταυτόχρονη μεταβολή όλων των σημείων του οπτικού πεδίου που συνοδεύει την κίνηση του υποκειμένου ονομάζεται *οπτική ροή* (optical flow). Κίνηση του υποκειμένου προς τα εμπρός έχει ως αποτέλεσμα την κίνηση όλων των στοιχείων του οπτικού πεδίου προς τα έξω· ακίνητο μένει μόνο ένα σημείο, εκείνο προς το οποίο κατευθυνόμαστε. Τα κοντινά στοιχεία του οπτικού πεδίου κινούνται ταχύτερα από τα πιο μακρινά. Έτσι, η οπτική ροή μας πληροφορεί για την κατεύθυνση της κίνησης και μπορεί να χρησιμοποιηθεί για πλοήγηση, διατήρηση ισορροπίας, αποφυγή εμποδίων κ.ά. Το άτομο αλληλεπιδρά με την οπτική ροή ρυθμίζοντας την κίνηση του σώματος και του κεφαλιού έτσι ώστε να εξερευνά επιλεκτικά τα στοιχεία του περιβάλλοντος και παράλληλα να ελέγχει την πορεία του. Έτσι, αντί για τη συνήθη εικόνα εισροής-εκροής των συστημάτων «επεξεργασίας πληροφοριών», το οικολογικό υποκείμενο συγκροτεί ένα ενιαίο αισθησιοκινητικό μηχανισμό για την ενεργή εμπλοκή του με το περιβάλλον.

Οι ιδέες της οικολογικής ψυχολογίας έχουν οδηγήσει σε πολύ ενδιαφέρουσες προσεγγίσεις στη ρομποτική, όπως το Cog project του Rodney Brooks στο MIT, στο οποίο επιχειρείται να δοθεί η δυνατότητα σε ένα τεχνητό σύστημα να αναπτυχθεί νοητικά διαμέσου της ενεργής αλληλεπίδρασης με το περιβάλλον και όχι χρησιμοποιώντας προγραμματισμένες συμβολικές δομές. Το ρομπότ διαθέτει κινητά μέλη ώστε να μπορεί να εξερευνά και να πειραματίζεται, καθώς και να χειρίζεται αντικείμενα και να αλληλεπιδρά με ανθρώπους. Η ενσώματη προσέγγιση στην αντίληψη έχει οδηγήσει και σε νέες εκδοχές της οικολογικής προσέγγισης κατά τις οποίες η επίγνωση των οπτικών αντιλημμάτων αναδύεται μέσα από την αισθησιοκινητική αλληλεπίδραση με το περιβάλλον. Η ιδέα είναι ότι η συνειδητή αντίληψη των αντικειμένων συνίσταται στο σύνολο των προσδοκιών που αφορούν στα αντικείμενα, για παράδειγμα πώς θα μεταβληθεί η όψη τους σε κάθε δυνατή κίνηση του κεφαλιού ή σε αλλαγή φωτισμού.¹⁹ Η προσέγγιση αυτή απορρίπτει την εσωτερική νοητική αναπαράσταση των αντικειμένων καθαυτών και την αντικαθιστά με άρρητη γνώση για

¹⁹Βλ. O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24, 5, 939–1031.

τις συνέπειες της δράσης, σε ένα κλειστό βρόχο αλληλεπίδρασης με το περιβάλλον, το οποίο διαρκώς παρέχει όλες τις πληροφορίες για κάθε αντικείμενο. Έτσι, το περιβάλλον παίζει το ρόλο του νοητικού σημειωματορίου, καθιστώντας περιττή την εσωτερίκευση των αντικειμένων, και αποτελεί μέρος του εκτεταμένου νου.

Αν η αντίληψη είναι θέμα αισθησιοκινητικών εξαρτήσεων (sensorimotor contingencies), τότε ποιος είναι ο ειδικός ρόλος των αισθητηρίων οργάνων; Πόση σημασία έχει η εισροή από τον αμφιβληστροειδή προκειμένου να παραχθεί συνειδητή οπτική αντίληψη; Οι μελέτες συστημάτων αισθητηριακής υποκατάστασης (sensory substitution) έχουν δείξει εδώ και δεκαετίες ότι ο αμφιβληστροειδής δεν είναι αναγκαίος, διότι «δεν βλέπουμε με τα μάτια, αλλά με τον εγκέφαλο».²⁰ Εκείνο που είναι αναγκαίο είναι μια διάταξη αλληλεπίδρασης με το περιβάλλον που να διατηρεί τις σχέσεις μεταξύ δράσης και αισθητηριακής εισροής, όπως για παράδειγμα μια βιντεοκάμερα που στηρίζεται στο κεφάλι ή κρατιέται στο χέρι. Το σήμα της εικόνας μπορεί να παρέχεται με ηλεκτρόδια στη γλώσσα ή με μηχανικούς μικροδονητές στην πλάτη.²¹ Όταν η εισερχόμενη πληροφορία έχει την κατάλληλη σχέση οπτικοχωρικής αλληλεπίδρασης με το περιβάλλον, το αισθητηριακό μέσο εισροής δεν παίζει ρόλο. Με τέτοιες συσκευές και μετά από κάποια εξάσκηση ένας τυφλός μπορεί να δει αρκετά καλά για να πλοηγηθεί στο χώρο, να αναγνωρίσει και να πιάσει αντικείμενα, και όλα αυτά με οπτικοχωρική επίγνωση όπως και οι βλέποντες.

6 Πιθανοκρατική (μπεϊζιανή) προσέγγιση

Μια προσέγγιση που γνωρίζει μεγάλη ανάπτυξη τα τελευταία χρόνια στη γνωσιακή επιστήμη εστιάζει στον πιθανοκρατικό χαρακτήρα της νόησης και της αποδίδει την ιδιότητα της ορθολογικότητας, δηλαδή της πληροφοριακά βέλτιστης λειτουργίας. Οι νοητικές λειτουργίες αντιμετωπίζονται ως πιθανοκρατικές συναγωγές, ενώ οι νοητικές αναπαραστάσεις αφορούν κατανομές πιθανοτήτων εναλλακτικών υποθέσεων. Τα «περιεχόμενα» των μοντέλων είναι τυχαίες μεταβλητές (random variables) και οι πιθανοτικές κατανομές των τιμών τους, οι οποίες είναι δεσμευμένες (conditional) μεταξύ των μεταβλητών, αντιστοιχώντας στις αλληλοεξαρτήσεις τους. Η προσέγγιση αυτή είναι γνωστή ως μπεϊζιανή διότι βασίζεται στον κανόνα του Μπέιζ (Bayes) για τον υπολογισμό των σχετικών πιθανοτήτων των εναλλακτικών υποθέσεων.

Τα δύο κυριότερα χαρακτηριστικά της προσέγγισης αυτής είναι (α) ο πιθανοκρατικός (probabilistic) χαρακτήρας της και (β) η υπόθεση της ορθολογικότητας (rationality). Λέγοντας «πιθανοκρατικό χαρακτήρα» εννοούμε δύο πράγματα: Αφενός ότι η νοητική λειτουργία λειτουργεί πάντοτε σε συνθήκες αβεβαιότητας (uncertainty). Και αφετέρου—και κυριότερο—ότι η ίδια η νοητική λειτουργία λειτουργεί με βάση πιθανότητες, καθώς αυτό

²⁰Bach-y-Rita, P. (2005). Emerging concepts of brain function. *Journal of Integrative Neuroscience*, 4, 2, 183-205.

²¹Bach-y-Rita, P., & Kercel, S. W. (2003). Sensory substitution and the human-machine interface. *Trends in Cognitive Sciences*, 7, 12, 541-546

που υπολογίζεται κάθε φορά δεν είναι «ποια είναι η σωστή απάντηση» αλλά ποια είναι η κατανομή πιθανοτήτων που αντιστοιχεί στις εναλλακτικές δυνατές επιλογές (είτε αυτές είναι διακριτές είτε σχηματίζουν ένα συνεχές). Δηλαδή οι μπειζιανοί υπολογισμοί δεν παράγουν σημειακές εκτιμήσεις αλλά ολόκληρες κατανομές. Πάνω σε αυτό το υπόβαθρο, η υπόθεση της ορθολογικότητας θέτει τα κριτήρια προσδιορισμού των κατανομών και επιλογής εναλλακτικών: Λέγοντας «ορθολογικότητα» εννοούμε ότι το σύστημα συνδυάζει όλες τις πληροφορίες που έχει στη διάθεσή του με τον βέλτιστο τρόπο. Δηλαδή, οι κατανομές που υπολογίζει και οι αποφάσεις που λαμβάνει είναι τέτοιες ώστε να πετυχαίνει το απολύτως μικρότερο σφάλμα που είναι θεωρητικά δυνατό με βάση τις υπάρχουσες πληροφορίες. Εξ ου και τα συστήματα αυτά ονομάζονται επίσης και *μοντέλα βέλτιστων αποφάσεων* (optimal decision models).

6.1 Πιθανοκρατική νόηση

Η πιθανοκρατική προσέγγιση εξετάζει ένα ευρύ φάσμα νοητικών φαινομένων, που συμπεριλαμβάνει αντίληψη, κατηγοριοποίηση, συλλογισμούς, αποφάσεις, και διάφορες περιπτώσεις μάθησης.²² Σε κάθε περίπτωση το ζήτημα προσεγγίζεται ως πρόβλημα συναγωγής, αναζητώντας την πιθανότερη «εξήγηση» για τα παρατηρούμενα δεδομένα. Για παράδειγμα, στην αντίληψη, η εισροή απέχει πολύ από το να προσδιορίσει με σαφήνεια τη δομή του περιβάλλοντος. Αυτό έχει γίνει σαφές από την πλήρη αποτυχία των πρώτων συστημάτων τεχνητής όρασης και ακοής και από την (περιορισμένη) επιτυχία των πιο πρόσφατων συστημάτων, η οποία οφείλεται στην απομάκρυνση από τις αρχικές υπολογιστικές φιλοδοξίες και την υιοθέτηση στατιστικών μεθόδων. Έτσι, μια οπτική εισροή δεν επαρκεί για να προσδιοριστούν τα περιεχόμενα του οπτικού πεδίου και να παραχθούν τα αντίστοιχα αντιλήμματα, μια ακουστική εισροή δεν επαρκεί για να προσδιοριστούν τα αντίστοιχα συμβάντα που προκάλεσαν τον ήχο, κλπ. Συνεπώς, δουλειά του νοητικού συστήματος είναι να κάνει υποθέσεις για το τι θα μπορούσε να έχει προκαλέσει την παρατηρούμενη οπτική, ακουστική, ή άλλη εισροή, και να αξιολογήσει τις υποθέσεις αυτές για να επιλέξει την πιθανότερη. Αυτή που επιλέγεται αποτελεί και το «αντίλημμα», ως αποτέλεσμα αυτής της (μη συνειδητής) διεργασίας συγκριτικής αξιολόγησης υποθέσεων.

Η περιγραφή αυτή ισχύει σε γενικές γραμμές και για τα πεδία της κατηγοριοποίησης, της αιτιακής αντίληψης, της γλωσσικής επεξεργασίας και άλλα. Για παράδειγμα, τα προβλήματα κατηγοριοποίησης συνήθως συνίστανται σε παρουσίαση κάποιων ερεθισμάτων τα οποία καλείται ο συμμετέχων να ομαδοποιήσει σε συγκεκριμένες κατηγορίες. Οι κατηγορίες μπορεί να είναι εκ των προτέρων γνωστές ή μπορεί να μαθαίνονται μέσα από την ίδια τη διαδικασία κατηγοριοποίησης μέσα από ανατροφοδότηση. Κάθε δείγμα που παρουσιάζεται προς κατηγοριοποίηση γίνεται αντιληπτό ως μια συστάδα αντιληπτικών γνωρισμάτων, ενώ κάθε γνωστή κατηγορία αναπαρίσταται ως μια κατανομή πιθανοτήτων για τα επιμέ-

²²Βασισμένο στο Chater, N., Oaksford, M., Hahn, U., & Heit, E. (2010). Bayesian models of cognition. *WIREs Cognitive Science*, 1, 811-823.

ρους γνωρίσματα και τους συνδυασμούς τους. Η κατανομή αυτή μπορεί να αντιστοιχεί χοντρικά σε ένα «πρότυπο» με την έννοια ότι υπάρχει μια σαφής κεντρική τάση, βάσει της οποίας παράγονται τα διάφορα φαινόμενα τυπικότητας. Εναλλακτικά, η κατανομή μπορεί να αντιστοιχεί σε επιμέρους συγκεκριμένα στιγμιότυπα-υποδείγματα, εάν η κεντρική τάση απουσιάζει ή είναι πολύ ασθενής και, αντίθετα, κυριαρχούν τοπικά μέγιστα στο τοπίο πιθανότητας των γνωρισμάτων. Έτσι, η πιθανοκρατική προσέγγιση μπορεί να επιλύσει την ένταση μεταξύ των δύο ειδών θεωριών, ενοποιώντας τις σε ένα συνεχές που επιτρέπει και ισχυρή επίδραση προτύπου αλλά και τοπικές επιδράσεις επιμέρους υποδειγμάτων.

Αντίστοιχα, η αιτιακή μάθηση, δηλαδή η μάθηση «αιτιών» για παρατηρούμενα συμβάντα-«αποτελέσματα» συνίσταται στη σύνδεση παρατηρούμενων μεταβλητών (γνωρισμάτων, συμβάντων, ή πιο αφαιρετικών συστάδων) με προβλεπτικές σχέσεις. Αυτό μπορεί να συμβεί αν κάποια συστάδα παρατηρήσεων συστηματικά συνεμφανίζεται ή ακολουθεί κάποια άλλη συστάδα παρατηρήσεων, οπότε το σύστημα μπορεί να προβλέψει επιτυχώς την εμφάνισή τους από τις άλλες. Η συνεμφάνιση μπορεί σταδιακά να αποδοθεί σε αιτιότητα, με την έννοια ότι το νοητικό σύστημα μαθαίνει τη συστηματική σχέση ανάμεσα στις παρατηρήσεις και τη συμπεριλαμβάνει στο μοντέλο του για τον κόσμο, δηλαδή στις κατανομές πιθανοτήτων των παρατηρήσεων και τις σχέσεις μεταξύ τους. Όταν μια σύνδεση μεταξύ πιθανοτήτων παρατηρήσεων είναι σταθερή και συμβατή με το υπόλοιπο πλέγμα αναπαραστάσεων τότε συνιστά μια νοητή αιτιότητα, δηλαδή, το σύστημα θεωρεί τις σχέσεις ανάμεσα στις παρατηρήσεις ως «αιτιακές». Αυτό δεν σημαίνει ότι οποιαδήποτε συγκυριακή συνεμφάνιση γνωρισμάτων θα οδηγήσει σε αιτιακή αναπαράσταση, καθώς για την υποστήριξη της αντιληπτής αιτιότητας φαίνεται ότι απαιτείται ένταξη σε ένα πιο πλούσιο σύμπλεγμα αλληλεπιδράσεων και συσχετίσεων. Ένα συμβάν μπορεί να γίνει αντιληπτό ως αιτία άλλου συμβάντος-«αποτελέσματος» μόνο όταν δεν υπάρχουν άλλες αιτίες που να το προκάλεσαν. Έτσι, δεν αντιλαμβανόμαστε το ζυπνητήρι μας ως αιτία της καμπάνας της εκκλησίας ακόμα και αν το έχουμε βάλει να χτυπάει ακριβώς και σταθερά την ίδια ώρα.

Στη γλωσσική επεξεργασία, η πιθανοκρατική προσέγγιση μπορεί να εφαρμοστεί τόσο σε αντιληπτικό όσο και σε νοηματικό επίπεδο. Στην αντίληψη της ομιλίας, οι ήχοι δεν επαρκούν για τον προσδιορισμό των φθόγγων που πρόφερε (ή, συνηθέστερα, είχε την πρόθεση να προφέρει) ο ομιλητής, διότι ο θόρυβος και η συνάρθρωση μεταξύ φθόγγων έχουν ως αποτέλεσμα την παραγωγή ήχων στους οποίους δεν διακρίνονται ξεκάθαρα οι επιμέρους φθόγγοι αλλά αντίθετα συνυπάρχουν διάφορα φωνητικά γνωρίσματα. Πρόκειται για ένα αντιληπτικό πρόβλημα παρόμοιο με τη γενική περίπτωση οπτικής ή ακουστικής αντίληψης, μόνο που τα «συμβάντα» που καλείται να αναγνωρίσει ο ακροατής δεν είναι οι κινήσεις της φωνητικής οδού του ομιλητή αλλά οι λέξεις που είχε την πρόθεση να προφέρει. Παρότι κάθε λέξη έχει θεωρητικά μια «κανονική» εκφορά, αποτελούμενη από συγκεκριμένη, ευδιάκριτη ακολουθία φθόγγων, ο παραγόμενος ήχος δεν έχει τα αντίστοιχα χαρακτηριστικά παρά μόνο σε πολύ αργή, σκόπιμα ευδιάκριτη εκφορά της μεμονωμένης λέξης, κάτι που δεν συμβαίνει ποτέ σε αυθόρμητη προφορική συζήτηση. Έτσι, το ζήτημα της αντίληψης ομιλίας ανάγεται σαφώς σε πρόβλημα πιθανοκρατικής συναγωγής.

Πέρα από την αντίληψη των μεμονωμένων λέξεων, στο επίπεδο του φραστικού νοήματος ο ακροατής πρέπει να αναγνωρίσει αφενός την κατάλληλη σημασία των λέξεων (που συχνά έχουν περισσότερες από μία) στο συγκεκριμένο πλαίσιο, και αφετέρου τη συντακτική δομή με την οποία έχουν συνδυαστεί, η οποία σε συνδυασμό με τη σημασία των λέξεων συνεισφέρει στο νόημα της φράσης. Οι αμφισημίες και στα δύο αυτά πεδία δεν επιτρέπουν τον μονοσήμαντο προσδιορισμό του νοήματος. Σε κλασικά παραδείγματα όπως «είδα το παιδί με τα κιάλια» ο ακροατής πρέπει να αντιληφθεί αν πρόκειται για ένα παιδί που κρατάει κιάλια ή αν ο ομιλητής χρησιμοποίησε κιάλια για να δει το παιδί. Οι αμφισημίες δεν αποτελούν μεμονωμένες ή ολιγάριθμες παραδοξότητες. Αντίθετα, εμφανίζονται σε όλα τα επίπεδα και συχνά στη γλωσσική επικοινωνία, με αποτέλεσμα την ανάγκη πιθανοκρατικής συναγωγής για την επεξεργασία και την επικοινωνία.

Στο πλαίσιο αυτό πρέπει να αναφερθεί ο ρόλος της πιθανοκρατικής προσέγγισης στη γλωσσική κατάκτηση, καθώς πρόκειται για σχετικά πρόσφατη εξέλιξη που καταρρίπτει παγιωμένες αντιλήψεις δεκαετιών. Συγκεκριμένα, όπως αναφέρεται στη σχετική ενότητα της γλωσσολογίας (σελ. 109), ορισμένες θεωρητικές προσεγγίσεις υποθέτουν ότι οι πληροφορίες στις οποίες εκτίθεται το βρέφος και το παιδί είναι θεμελιωδώς ανεπαρκείς για να οδηγήσουν στην απόκτηση ενός ολοκληρωμένου γλωσσικού συστήματος. Έτσι, οδηγούνται στην παραδοχή ότι οι βασικές γλωσσικές δομές πρέπει να είναι έμφυτες στον άνθρωπο. Σε συνδυασμό με ολοένα λεπτομερέστερη καταγραφή των πραγματικών γλωσσικών δεδομένων στα οποία εκτίθεται ένα παιδί, οι εξελίξεις στη στατιστική επεξεργασία και ειδικότερα στα πιθανοκρατικά μοντέλα έχουν δείξει ότι η υπόθεση αυτή είναι λανθασμένη. Δηλαδή, στην πραγματικότητα είναι εφικτό να συναχθούν οι ορθές δομές με βάση την κατανομή των δεδομένων εισροής, τουλάχιστον για κάποιες περιπτώσεις.²³ Συνεπώς δεν ισχύει το επιχείρημα της ένδειας του ερεθίσματος και δεν είναι απαραίτητη η παραδοχή της εμφυτότητας.

Η πιθανοκρατική προσέγγιση εφαρμόζεται επίσης με μεγάλη επιτυχία στο πεδίο της λογικής συναγωγής, τόσο της επαγωγικής όσο και της παραγωγικής, πετυχαίνοντας αποτελέσματα πολύ κοντινότερα στις ανθρώπινες επιδόσεις από εναλλακτικές που βασίζονται στην τυπική λογική. Σε κάθε περίπτωση, η πιθανοκρατική προσέγγιση λαμβάνει υπόψη τη βαρύτητα όλων των διαθέσιμων στοιχείων για την εξαγωγή του συμπεράσματος, ακόμα και όταν αυτό παραβιάζει τη λογική δομή του προβλήματος (όπως αυτή εννοείται υπό το πρίσμα του φορμαλισμού). Έτσι, μπορεί να προσομοιώσει φαινόμενα επαγωγικής ισχύος, δηλαδή διαφορές στη βεβαιότητα με την οποία συνάγονται συμπεράσματα, οι οποίες παρατηρούνται στην ανθρώπινη σκέψη και δικαιολογούνται αν λάβουμε υπόψη την κατανομή των πληροφοριών στο περιβάλλον. Επίσης μπορεί να προσομοιώσει περιπτώσεις «εξαιρέσεων» σε υποθετικά γενικούς «κανόνες» με το να αναπαριστά τους κανόνες ως πιθανοκρατική (και όχι απόλυτη) σύνδεση μεταξύ παρατηρήσεων. Τα φαινόμενα αυτά δείχνουν ότι η αν-

²³βλ., λ.χ., Perfors, A., Tenenbaum, J., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, 118, 306–338. Hsu, A. S., & Chater, N. (2010). The logical problem of language acquisition: A probabilistic perspective. *Cognitive Science*, 34, 972–1016.

θρόπινη σκέψη είναι προσαρμοσμένη για να λειτουργεί αποτελεσματικά στο περιβάλλον της, με συγκεκριμένες κατανομές πιθανοτήτων, υποστηρίζοντας βέλτιστες αποφάσεις υπό συνθήκες αβεβαιότητας, αντίθετα με τα φορμαλιστικά μοντέλα που εστιάζουν μόνο στη δομή και όχι στο περιεχόμενο των συλλογισμών.

Τέλος, πρέπει να αναφερθεί ότι η πιθανοκρατική προσέγγιση υποστηρίζει διαφορετικούς τύπους μάθησης. Όπως προαναφέρθηκε, τα πιθανοκρατικά μοντέλα περιέχουν τυχαίες μεταβλητές και τις δεσμευμένες κατανομές πιθανοτήτων των τιμών τους. Οι μεταβλητές μπορούν να αντιστοιχούν σε λανθάνουσες υποθέσεις για τη δομή του κόσμου καθώς και σε εμπειρικές παρατηρήσεις (εισορές). Οι αλληλεξαρτήσεις μεταξύ των μεταβλητών δεσμεύουν τις πιθανοτικές κατανομές των τιμών τους. Δηλαδή το ποιες τιμές μπορεί να πάρει μια μεταβλητή και με τι πιθανότητα μπορεί να εξαρτάται από άλλες μεταβλητές του συστήματος. Όταν ισχύει αυτό λέμε ότι υπάρχει αιτιακή σύνδεση ανάμεσα στις μεταβλητές.

Με βάση αυτά τα γενικά χαρακτηριστικά της δομής κάθε πιθανοκρατικού μοντέλου, μπορούν να οριστούν δύο διαφορετικά είδη μάθησης:²⁴ Το πρώτο ονομάζεται μάθηση παραμέτρων (parameter learning), και αντιστοιχεί στη μεταβολή των κατανομών πιθανοτήτων των τιμών των μεταβλητών. Μαθαίνει, δηλαδή, το σύστημα, ποιες είναι οι πιθανότερες τιμές των μεταβλητών σε διαφορετικές περιστάσεις, προσαρμοζόμενο έτσι στην κατανομή των παρατηρήσεων που κάνει. Για τη μάθηση αυτή οι αιτιακές συνδέσεις του μοντέλου παραμένουν σταθερές. Δηλαδή παραμένει σταθερό το ποιες μεταβλητές πρέπει να ληφθούν υπόψη για τον υπολογισμό κάθε πιθανοτικής κατανομής. Το δεύτερο είδος μάθησης ονομάζεται μάθηση δομών (structure learning), και αντιστοιχεί στις αιτιακές συνδέσεις μεταξύ των μεταβλητών. Δηλαδή το σύστημα εξετάζει διαφορετικές εναλλακτικές δομές διασύνδεσης ανάμεσα στις μεταβλητές, επιδιώκοντας να «μάθει» (ακριβέστερα, να επιλέξει) ποιες μεταβλητές εξαρτώνται από ποιες άλλες· με άλλα λόγια, ποιες μεταβλητές πρέπει να λαμβάνονται υπόψη στον υπολογισμό των δεσμευμένων κατανομών καθεμιάς. Μοντέλα μάθησης δομών έχουν εφαρμοστεί με επιτυχία στη διερεύνηση των παραγόντων που λαμβάνουν υπόψη οι άνθρωποι για να λάβουν τις αποφάσεις τους.

Η υπόθεση της βέλτιστης αξιοποίησης των δεδομένων για τη μάθηση δομών και παραμέτρων παρέχει στα πιθανοκρατικά μοντέλα τη δυνατότητα «ενεργητικής μάθησης» (active learning). Η ενεργητική μάθηση αποτελεί στοχευμένη συλλογή δεδομένων, δηλαδή σκόπιμη παρατηρησιακή πράξη, η οποία καθοδηγείται από την τρέχουσα γνώση του μοντέλου. Συγκεκριμένα, όταν το σύστημα διαθέτει δύο ή περισσότερες εναλλακτικές υποθέσεις, μπορεί να υπολογίσει ποιο στοιχείο πρέπει να παρατηρηθεί προκειμένου να επιλεγεί η καλύτερη υπόθεση. Το στοιχείο αυτό θα πρέπει να είναι εκείνο για το οποίο οι δύο εναλλακτικές υποθέσεις παρουσιάζουν μέγιστη απόκλιση (είναι πολύ πιθανό για τη μία υπόθεση αλλά απίθανο για την άλλη). Αντί να καταγράφει παθητικά παρατηρήσεις, οι οποίες μπορεί απλώς να επαναλαμβάνουν προηγούμενες παρατηρήσεις και να εξηγούνται

²⁴Βασισμένο στο Jacobs, R. A., & Kruschke, J. K. (2011). Bayesian learning applied to human cognition. *WIREs Cognitive Science*, 2, 8-21.

εξίσου από τις εναλλακτικές υποθέσεις, ένα πιθανοκρατικό σύστημα ενεργητικής μάθησης μπορεί να συγκλίνει με μεγάλη ταχύτητα στη βέλτιστη υπόθεση επιδρώντας στον παρατηρησιακό χώρο και συλλέγοντας τα πιο πληροφοριακά (διαγνωστικά) δεδομένα.

6.2 Αναθεώρηση πεποιθήσεων

Όπως προαναφέρθηκε, η πιθανοκρατική προσέγγιση θέτει τα νοητικά ζητήματα ως προβλήματα συναγωγής: Με δεδομένα κάποια παρατηρούμενα στοιχεία, ο νους βγάζει συμπεράσματα σχετικά με την υποκείμενη δομή που μπορεί να έχει παράξει αυτά τα στοιχεία και εν συνεχεία χρησιμοποιεί τη γνώση αυτή προκειμένου να κάνει προβλέψεις για μελλοντικές παρατηρήσεις.²⁵ Η λύση αυτών των προβλημάτων περνάει μέσα από την αναθεώρηση των πεποιθήσεων με βάση το ορθολογικό κριτήριο του κανόνα του Μπέιζ. Ο κανόνας αυτός δεν είναι παρά ένα στοιχειώδες πόρισμα της θεωρίας πιθανοτήτων σχετικά με την αντιστροφή της σχέσης της δεσμευμένης πιθανότητας δύο μεταβλητών. Συγκεκριμένα, αν έχουμε κάποια παρατηρούμενα στοιχεία d και μια υπόθεση h για τη διεργασία που μπορεί να έχει παράξει τα στοιχεία d , τότε η εκ των υστέρων πιθανότητα της υπόθεσης h δεδομένων των στοιχείων d δίνεται από τον τύπο

$$P(h|d) = \frac{P(d|h)P(h)}{P(d)}.$$

Ο τύπος αυτός συνδέει τη ζητούμενη πιθανότητα (για την υποθετική διεργασία) με την εκ των προτέρων πιθανότητα της υπόθεσης $p(h)$, την *πιθανοφάνεια* (likelihood) των παρατηρούμενων στοιχείων (πιθανότητα των στοιχείων δεδομένης της υπόθεσης) και έναν παράγοντα κανονικοποίησης που έχει να κάνει με τη συνολική (αδέσμευτη) πιθανότητα των παρατηρούμενων στοιχείων. Η συνολική πιθανότητα των στοιχείων ισούται με το σταθμισμένο άθροισμα των δεσμευμένων πιθανοτήτων τους για όλες τις δυνατές εναλλακτικές υποθέσεις h' (ανάλογα με την πιθανότητα καθεμιάς), δηλαδή $P(d) = \sum_{h'} P(d|h')P(h')$. Με άλλα λόγια, η συνολική (περιθωριακή, marginal) πιθανότητα των στοιχείων d είναι το άθροισμα της πιθανότητας να παρατηρηθούν με δεδομένη κάθε δυνατή υπόθεση h' επί την πιθανότητα να ισχύει η εκάστοτε υπόθεση.

Για παράδειγμα, ας υποθέσουμε ότι περπατάμε στο δρόμο και ακούμε ένα σφύριγμα (ας το συμβολίσουμε ϕ). Πόσο πιθανό είναι να βρίσκεται ο φίλος μας Φ πίσω μας, με δεδομένο ότι ακούστηκε αυτό το σφύριγμα ϕ ; Η απάντηση $P(\Phi|\phi)$ μπορεί να υπολογιστεί αν γνωρίζουμε (α) πόσο πιθανό είναι γενικά να βρίσκεται εκεί ο Φ , (β) πόσο πιθανό είναι να σφυρίξει ο Φ , και (γ) πόσο συχνά ακούγονται σφυρίγματα εκεί που βρισκόμαστε. Αν ο Φ μένει σε αυτή τη γειτονιά, η εκ των προτέρων πιθανότητα $P(\Phi)$ είναι μεγάλη, αν όμως ζει στο εξωτερικό είναι πολύ μικρή, και αυτό πρέπει να επηρεάσει το κατά πόσο

²⁵Βασισμένο στο Griffiths, T. L., Kemp, C., & Tenenbaum, J. B. (2008). Bayesian models of cognition. Στο R. Sun (Επιμ.), *The Cambridge handbook of computational psychology* (σελ. 59–100). Cambridge University Press.

θα σκεφτούμε τον Φ ακούγοντας το σφύριγμα. Αν ο Φ είναι τύπος που σφυρίζει στους γνωστούς, τότε η πιθανοφάνεια $P(\phi|\Phi)$ είναι μεγάλη, αν όμως δεν συνηθίζει να σφυρίζει τότε το άκουσμα του σφυρίγματος δεν πρέπει να πάει τη σκέψη μας σε αυτόν. Τέλος, αν είμαστε σε ένα μέρος όπου δεν ακούγονται σφυρίγματα, η παρατήρηση ενός σφυρίγματος είναι «διαγνωστική», με την έννοια ότι σηματοδοτεί κάτι ιδιαίτερο, το οποίο χρειάζεται μια επεξηγηματική υπόθεση για να δικαιολογηθεί (για παράδειγμα, την υπόθεση ότι βρίσκεται εκεί ένας φίλος που σφυρίζει). Αν όμως βρισκόμαστε σε μια γειτονιά όπου συχνά οι περίοικοι σφυρίζουν ο ένας στον άλλον (μεγάλη πιθανότητα $p(\phi)$ να παρατηρηθεί σφύριγμα ούτως ή άλλως) τότε το άκουσμα ενός σφυρίγματος συνήθως δεν σημαίνει τίποτα και δεν απαιτεί ιδιαίτερη επεξηγηματική υπόθεση.

Ένα πολύ σημαντικό σημείο της πιθανοκρατικής προσέγγισης είναι ότι μιλάμε πάντοτε για «αναθεώρηση» πιθανοτήτων (και όχι για γενικό υπολογισμό) καθώς ξεκινάμε υποχρεωτικά από μια αρχική κατανομή (την εκ των προτέρων πιθανότητα) και από αυτήν περνάμε σε μια νέα κατανομή (την εκ των υστέρων πιθανότητα). Αρχικά, πριν λάβουμε γνώση των παρατηρήσεων, δίνεται στην υπόθεση h μια πιθανότητα που αντιστοιχεί στο βαθμό πεποίθησής μας γι' αυτήν. Μπορεί να είναι μια «μη πληροφοριακή» εκ των προτέρων πιθανότητα (non-informative prior) αν η συγκεκριμένη υπόθεση δεν θεωρείται πιθανότερη από άλλες εναλλακτικές. Στην περίπτωση αυτή μπορεί οι εκ των προτέρων πιθανότητες όλων των εναλλακτικών να τίθενται ίσες μεταξύ τους. Ή μπορεί για κάποιους λόγους ορισμένες εναλλακτικές να είναι εκ των προτέρων πιθανότερες από άλλες. Αυτό αντικατοπτρίζει την αρχική γνώση με την οποία το σύστημα προσεγγίζει το πεδίο του προβλήματος. Εάν η φύση του προβλήματος είναι τέτοια που δεν υπάρχουν άλλες εναλλακτικές διεργασίες αλλά η μόνη εναλλακτική είναι να μην ισχύει η υπόθεση h τότε οι εκ των προτέρων πιθανότητες είναι οι $P(h)$ και $P(\neg h) = 1 - P(h)$. Σε κάθε περίπτωση η υπόθεση h δεν στέκεται από μόνη της στο πρόβλημα αλλά τίθεται ως μία εναλλακτική. Με άλλα λόγια, η διατύπωση ενός οποιουδήποτε προβλήματος σε αυτό το πιθανοκρατικό πλαίσιο απαιτεί τη ρητή και εξαντλητική απαρίθμηση όλων των δυνατών εναλλακτικών, προκειμένου να υπολογιστούν οι πιθανότητες που αντιστοιχούν σε καθεμιά (τόσο οι εκ των προτέρων πιθανότητες κάθε υπόθεσης όσο και η πιθανοφάνεια των παρατηρήσεων με δεδομένη την κάθε υπόθεση).

Η εκ των προτέρων πιθανότητα, η οποία είναι υποχρεωτική για την πιθανοκρατική προσέγγιση οποιουδήποτε προβλήματος σε μπειζιανό πλαίσιο, αποτελεί πολύ σημαντικό παράγοντα για την ορθολογική αναθεώρηση των πεποιθήσεων. Θεωρητικά, ένα σύστημα τέτοιου είδους δεν «γεννά» νέες υποθέσεις, καθώς δεν υπάρχει διαδικασία κατασκευής υποθέσεων. Όλες οι δυνατές υποθέσεις πρέπει με κάποιον τρόπο να προϋπάρχουν μέσα στο σύστημα. Και, καθώς οι αναπαραστάσεις σε τέτοιου είδους συστήματα δεν είναι οντότητες ούτε σημειακές εκτιμήσεις αλλά κατανομές πιθανοτήτων, η ύπαρξη μιας εναλλακτικής υπόθεσης ανάγεται στην ύπαρξη μιας εκ των προτέρων πιθανότητας για την υπόθεση αυτή. Ουσιαστικά, η ύπαρξη της πιθανότητας ταυτίζεται με την αναπαράσταση της υπόθεσης. Από πρακτική σκοπιά, αυτό καθιστά την απαίτηση για προσδιορισμό εκ των προτέρων πιθανοτήτων ευάλωτο σημείο της προσέγγισης, καθώς δεν είναι πάντοτε προφανές ποιες

πρέπει να είναι αυτές οι εκ των προτέρων πιθανότητες και με ποια κριτήρια θα πρέπει να υπολογιστούν ή να ενταχθούν στο σύστημα. Στην πράξη, επειδή τα προβλήματα που εξετάζονται υπό το πρίσμα της πιθανοκρατικής προσέγγισης δεν αντιστοιχούν σε καταστάσεις εκ του μηδενός, μπορούν να εφαρμοστούν δύο εύλογες μέθοδοι, με κυμαινόμενο βαθμό επιτυχίας: Η μία είναι η προαναφερθείσα μέθοδος μη πληροφοριακών πρότερων πιθανοτήτων. Η άλλη είναι η μέθοδος εμπειρικών πρότερων πιθανοτήτων (*empirical priors*), κατά την οποία η εκ των προτέρων πιθανότητα συνάγεται από την κατανομή των παρατηρήσεων που προηγούνται της τρέχουσας κατάστασης ή του τρέχοντος προβλήματος.

6.3 Ορθολογικότητα

Στη μπειζιανή προσέγγιση η ορθολογικότητα αποτελεί θεμελιώδη υπόθεση. Χάρη σε αυτή την υπόθεση μπορεί να εφαρμοστεί αδιακρίτως ο κανόνας για την αναθεώρηση των πεποιθήσεων, καθώς η εφαρμογή του κανόνα αποτελεί ορθολογική αξιοποίηση των παρατηρούμενων στοιχείων και εγγυάται βέλτιστους βαθμούς πεποίθησης. Ταυτόχρονα, η ορθολογικότητα αποτελεί βαθιά θεωρητική δέσμευση για τη φύση του νοητικού συστήματος: θεωρείται ότι το σύστημα έχει εξελιχθεί έτσι ώστε να συνδυάζει με βέλτιστο τρόπο τις διαθέσιμες πληροφορίες και να λαμβάνει αποφάσεις που ελαχιστοποιούν το σφάλμα. Η ιδιότητα αυτή είναι πολύ σημαντική μέσα σε ένα περιβάλλον αβεβαιότητας, όπου το κόστος κάθε σφάλματος μπορεί να είναι μεγάλο.

Όμως τι ακριβώς είναι ορθολογικότητα; Με ποια κριτήρια αξιολογείται; Και μπορούμε, πράγματι, να αποδώσουμε την ιδιότητα αυτή στο νοητικό σύστημα; Τα ερωτήματα αυτά έχουν απασχολήσει την ερευνητική κοινότητα από παλιά—σίγουρα πολύ πριν από την επέλαση της πιθανοκρατικής προσέγγισης και την εφαρμογή μπειζιανών μοντέλων στην προσομοίωση της νοητικής λειτουργίας.²⁶ Λέγοντας ορθολογικότητα (*rationality*) εννοούμε, πολύ γενικά, την ιδιότητα του συστήματος να δρα σε συνέπεια με τους στόχους του. Δηλαδή οι αποφάσεις και οι πράξεις του να είναι τέτοιες που να οδηγούν στο μέγιστο δυνατό βαθμό προς την προώθηση των προτεραιοτήτων του και την επίτευξη των στόχων του, όπως αυτές έχουν οριστεί και αναπαρίστανται στο ίδιο το σύστημα. Με άλλα λόγια, η ορθολογικότητα αντιστοιχεί στη λήψη των «σωστών» αποφάσεων. Όμως το τι συνιστά σωστή απόφαση εξαρτάται από τις πληροφορίες που είναι διαθέσιμες στο σύστημα και από το ιστορικό λειτουργίας του συστήματος (ή των εξελικτικών του προγόνων).

Μια ερευνητική προσέγγιση που κυριάρχησε στη μελέτη της νόησης πριν λίγες δεκαετίες εστιάζει στην κανονιστική (*normative*) ορθολογικότητα. Σύμφωνα με αυτή την προσέγγιση, το σωστό και το λάθος ορίζονται έξωθεν, με βάση κάποιο απόλυτο (εκ των προτέρων) πλαίσιο όπως η τυπική λογική ή η μαθηματική θεωρία των πιθανοτήτων. Από τέτοια πλαίσια μπορεί να υπολογιστεί ποια είναι κάθε φορά η σωστή απάντηση. Με άλλα λόγια, μια κανονιστική προσέγγιση ορίζει το πώς *πρέπει* να λειτουργεί ένα σύστημα, πα-

²⁶Βλ. σχετική συζήτηση στο Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23, 645–726.

ρέχοντας ένα κριτήριο ορθότητας. Όταν το νοητικό σύστημα λαμβάνει μια απόφαση που δεν ταυτίζεται με εκείνη της λογικής ή των μαθηματικών τότε θεωρείται ότι δρα ανορθολογικά, καθώς η απόφαση είναι—με βάση το απόλυτο εξωτερικό κριτήριο—λανθασμένη. Οι λανθασμένες αποφάσεις είναι με κάποιον τρόπο ασυνεπείς, είτε με την πραγματικότητα του περιβάλλοντος είτε με τις προτεραιότητες και τους στόχους του ίδιου του συστήματος. Το πλήθος των ερευνητικών ευρημάτων που δείχνουν ότι η ανθρώπινη νόηση δεν συνάδει με τις επιταγές της λογικής και των μαθηματικών²⁷ υπαγορεύει το συμπέρασμα ότι ο νους δεν είναι ορθολογικός και για το λόγο αυτό οι άνθρωποι δεν λειτουργούν με βέλτιστο (για τους εαυτούς τους) τρόπο. Το συμπέρασμα αυτό δεν έρχεται σε αντίθεση με την εξελικτική προέλευση του νου, καθώς η εξέλιξη δεν εγγυάται ούτε ορθολογικότητα ούτε βέλτιστη λειτουργία, παρά μόνο αρμοστικότητα σε συγκεκριμένο περιβάλλον, οριζόμενη σε σχέση με τις πιέσεις του περιβάλλοντος, συγκριτικά με άλλα μέλη του ίδιου είδους.

Η θέση της ανορθολογικότητας του νου έχει βρει σθεναρή αντίσταση από άλλους ερευνητές που αντιτίθενται στην κανονιστική προσέγγιση και υποστηρίζουν την εναλλακτική προσέγγιση της περιγραφικής (descriptive) ορθολογικότητας.²⁸ Ουσιαστικά θεωρούν εκ των προτέρων ότι η νόηση λειτουργεί ορθολογικά, και επιχειρούν να δικαιολογήσουν τις αποφάσεις που λαμβάνονται σε κάθε περίπτωση με βάση το εξελικτικό ιστορικό της νόησης και τις συγκεκριμένες συνθήκες μέσα στις οποίες λειτουργεί. Η ερευνητική αυτή κατεύθυνση έχει πετύχει σε αρκετές περιπτώσεις να ανατρέψει τα συμπεράσματα της κανονιστικής προσέγγισης, δείχνοντας ότι οι κανονιστικά «λανθασμένες» επιλογές μπορεί να αποτελούν βέλτιστες αποφάσεις μέσα στο συγκεκριμένο (ιστορικό και περιβαλλοντικό) πλαίσιο στο οποίο λαμβάνονται, διότι αποτελούν προσαρμογές σε συγκεκριμένες συνθήκες (ενώ οι τυπικές λύσεις που δίνουν τα μαθηματικά και η λογική είναι εξ ορισμού τυφλές στο περιεχόμενο και συνεπώς δεν μπορούν να προσαρμοστούν σε διαφορετικές περιστάσεις). Από θεωρητική σκοπιά, η προσέγγιση της περιγραφικής ορθολογικότητας θεωρεί ότι το κριτήριο δεν μπορεί να είναι εξωτερικό και απόλυτο αλλά να βασίζεται στην ίδια τη νοητική λειτουργία, δεδομένου ότι αυτή έχει αποδεδειγμένα οδηγήσει σε εξελικτική επιτυχία και σε οργανισμούς (άρα και νοητικές λειτουργίες) καλά προσαρμοσμένους στο περιβάλλον τους. Με άλλα λόγια, «ορθολογικό» ορίζεται αυτό που τείνει να κάνει ο νους στη φυσιολογική λειτουργία του. Ο προσδιορισμός του τι είναι «σωστό» και τι «λάθος» θα πρέπει να βασιστεί στη μελέτη της ίδιας της νοητικής λειτουργίας (εξ ου και ο όρος «περιγραφική»). Υπό αυτό το πρίσμα, μη ορθολογική είναι η νόηση όταν αποκλίνει από εκείνη της πλειονότητας.

Από πρακτική σκοπιά, η αξιοποίηση της υπόθεσης της ορθολογικότητας στην πιθανοκρατική προσέγγιση επιτρέπει τη διερεύνηση της πληροφοριακής διαθεσιμότητας στη νοητική λειτουργία. Έχοντας υποθέσει ότι η νόηση λαμβάνει αποφάσεις συνδυάζοντας πάντοτε τα δεδομένα με βέλτιστο τρόπο, μπορούμε να διερευνήσουμε ποια πρέπει να είναι τα διαθέσιμα δεδομένα προκειμένου να παραχθούν οι αποφάσεις που λαμβάνουν οι άνθρωποι. Ας υποθέσουμε ότι σε ένα έργο οι συμμετέχοντες δίνουν τις απαντήσεις Α και

²⁷Βλ. Kahneman, D. (2011). *Thinking, fast and slow*. Penguin.

²⁸Π.χ., Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world*. Oxford University Press.

Β με μια συγκεκριμένη αναλογία, ανάλογα με τις συνθήκες του προβλήματος. Αν ένα πιθανοκρατικό μοντέλο δίνει διαφορετική αναλογία απαντήσεων, τότε αυτό σημαίνει ότι τα δεδομένα που επεξεργάζεται το μοντέλο δεν ταυτίζονται με τα δεδομένα που είναι διαθέσιμα (ή που λαμβάνονται υπόψη) κατά τη νοητική λειτουργία. Αν ένα μοντέλο δίνει περισσότερες (ή λιγότερες) κανονιστικά ορθές απαντήσεις από τους ανθρώπους, αυτό σημαίνει ότι το μοντέλο έχει περισσότερες (ή λιγότερες) πληροφορίες από το νου. Η καλή προσαρμογή του μοντέλου σε μεγάλο εύρος συνθηκών μπορεί να επιτευχθεί μόνο αν οι πληροφορίες τις οποίες συνδυάζει βέλτιστα είναι οι ίδιες με τις πληροφορίες που έχει στη διάθεσή του ο νους. Έτσι, μπορούμε να εξετάσουμε πειραματικά, με αξιοποίηση κατάλληλων υπολογιστικών μοντέλων, ποιες ακριβώς πληροφορίες διαθέτει ο νους σε κάθε περίπτωση και ποιες επιστρατεύει για τη λύση κάθε προβλήματος. Εν συνεχεία μπορεί να εξεταστεί πώς η νόηση αποκτά πρόσβαση στις πληροφορίες αυτές, πώς τις διαχειρίζεται, πόσο μπορεί να τις συγκρατήσει, πώς αυτό επηρεάζεται από τις περιστάσεις του προβλήματος κλπ. Με άλλα λόγια, η υπόθεση της ορθολογικής λειτουργίας μεταθέτει το πρόβλημα της κανονιστικής ανορθολογικότητας από το μηχανισμό στη διαθεσιμότητα των πληροφοριών.

6.4 Επίπεδο ανάλυσης

Ένα ζήτημα που έχει απασχολήσει πολύ την ερευνητική κοινότητα²⁹ έχει να κάνει με το επίπεδο ανάλυσης στο οποίο αντιστοιχούν οι περιγραφές των πιθανοκρατικών μοντέλων. Όπως είναι προφανές από τις παραπάνω παρατηρήσεις, τα μπειζιανά μοντέλα δεν προσομοιώνουν νοητικούς μηχανισμούς αλλά παράγουν τις προβλέψεις τους με βάσει γενικότερες αρχές λειτουργίας όπως είναι η ορθολογικότητα. Συνεπώς δεν αποτελούν μοντέλα στο (κατά Marr) «αλγοριθμικό» επίπεδο περιγραφής, όπως τα συμβολικά ή τα συνδυαστικά, αλλά εστιάζουν απευθείας στο «υπολογιστικό» επίπεδο περιγραφής, παράγοντας προβλέψεις για τα αποτελέσματα των νοητικών διεργασιών.

Η ιδιότητα αυτή έχει οδηγήσει σε σημαντικές επικρίσεις από τους ερευνητές που θεωρούν ότι το αντικείμενο της γνωσιακής επιστήμης (και άρα ο στόχος της μελέτης) είναι η περιγραφή νοητικών μηχανισμών. Από τη σκοπιά αυτή τα μπειζιανά μοντέλα δεν μας δίνουν κάποια χρήσιμη πληροφορία για την κατανόηση των νοητικών φαινομένων, εφόσον δεν μπορούν να εξηγήσουν ούτε πώς αποκτά το σύστημα τις πληροφορίες από το περιβάλλον ούτε με ποιον τρόπο τις επεξεργάζεται προκειμένου να καταλήξει στις αποφάσεις του. Οι υποστηρικτές της πιθανοκρατικής προσέγγισης, αντίθετα, θεωρούν ότι γνωσιακή επιστήμη δεν μπορεί να περιοριστεί στην περιγραφή των μηχανισμών, ειδικά όταν δεν είναι προηγουμένως σαφές ποιο είναι το αντικείμενο και το αποτέλεσμα των υπολογισμών. Η μπειζιανή προσέγγιση μπορεί να αποσαφηνίσει τόσο τους σκοπούς όσο και τις βασικές αρχές της νοητικής λειτουργίας, στο βαθμό που μπορεί να ελέγξει (α) κατά πόσο τα νοητι-

²⁹Βλ. σχετική συζήτηση στο Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences*, 34, 169–231.

κά συστήματα όντως συμπεριφέρονται ως βελτιστοποιητές και (β) ποιες πληροφορίες είναι διαθέσιμες για την ανταπόκριση σε κάθε συγκεκριμένη περίπτωση. Άλλες προσεγγίσεις θα πρέπει στη συνέχεια να εξετάσουν με ποιον τρόπο το νοητικό (ή το νευρικό) σύστημα υλοποιεί τις αρχές αυτές και παράγει τις παρατηρούμενες αποκρίσεις. Συνεπώς, από τη σκοπιά αυτή η μπειζιανή προσέγγιση δεν αποτελεί άμεσο «αντίπαλο» άλλων προσεγγίσεων αλλά μάλλον μια συμπληρωματική κατεύθυνση η οποία δίνει πληροφορίες για τη φύση και το στόχο των υπολογισμών, καθώς και για τη διαθεσιμότητα των αντίστοιχων πληροφοριών.

7 Εξελικτική ψυχολογία

Μια προσέγγιση στην περιγραφή του νου προέρχεται από την εξελικτική θεώρηση των νοητικών λειτουργιών. Λέγοντας «εξελικτική θεώρηση» εννοούμε υπό το πρίσμα της δαρβινικής εξέλιξης του είδους, με την παραδοχή (ή υπόθεση) ότι οι νοητικές λειτουργίες αποτελούν στοιχεία του οργανισμού που έχουν γίνει χαρακτηριστικά γνωρίσματα του ανθρώπινου είδους επειδή αποδείχτηκαν αποτελεσματικά στην αντιμετώπιση ή επίλυση συγκεκριμένων προβλημάτων επιβίωσης και αναπαραγωγής κατά το διάστημα διαμόρφωσης του είδους.³⁰ Ερμηνεύοντας το νοητικό δυναμικό στα πλαίσια του οικολογικού περιβάλλοντος στο οποίο υποθέτουμε ότι εξελίχθηκε, η εξελικτική θεώρηση έχει κερδίσει μια κεντρική θέση στις συζητήσεις για την αρχιτεκτονική του νου. Προσεγγίζοντας το αντίστοιχο σκεπτικό των ζωολόγων, που αναλύουν τις συμπεριφορές των ζώων με αναφορά σε υποθετικούς σκοπούς τους οποίους αυτές εξυπηρετούν διαχρονικά, η εξελικτική ψυχολογία επιχειρεί να εντοπίσει τις κύριες συνιστώσες της νοητικής λειτουργίας σε αντιστοιχία με εξελικτικές πιέσεις που οδήγησαν στην εμφάνισή τους. Έτσι, κεντρικό ρόλο στη θεωρητική συγκρότηση παίζει η περιγραφή των περιβαλλοντικών απαιτήσεων και περιορισμών, και των ατομικών επιδιώξεων, που οδήγησαν στη διαμόρφωση νοητικών μηχανισμών, και κατά συνέπεια συμπεριφορών, για την αντιμετώπισή τους.

7.1 Εξέλιξη και φυσική επιλογή

Εξέλιξη ονομάζεται η αλλαγή των χαρακτηριστικών (μορφή και συμπεριφορά) των οργανισμών από γενιά σε γενιά. Στη γενετική, εξέλιξη είναι η φυλογονική μεταβολή των γονοτύπων,³¹ δηλαδή η αλλαγή στη σύσταση και κατανομή των γονιδιακών συστατικών

³⁰Ο όρος «εξελικτική ψυχολογία» αποδίδει το αγγλικό evolutionary psychology και αναφέρεται στην εφαρμογή της θεωρίας της εξέλιξης (evolution) στις νοητικές λειτουργίες, εστιάζοντας στη *φυλογένεση* (phylogeny), δηλαδή τη διαμόρφωση του είδους. Δεν έχει σχέση με την αναπτυξιακή ψυχολογία (developmental psychology), η οποία αναφέρεται στην ανάπτυξη των παιδιών, ή καλύτερα του ανθρώπινου νου από τη γέννηση ως την ενηλικίωση, εστιάζοντας στην οντογένεση (ontogeny). Σε παλαιότερα ελληνικά εγχειρίδια η αναπτυξιακή ψυχολογία ονομάζεται «εξελικτική» και η ανάπτυξη των παιδιών «εξέλιξη», οδηγώντας σε σύγχυση.

³¹*Γονότυπος* ονομάζεται η γενετική σύσταση ενός οργανισμού (ατόμου), δηλαδή η σύσταση του γενετικού υλικού (DNA). *Φαινότυπος* ονομάζεται το σύνολο των χαρακτηριστικών ενός οργανισμού στη μορφή.

από έναν πληθυσμό στον απόγονό του. Πολύ χοντρικά, όταν αλλάζουν τα γονίδια ή οι σχετικές αναλογίες τους, υπάρχει εξέλιξη. Κάθε μακροσκοπική μεταβολή στην κατανομή διαφορετικών γονοτύπων σε έναν πληθυσμό αποτελεί εξέλιξη, ανεξαρτήτως αιτίας και αποτελέσματος. Η εξέλιξη περιλαμβάνει διαδικασίες αύξησης και μείωσης της ποικιλίας. Η πιο γνωστή περίπτωση μεταβολής είναι η μετάλλαξη (mutation), που δημιουργεί ένα νέο γονότυπο και αυξάνει την ποικιλία. Ακολουθώντας, η φυσική επιλογή (natural selection) του καταλληλότερα προσαρμοσμένου στο περιβάλλον αφαιρεί, λόγω ανταγωνισμού, τις λιγότερο αρμοστικές εκδοχές και μειώνει την ποικιλία. Η περίπτωση αυτή δεν είναι η μοναδική ούτε και η πιο κοινή. Στους σεξουαλικά αναπαραγόμενους οργανισμούς η πιο συνηθισμένη διεργασία αύξησης της γονοτυπικής ποικιλίας είναι ο ανασυνδυασμός (recombination) του γενετικού υλικού στους γαμέτες κατά τη μείωση (κυτταρική διαίρεση). Άλλες διεργασίες αύξησης της ποικιλίας είναι η γονιδιακή ροή (gene flow) και η οριζόντια μεταφορά (horizontal transfer). Οι διεργασίες μείωσης της ποικιλίας περιλαμβάνουν, εκτός από τη φυσική επιλογή, την τυχαία ολίσθηση (random drift) και άλλες ειδικές περιπτώσεις.

Η *φυσική επιλογή* είναι ένα φαινόμενο που συμβαίνει υποχρεωτικά όποτε συνυπάρχουν οι εξής προϋποθέσεις: Πρώτον, ένας αναπαραγόμενος πληθυσμός, φυσικών οργανισμών ή τεχνητών οντοτήτων, στον οποίο υπάρχει κάποια ποικιλομορφία (διακύμανση), δηλαδή δεν είναι όλα τα άτομα πανομοιότυπα. Δεύτερον, κληρονομικότητα, δηλαδή συστηματική διατήρηση της ποικιλομορφίας από γενιά σε γενιά. Πιο απλά, αυτό σημαίνει ότι οι απόγονοι είναι κατά μέσο όρο πιο όμοιοι στους προγόνους («γονείς») τους παρά σε άλλα άτομα της προηγούμενης γενιάς. Τρίτον, περιορισμένοι πόροι κάποιας μορφής, έτσι ώστε να μην είναι δυνατό να επιβιώσουν και να αναπαραχθούν όλοι οι απόγονοι. Αυτό μπορεί να σημαίνει περιορισμένη τροφή στο περιβάλλον ή περιορισμένες ευκαιρίες αναπαραγωγής ή οποιαδήποτε άλλη διαφορά που τελικά να έχει επιπτώσεις στην αναπαραγωγή. Οποιοδήποτε συνεμφανίζονται αυτές οι ιδιότητες (αναπαραγωγή-ποικιλία-κληρονομικότητα-περιορισμός πόρων) τότε υποχρεωτικά θα παρατηρείται φυσική επιλογή, δηλαδή αύξηση της ομοιογένειας του πληθυσμού έτσι ώστε τα χαρακτηριστικά (traits) των απογόνων να είναι μεταξύ τους πιο όμοια και καλύτερα ταιριασμένα στο περιβάλλον σε σύγκριση με τους προγόνους τους. (Λέγοντας «περιβάλλον» εννοούμε και τα άλλα άτομα του πληθυσμού.) Αυτό συμβαίνει διότι οι περιορισμένοι πόροι έχουν ως άμεση επίπτωση τη διαφορική επιβίωση και αναπαραγωγή—για παράδειγμα, αν δεν υπάρχει αρκετή τροφή για όλους τότε κάποιοι υποχρεωτικά θα πεθάνουν από ασιτία.

Οποιοδήποτε χαρακτηριστικό των οργανισμών επιδρά διαφορικά στην αξιοποίηση των πόρων μπορεί να έχει επιπτώσεις στην αναπαραγωγή (λ.χ. ανάγκη για λιγότερη τροφή λόγω πιο αποδοτικού μεταβολισμού, σωματική δύναμη που αποκλείει βίαια ανταγωνιστές,

τη φυσιολογία, τη χημεία και τη συμπεριφορά. Στην εξελικτική βιολογία συνήθως αναφερόμαστε σε συγκεκριμένη περιοχή του γονιδιώματος και αντίστοιχα, σε μεμονωμένα μορφολογικά ή άλλα φαινοτυπικά χαρακτηριστικά, και όχι απαραίτητα στο σύνολο. Η σημαντική διάκριση είναι ότι ο γονότυπος καθορίζεται κατά τη σύλληψη και προϋπάρχει της ανάπτυξης, ενώ ο φαινότυπος είναι αποτέλεσμα της ανάπτυξης μέσα σε συγκεκριμένο περιβάλλον, με δεδομένες τις «προδιαγραφές» ή «οδηγίες» που καθορίζει ο γονότυπος. Υπό την έννοια αυτή, ο γονότυπος θεωρείται πληροφορία και ο φαινότυπος αποτελεί τη φυσική έκφρασή του.

πιο αποδοτική αναζήτηση τροφής λόγω ευφυΐας ή οξύτερων αισθήσεων κλπ.). Λόγω της κληρονομικότητας, οι απόγονοι των πιο αποδοτικών ατόμων θα έχουν κατά πάσα πιθανότητα τα χαρακτηριστικά που οδήγησαν στην αυξημένη αναπαραγωγή των προγόνων τους. Άρα η σύνθεση του πληθυσμού θα αλλάζει διαρκώς, «επιλέγοντας» τα χαρακτηριστικά εκείνα που έχουν ως αποτέλεσμα συγκριτικά καλύτερη αναπαραγωγή, τα οποία ονομάζονται *αρμοστικά* (adaptive). Αρμοστικό είναι κάθε φαινοτυπικό γνώρισμα που έχει επικρατήσει σε έναν πληθυσμό ως αποτέλεσμα επιλογής, επειδή συνιστά λειτουργικό πλεονέκτημα, δηλαδή επειδή η ύπαρξή του ευνοεί με συγκεκριμένο τρόπο τον κάτοχό του και αυτό έχει άμεση ή έμμεση θετική επίδραση στην αναπαραγωγική του αποδοτικότητα. Ταυτόχρονα, μπορεί να κληρονομούνται και ουδέτερα ή και ανεπιθύμητα χαρακτηριστικά, αν είναι με κάποιον τρόπο συνδεδεμένα με εκείνα που ευνοούν την αναπαραγωγή, εφόσον η τυχόν αρνητική επίδρασή τους δεν ξεπερνά το όφελος των ευνοϊκών χαρακτηριστικών. Προσέξτε ότι δεν υπάρχει σχέδιο ή προκαθορισμένη κατεύθυνση, ούτε κάποιος ειδικός μηχανισμός επιλογής, καθώς η μεταβολή στην κατανομή των χαρακτηριστικών είναι αποτέλεσμα της σύμπτωσης συγκεκριμένων συνθηκών. Η θεώρηση αυτή μπορεί να δικαιολογήσει την εμφάνιση εξαιρετικά σύνθετων φυσικών γνωρισμάτων καθώς και περίπλοκων συμπεριφορών, πολύ καλά ταιριασμένων στο περιβάλλον, που δίνουν την εντύπωση πρόθεσης ή σχεδίασης. Μπορεί επίσης να δικαιολογήσει τη διατήρηση μη βέλτιστων χαρακτηριστικών, μέσω της κατανόησης των αναπτυξιακών μηχανισμών που παράγουν τους φαινότυπους με διαρκή αλληλεπίδραση μεταξύ γονοτύπου και περιβάλλοντος.

Η φυσική επιλογή είναι απαραίτητη για την εξήγηση της αρμοστικότητας (fitness) των οργανισμών, δηλαδή του παρατηρούμενου ταιριάσματος μεταξύ των φαινοτυπικών χαρακτηριστικών των οργανισμών και του περιβάλλοντος και του τρόπου ζωής τους. Η αρμοστικότητα δεν είναι εκ των προτέρων δεδομένη ιδιότητα επιμέρους χαρακτηριστικών. Μπορεί να εξεταστεί η ισχύς της για κάθε χαρακτηριστικό βάσει συγκεκριμένων υποθέσεων. Αν ένα χαρακτηριστικό φαίνεται ιδιαίτερα αρμοστικό και κατάλληλο για συγκεκριμένη λειτουργία (π.χ. το μάτι για οπτική αντίληψη, το φτερό για πτήση κλπ.) τότε αναζητούνται εξελικτικές εξηγήσεις βάσει φυσικής επιλογής για την απόκτηση και διατήρησή του στο είδος, λαμβάνοντας υπόψη και τυχόν κόστος στην ανάπτυξη και διατήρηση της λειτουργίας του. Υπάρχουν σημαντικοί περιορισμοί στη φυσική επιλογή, και στη δυνατότητα εξέλιξης διαρκώς πιο αρμοστικών οργανισμών. Τέτοιοι είναι οι αναπτυξιακοί και εμβρυολογικοί περιορισμοί (τι είναι εφικτό αποτέλεσμα αλληλεπίδρασης μεταξύ γονοτύπου, περιβάλλοντος και φυσικοχημικών συνθηκών), ιστορικοί-φυλογονικοί περιορισμοί (ποιο είναι το διαθέσιμο γονοτυπικό υπόστρωμα της προηγούμενης γενιάς, πάνω στο οποίο διαμορφώνεται ένα νέο χαρακτηριστικό), αλληλεπιδράσεις μεταξύ γονιδίων (αν υπάρχουν ανταγωνιστικές τάσεις ή εμπόδια από άλλα φαινοτυπικά χαρακτηριστικά), και περιορισμένη ποικιλία στον πληθυσμό (υπερβολικά στενό γονοτυπικό δείγμα λόγω προηγούμενης επιλογής). Θα πρέπει λοιπόν να κάνουμε μια σαφή διάκριση μεταξύ εξέλιξης (που είναι η αλλαγή από γενιά σε γενιά) και φυσικής επιλογής (που είναι ένας μηχανισμός αλλαγής).

7.2 Εμφυτότητα και περιβάλλον

Μια κλασική διαμάχη στο χώρο της γνωσιακής επιστήμης αφορά στην εμφυτότητα (innateness) των χαρακτηριστικών, η οποία θεωρούνταν παλιότερα ως αντίθετη στην ιδέα της μάθησης ή της περιβαλλοντικής επίδρασης γενικότερα. Έτσι, διαμορφώθηκαν δύο φιλοσοφικά «στρατόπεδα», οι εμφυτοκράτες (nativists), που θεωρούσαν ότι η νόηση είναι κατά βάση γενετικά προκαθορισμένη, και οι εμπειριοκράτες (empiricists), που θεωρούσαν ότι η νόηση είναι κατά βάση αποτέλεσμα αλληλεπίδρασης με το περιβάλλον, το οποίο μέσω των εμπειριών διαμορφώνει τόσο τη γνώση όσο και τους μηχανισμούς επεξεργασίας της. Η διαμάχη αυτή πέρασε στη γνωσιακή επιστήμη (βλ. Ενότητα 11.3, σελ. 108) οριοθετώντας σε ένα βαθμό και τη διάκριση μεταξύ συμβολισμού, που απαιτεί έμφυτο κεντρικό επεξεργαστή και προκαθορισμένες διεργασίες, και συνδετισμού, που βασίζεται στη ρύθμιση των βαρών των συνδέσεων μέσα από την επεξεργασία συγκεκριμένων δειγμάτων (περιβαλλοντικών εισροών). Όπως συμβαίνει καμιά φορά με τα αναπάντητα φιλοσοφικά ερωτήματα, η απουσία απάντησης οφείλεται σε λάθος ερώτηση, η οποία διαλύεται με την καλύτερη επιστημονική κατανόηση. Στην προκειμένη περίπτωση η αντίθεση μεταξύ εμφυτοκρατίας και εμπειριοκρατίας εξαφανίζεται λόγω της αναγνώρισης της διαρκούς αλληλεπίδρασης και αλληλεξάρτησης μεταξύ γονιδιώματος και περιβάλλοντος ανάπτυξης, κάτι που γίνεται σαφές από τη μελέτη της σύγχρονης αναπτυξιακής βιολογίας.

Σε γενικές γραμμές, η σύγχρονη εικόνα δείχνει ότι η ανάπτυξη με βάση το γονιδίωμα δεν συνίσταται στην εκτέλεση μιας προδιαγεγραμμένης σειράς βημάτων ανεξάρτητα από το περιβάλλον και τις συνθήκες. Όπως και σε κάθε συνταγή, που χρειάζεται να υπάρχουν παράλληλα με τα κατάλληλα υλικά και ο εξοπλισμός και οι συνθήκες (θερμοκρασία, υγρασία, πίεση—δεν φτιάχνονται κουραμπιέδες στο διάστημα!), πολύ περισσότερο στην ανάπτυξη των οργανισμών η έκφραση του γονιδιώματος προϋποθέτει ένα κατάλληλα οργανωμένο χημικό περιβάλλον (αρχικά το ζυγωτό κύτταρο και όλο το ειδικό περιβάλλον γύρω από αυτό, π.χ. η μήτρα στα θηλαστικά) με τα απαραίτητα συστατικά. Η ανάπτυξη των οργανισμών γίνεται σε ένα εγγυημένο σταθερό περιβάλλον (βαρύτητας, θερμοκρασίας), το οποίο για τα κοινωνικά είδη συμπεριλαμβάνει και άλλα άτομα του ίδιου είδους. Από το χαμηλότερο επίπεδο, αυτό της γενετικής μεταγραφής, όπου απαιτείται η παρουσία ειδικών ρυθμιστικών μορίων για την επιλογή και ενεργοποίηση των κατάλληλων κάθε στιγμή κομματιών του γονιδιώματος, μέχρι το ανώτατο επίπεδο της κοινωνικής αλληλεπίδρασης, όπου απαιτείται η γλωσσική κοινότητα για την ανάπτυξη της γλωσσικής ικανότητας, η ύπαρξη και επίδραση του περιβάλλοντος δεν μπορεί να νοηθεί ως εξωτερική ή επικουρική αλλά ως μέρος της διαδικασίας της ανάπτυξης.

Η εξέλιξη, λοιπόν, δεν διαμορφώνει απλώς γονοτύπους επιλέγοντας φαινοτύπους. Η εξέλιξη διαμορφώνει γονοτύπους επιλέγοντας τα αποτελέσματα αλληλεπιδράσεων μεταξύ των αναπτυσσόμενων φαινοτύπων και των σταθερών χαρακτηριστικών του περιβάλλοντος ανάπτυξης. Με άλλα λόγια, η εξέλιξη οδηγεί σε μια ελεγχόμενη πορεία γονιδιακής έκφρασης για την ανάπτυξη του οργανισμού σε δεδομένο περιβάλλον. Αν αλλάξει κάτι στο περιβάλλον, αλλάζει ο οργανισμός (πιο συχνά, αδυνατεί να αναπτυχθεί βιώσιμα). Η θεώρη-

ση αυτή τονίζει τη σημασία της αναπτυξιακής διαδικασίας στην εξέλιξη και δίνει έμφαση στο ρόλο των εξελικτικών πιέσεων κατά την έκφραση του γονιδιώματος στο πλαίσιο της ανάπτυξης. Η ονομασία αυτής της σύγχρονης θεώρησης είναι εξελικτική αναπτυξιακή βιολογία (evolutionary developmental biology, ή *evo-devo*).

Όπως γίνεται φανερό, η προσέγγιση *evo-devo* ανατρέπει εκ θεμελίων το ερώτημα αν κάποιο χαρακτηριστικό είναι έμφυτο ή περιβαλλοντικά επαγόμενο, διότι δεν νοείται έμφυτο χαρακτηριστικό που να μην απαιτεί αλληλεπίδραση με συγκεκριμένο περιβάλλον για να αναπτυχθεί, και διότι δε νοείται περιβαλλοντική επίδραση που να μη λαμβάνει χώρα σε κάποιο προϋπάρχον υπόβαθρο. Τα πάντα είναι έμφυτα και εμπειρικά ταυτόχρονα, και όσο πιο σύνθετες είναι οι περιβαλλοντικές επιδράσεις τόσο πιο σύνθετο πρέπει να είναι το υπόβαθρο που θα τις επεξεργαστεί προκειμένου να τις ενσωματώσει. Καμία μάθηση δεν μπορεί να λάβει χώρα αν δεν υπάρχει μηχανισμός μάθησης. Εφόσον η επίδραση του περιβάλλοντος συμβάλλει διαρκώς στον καθορισμό του οργανισμού, ο αναπτυσσόμενος οργανισμός είναι κάθε στιγμή αποτέλεσμα προηγούμενων αλληλεπιδράσεων γενετικών και περιβαλλοντικών πληροφοριών και μηχανισμών. Άρα η αλληλεπίδραση γονιδιώματος και περιβάλλοντος οδηγεί συνεχώς στο νέο υπόστρωμα που θα αλληλεπιδράσει με το περιβάλλον την επόμενη στιγμή κ.ο.κ., συνδιαμορφώνοντας την ανάπτυξη του οργανισμού σε βαθμό που δεν έχει κανένα νόημα η απόπειρα διάκρισης επιμέρους αιτιακών παραγόντων και απόδοσής τους σε ιδιότητες αποκλειστικά του γονιδιώματος ή του περιβάλλοντος.

Η θεώρηση αυτή έχει σημαντικές συνέπειες για τον τρόπο που συλλαμβάνεται η νόηση στο πλαίσιο της εξελικτικής προσέγγισης, καθώς συνεπάγεται την ύπαρξη «ανθρώπινης φύσης» και ταυτόχρονα την ενσωμάτωση σε αυτήν των χαρακτηριστικών του περιβάλλοντος εξέλιξης. Λέγοντας *ανθρώπινη φύση* εννοούμε σταθερά χαρακτηριστικά του είδους που απαντώνται σε όλα τα άτομα, με μικρές σχετικά διακυμάνσεις. Αυτά δεν περιλαμβάνουν μόνο τη μορφή και τη φυσιολογία αλλά και τη νοητική λειτουργία, καθώς και αυτή υπόκειται σε εξελικτικές πιέσεις όπως και όλες οι υπόλοιπες. Το αποτέλεσμα είναι οι χαρακτηριστικά ανθρώπινες συμπεριφορές και τα καθολικά πολιτισμικά γνωρίσματα των ανθρώπινων κοινωνιών. Υπό το πρίσμα αυτό, το περιθώριο πολιτισμικής διακύμανσης είναι περιορισμένο και δεν υπάρχει σημαντική αυθαιρεσία στα πολιτισμικά γνωρίσματα που να διακρίνει θεμελιωδώς μεταξύ κοινωνιών. Η ανθρώπινη φύση, επειδή εξελίχθηκε μέσω ανάπτυξης σε συγκεκριμένο περιβάλλον, αντικατοπτρίζει τις ιδιότητες του περιβάλλοντος που ήταν σταθερές κατά την εξέλιξή της, με αποτέλεσμα η ανθρώπινη φύση να είναι έμφυτα ταιριασμένη στο εξελικτικό της περιβάλλον.

Με δεδομένη την αναγνώριση του ρόλου της περιβαλλοντικής αλληλεπίδρασης, και επειδή τα βασικά χαρακτηριστικά του περιβάλλοντος παραμένουν σταθερά, μπορούμε να συλλάβουμε την έννοια του έμφυτου ως κάτι που είναι προδιαγεγραμμένο να συμβεί αν δεν μεταβληθούν ριζικά οι συνθήκες. Έτσι, εξουδετερώνεται η ένταση με την εμπειριοκρατία, καθώς οι συνθήκες αποτελούν απαραίτητο πλαίσιο ανάπτυξης των έμφυτων γνωρισμάτων. Έτσι, έμφυτη είναι η μορφή του σώματος, η ύπαρξη των οργάνων κλπ. Λέγοντας ότι ένα γνώρισμα είναι έμφυτο, δεν εννοούμε ότι αυτό είναι παρόν κατά τη γέννηση, διότι πολλά

φαινόμενα και γνωρίσματα αναπτύσσονται αργότερα, όπως π.χ. τα δόντια, η τριχοφυΐα σε ορισμένα σημεία του σώματος, η σεξουαλική ωρίμανση κλπ., τα οποία ουδείς διανοείται να χαρακτηρίσει αποτέλεσμα περιβαλλοντικής επίδρασης (αν και κάποια, όπως η έναρξη της έμμηνης ρύσης, εξαρτώνται εν μέρει και από άλλες συνθήκες όπως είναι η συγκέντρωση λίπους). Υπό την έννοια αυτή, έμφυτη μπορεί να είναι σε μεγάλο βαθμό και η γλωσσική ανάπτυξη, καθώς εξαρτάται από την αλληλεπίδραση με μια γλωσσική κοινότητα, χωρίς όμως αυτό να σημαίνει ότι είναι προκαθορισμένη στο γονιδίωμα οποιαδήποτε συγκεκριμένη δομική περιγραφή της γλώσσας.

Μια σημαντική διάκριση στο πλαίσιο της εξελικτικής θεώρησης είναι η αντίθεση μεταξύ μάθησης και επιλογής, καθώς η έμφαση στις εξελικτικές διεργασίες υποβαθμίζει το ρόλο της εγκεφαλικής ευπλαστότητας και κατ' επέκταση της νοητικής μάθησης. Συγκεκριμένα, η υπόθεση της κυρίαρχης ανθρώπινης φύσης ως σταθερό χαρακτηριστικό του είδους που ανταποκρίνεται σε απαιτήσεις του περιβάλλοντος εξέλιξης σημαίνει ότι οι εξελιγμένοι νοητικοί μηχανισμοί είναι κατά βάση παγιωμένοι, για να εξυπηρετούν τις ανάγκες που οδήγησαν στη διαμόρφωσή τους, και άρα δεν υπόκεινται σε τροποποιήσεις βάσει μάθησης. Αν ο νους ήταν απεριόριστα εύπλαστος και προσαρμοσίμος στις τρέχουσες συνθήκες, μέσω της μάθησης, τότε η εξελικτική προσέγγιση δεν θα προσέφερε κάποιο ιδιαίτερο θεωρητικό πλεονέκτημα. Η ισχύς της εξελικτικής θεώρησης βασίζεται στην παραδοχή ότι δεν υπάρχει μάθηση γενικής φύσης και ότι οι νοητικές λειτουργίες είναι σε μεγάλο βαθμό προκαθορισμένες εξελικτικά. Έτσι, η αλληλεπίδραση κάθε οργανισμού με το περιβάλλον αποτελεί απλώς το αναπτυξιακό κομμάτι της έκφρασης εξελικτικά παγιωμένων τάσεων και χαρακτηριστικών, με συγκεκριμένο και εκ των προτέρων δεδομένο στόχο.

7.3 Λειτουργικός κατακερματισμός

Κατά την πιο συνηθισμένη εκδοχή της, η εξελικτική θεώρηση περιλαμβάνει την υπόθεση ότι ο νους δεν είναι μια ενιαία λειτουργική μονάδα, και δεν πρέπει να αντιμετωπίζεται ως ένας κεντρικός επεξεργαστής αλλά ως μια συλλογή εξειδικευμένων νοητικών εργαλείων. Η υπόθεση αυτή δεν αντιστοιχεί στη λειτουργική διαφοροποίηση μεγάλης κλίμακας που κάνουμε συνήθως στην ψυχολογία, όπου διαχωρίζουμε π.χ. αντιληπτικές από εννοιολογικές αναπαραστάσεις και λειτουργίες. Επίσης δεν αντιστοιχεί σε διαχωρισμούς μεγάλης κλίμακας από τη γνωσιακή νευροεπιστήμη, όπου συναντά κανείς π.χ. διάφορα συστήματα μνήμης με διαφορετικές λειτουργίες και διαφορετικούς σκοπούς σε διαφορετικές περιοχές του εγκεφάλου. Ο λειτουργικός κατακερματισμός της εξελικτικής θεώρησης είναι πολύ βαθύτερος και αγγίζει επιμέρους νοητικές διεργασίες σε εξαιρετικά λεπτομερειακό επίπεδο. Ουσιαστικά υποστηρίζεται ότι ο νους αποτελείται από διαφορετικά μεταξύ τους εργαλεία, καθένα από τα οποία έχει μια συγκεκριμένη λειτουργία και δεν συνδέεται με τα υπόλοιπα πέρα από το γεγονός της συνύπαρξής τους στην ίδια συσκευασία.

Η εξελικτική θεώρηση εστιάζει στο ρόλο του νου (ως λειτουργίας του νευρικού συστήματος) για την επιβίωση και αναπαραγωγή μέσα σε ένα οικολογικό πλαίσιο. Αντιμετωπίζει το νου ως απάντηση στο πρόβλημα της αρμοστικότητας, ακολουθώντας την τρέχουσα προσέγγιση

γηση στη βιολογία. Συνεπώς, η νοητική λειτουργία οφείλει να περιλαμβάνει συγκεκριμένες αρμοστικές λύσεις σε συγκεκριμένα προβλήματα, όπως κάθε άλλο όργανο ή λειτουργία του οργανισμού. Έτσι, η εξελικτική προσέγγιση δεν θεωρεί την εξέλιξη ενός κεντρικού επεξεργαστή ως πιθανό αποτέλεσμα της εξελικτικής πορείας του νου, διότι ο μοναδικός κεντρικός επεξεργαστής δεν φαίνεται να αποτελεί λύση σε κάποιο συγκεκριμένο πρόβλημα. Η εξέλιξη ενός «γενικού διαχειριστή συμβόλων» φαντάζει ασαφής και αρμοστικά αβέβαιη. Δεδομένου ότι η προσαρμοστική εξέλιξη μιας λειτουργίας απαιτεί σταδιακή διαμόρφωση και βελτίωσή της, αναρωτιέται κανείς πώς θα μπορούσε να είναι ένας «μερικός» επεξεργαστής. Πώς θα μπορούσε να προκύψει, τι είδους γενετική κωδικοποίηση θα μπορούσε να τον υποστηρίξει, και ποιες εξελικτικές πιέσεις θα οδηγούσαν στη βελτίωσή του.

Επιπλέον, ο ρόλος ενός κεντρικού επεξεργαστή γενικής χρήσεως φαντάζει ακατανόητος μέσα στο οικολογικό πλαίσιο της εξέλιξης διότι η γενικότητα αντιστρατεύεται την αποτελεσματικότητα: κάθε εργαλείο υψηλής απόδοσης είναι ταυτόχρονα και υψηλής εξειδίκευσης, διότι κάθε τροποποίηση που αποτελεί βελτίωση προς συγκεκριμένο στόχο ταυτόχρονα αποτελεί απομάκρυνση από οποιοδήποτε εναλλακτικό στόχο. Ένα αισθητήριο αυξάνει την ευαισθησία του με προσαρμογή σε συγκεκριμένο είδος φυσικής ενέργειας, ένα εργαλείο αυξάνει την αποτελεσματικότητά του με εξειδίκευση του σχήματος, και ένας οργανισμός επιτυγχάνει καλύτερη προσαρμογή στο περιβάλλον εξειδικεύοντας, όχι γενικεύοντας, τη μορφή και τη συμπεριφορά του ώστε να είναι ταιριαστή στα χαρακτηριστικά του περιβάλλοντος και των ομοίων του. Έτσι η εξελικτική θεώρηση απορρίπτει τον κεντρικό επεξεργαστή που υποτίθεται πως «τα κάνει όλα» και στη θέση του προκρίνει ένα συνονθύλευμα διακριτών λειτουργικών μονάδων με απλή σχέση «συγκατοίκησης» μεταξύ τους. Άλλωστε, τα προβλήματα που πραγματικά έχουν μελετηθεί στη γνωσιακή επιστήμη υπό το πρίσμα του κεντρικού επεξεργαστή δεν αποτελούν δείγμα ευρέος φάσματος αλλά μια εξαιρετικά επιλεκτική και περιοριστική συλλογή έργων απομακρυσμένων από τη φυσική λειτουργία και τη νοητική καθημερινότητα. Ουσιαστικά, δηλαδή, ο υποτιθέμενος κεντρικός επεξεργαστής είναι και αυτός εξειδικευμένος σε συγκεκριμένου τύπου προβλήματα (αριθμητικής, λογικής-συλλογισμού, νοημοσύνης, λεκτικής-προτασιακής επεξεργασίας κλπ.) και μάλιστα σε προβλήματα σχολικού και όχι κοινωνικού ή οικολογικού τύπου. Για το λόγο αυτό, οι επιτυχίες και οι αποτυχίες της τεχνητής νοημοσύνης, που βασίζεται στην προσέγγιση του γενικού κεντρικού επεξεργαστή, αναδεικνύουν τα πεδία φυσικής νόησης (βλ. Ενότητα 1.2, σελ. 4) ως εκείνα τα προβλήματα που φαντάζουν πολύ απλά, σχεδόν αυτόνοχα, κι όμως δεν μπορούν να λυθούν με τις γνωστές υπολογιστικές μεθόδους. Αντίθετα, η τεχνητή νοημοσύνη είναι πολύ αποτελεσματική σε πεδία όπου ο φυσικός νους δεν τα καταφέρνει καλά και ως εκ τούτου μας φαίνονται δύσκολα, ενώ στην πραγματικότητα είναι πολύ ευκολότερα από υπολογιστική σκοπιά. Η διάκριση αυτή δείχνει ότι ο νους έχει εξελίξει αποτελεσματικές λύσεις για ιδιαίτερα απαιτητικά προβλήματα, δεν είναι όμως προσαρμοσμένος σε πεδία συμβολικής επεξεργασίας και κεντρικών συντακτικών επεξεργαστών.

Η θεώρηση του κατακερματισμού καθεαυτή δεν είναι πρωτότυπη ούτε ριζοσπαστική, και επιτρέπει ένα ευρύ φάσμα υποθέσεων για τις σχέσεις μεταξύ των λειτουργικών μο-

νάδων. Σε επίπεδο οργανισμού, κάθε όργανο έχει συγκεκριμένες λειτουργίες και αντιμετωπίζει ένα σχετικά στενό φάσμα αναγκών, για τις οποίες υποθέτουμε ότι εξελίχθηκε. Συνεργάζεται στενά με τα υπόλοιπα όργανα και διαθέτει μηχανισμούς «επικοινωνίας» και συντονισμού με το σύνολο, όμως στην ουσία κάνει τη δικιά του δουλειά με το δικό του τρόπο, και για το σκοπό αυτό έχει ξεχωριστή μορφή, δομή, και λειτουργικά χαρακτηριστικά.

Από το χώρο της νευροψυχολογίας μπορεί κανείς να αντλήσει υποστηρικτικά στοιχεία για τη θέση ότι ο νους είναι ένα σύστημα επιμέρους εργαλείων και όχι ένας κεντρικός επεξεργαστής. Συγκεκριμένα, από τις διάφορες εγκεφαλικές βλάβες όλων των ειδών που έχουν μελετηθεί, φαίνεται ότι οι διαφορετικές περιοχές του εγκεφάλου εξυπηρετούν διαφορετικές εξειδικευμένες λειτουργίες, καθεμιά από τις οποίες μπορεί να υποστεί βλάβη και έτσι να απωλεσθεί μια ενδεχομένως πολύ συγκεκριμένη ικανότητα. Βέβαια η κλινική πρακτική δείχνει ότι εντελώς περιγεγραμμένα λειτουργικά ελλείμματα είναι σπάνια, δηλαδή συνήθως μια εγκεφαλική βλάβη έχει ένα εύρος συνεπειών, αλλά η βαρύτητα των επιπτώσεων δεν εξαρτάται μόνο από το μέγεθος της βλάβης. Μπορεί μια σχετικά εντοπισμένη βλάβη ιστού να συνδέεται με έντονη λειτουργική διαταραχή σε κάποιο συγκεκριμένο τομέα, έως και απώλεια κάποιας ικανότητας, έστω και σε συνδυασμό με μικρότερης έκτασης απώλειες σε άλλους λειτουργικούς τομείς. Αυτό σημαίνει ότι ορισμένες εγκεφαλικές περιοχές αντιστοιχούν σε συγκεκριμένες νοητικές λειτουργίες και κατ'επέκταση ότι υπάρχουν λειτουργικά διαχωρίσιμες νοητικές λειτουργίες. Πρέπει όμως να διευκρινιστεί ότι η εξελικτική θεώρηση του νου δεν αφορά σε εγκεφαλικές περιοχές (ούτε, γενικότερα, στη δομή του νευρικού συστήματος). Απλώς τα στοιχεία από εγκεφαλικές βλάβες μπορούν να χρησιμοποιηθούν για να υποστηριχθεί η θέση ότι ο νους περιλαμβάνει ένα σύνολο εξειδικευμένων νοητικών εργαλείων.

7.4 Παλαιοανθρωπολογία

Ένα θεμελιώδες χαρακτηριστικό της εξελικτικής θεώρησης του νου έχει να κάνει με το πεδίο στο οποίο εξελίχθηκε ο νους, δηλαδή τον οικολογικό περίγυρο μέσα στα πλαίσια του οποίου θεωρείται ότι έδρασαν οι εξελικτικές πιέσεις που οδήγησαν στη σημερινή μορφή του. Εδώ πρέπει να διακρίνουμε μεταξύ νοητικών λειτουργιών που μοιραζόμαστε με άλλα είδη, δηλαδή που κληρονομήσαμε από τους κοινούς μας προγόνους στην εξελικτική ιστορία, και νοητικών λειτουργιών που αφορούν σε αποκλειστικά ανθρώπινες προσαρμογές, δηλαδή προέκυψαν ως απαντήσεις σε προβλήματα των πρωτοανθρώπων μετά τη διάσπαση της εξελικτικής πορείας του γένους *homo* από τον κορμό των πρωτευόντων. Η διάκριση αυτή δεν αλλάζει την ουσία της προσέγγισης, έχει όμως σημαντικές συνέπειες για το είδος του οικολογικού περιγύρου μέσα στο οποίο θεωρούμε ότι έλαβαν χώρα οι αντίστοιχες εξελικτικές πιέσεις. Οποσδήποτε, η εξελικτική προσέγγιση δεσμεύεται στην αναγνώριση μιας ανθρώπινης φύσης, δηλαδή ενός χαρακτηριστικού νοητικού περιγράμματος που είναι σε μεγάλο βαθμό γενετικά προκαθορισμένο και δεν υπόκειται σε σημαντικές κοινωνικές ή πολιτισμικές επιδράσεις. Ο νοητικός αυτός «πυρήνας» χαρακτηρίζει την ανθρώπινη φύση όπως η μορφή του σώματος και οι λειτουργίες των άλλων οργάνων του.

Αναφερόμενοι σε αποκλειστικά ανθρώπινες νοητικές λειτουργίες, ή σε νοητικές λειτουργίες οι οποίες θεωρούμε ότι έχουν τροποποιηθεί σε μεγάλο βαθμό κατά την πρόσφατη εξελικτική πορεία του ανθρώπινου είδους, είναι προφανές ότι ο χρόνος και ο τόπος αρμοστικής δράσης, δηλαδή το διαθέσιμο πεδίο της εξέλιξης, υπήρξε εξαιρετικά περιορισμένο. Ο ανθρώπινος κλάδος του δέντρου της εξέλιξης έχει ηλικία λίγων μόνο εκατομμυρίων ετών (μεταξύ 5 και 8 εκατομμυρίων ετών υπολογίζεται η εξελικτική μας απόσταση από τον κοινό πρόγονό μας με τον κοντινότερο «συγγενή», τον χιμπατζή). Το σημερινό ανθρώπινο είδος, ο *homo sapiens sapiens*, έχει εξελικτική ηλικία της τάξης των 100–150 χιλιάδων ετών, και σύμφωνα με την επικρατέστερη υπόθεση η ιστορία μεταξύ του κοινού προγόνου μας με τον χιμπατζή και του πρώτου σύγχρονου ανθρώπου διαδραματίστηκε στις πεδιάδες της Αφρικής, μέσα σε μια πανίδα και χλωρίδα που δεν διέφερε και πάρα πολύ από τη σημερινή, καθώς η πυκνή βλάστηση της ζούγκλας έδωσε τη θέση της σε σαβάνες με πιο αραιή και χαμηλή βλάστηση, με αποτέλεσμα την εξέλιξη του χιμπατζή στις ζούγκλες και των ανθρωπίδων στις σαβάνες. Αυτό είναι το «περιβάλλον εξελικτικής προσαρμογής» (*environment of evolutionary adaptedness*) του ανθρώπου και μόνο με αναφορά σε τέτοιες συνθήκες μπορούμε να εξετάσουμε τη λειτουργικότητα των βιολογικών και νοητικών χαρακτηριστικών που διαφοροποιούν τον άνθρωπο από το χιμπατζή.

Δεδομένης της κοινωνικότητας των χιμπατζήδων και των άλλων πρωτευόντων, και των διαφόρων παλαιοντολογικών ευρημάτων που υποδεικνύουν κοινωνική οργάνωση των προανθρώπων και πρωτανθρώπων, θεωρείται δεδομένο ότι η εξέλιξη των νοητικών λειτουργιών έλαβε χώρα στο πλαίσιο ευρέων κοινωνικών αλληλεπιδράσεων. Την εικόνα συμπληρώνουν οι αυξανόμενες γνώσεις μας για τις ικανότητες και τις συνήθειες των διαφόρων ενδιάμεσων ειδών, π.χ. σχετικά με το αν κατασκεύαζαν εργαλεία και τι είδους ήταν αυτά, ποιες ήταν οι διατροφικές προτιμήσεις τους, ο νομαδικός χαρακτήρας, το μέγεθος και η σύσταση των κοινωνικών ομάδων κλπ. Καθώς οι διαφορές μας από το χιμπατζή προέκυψαν μέσα σε αυτό το συγκεκριμένο φυσικό και κοινωνικό περιβάλλον και μέσα σε συγκεκριμένο χρονικό διάστημα, αυτό διευκολύνει τη μελέτη καθώς περιορίζεται σημαντικά τόσο το εύρος όσο και η αιτιολόγηση των εξελικτικών μεταβολών. Συγκεκριμένα, οι ειδικά ανθρώπινες λειτουργίες θα πρέπει να αποτελούν αρμοστικές λύσεις σε προβλήματα του περιβάλλοντος εξελικτικής προσαρμογής, διαφορετικά δεν έχουν λόγο εμφάνισης και παγίωσης ως χαρακτηριστικά του είδους και μάλιστα σε τόσο σύντομο χρονικό διάστημα. Επίσης, η έκταση των εξελικτικών μεταβολών πρέπει αναγκαστικά να είναι περιορισμένη λόγω του σύντομου χρονικού διαστήματος, καθώς η πιθανότητα τυχαίων ευεργετικών μεταλλάξεων είναι χαμηλή και ο ρυθμός εξάπλωσης και εδραίωσης χαρακτηριστικών δεν είναι ραγδαίος.

Η κύρια θεωρητική δέσμευση της εξελικτικής ψυχολογίας είναι ότι οι ανθρώπινες νοητικές λειτουργίες αποτελούν λύσεις σε εξελικτικά πρόβληματα του περιβάλλοντος εξελικτικής προσαρμογής και όχι του σημερινού περιβάλλοντος, το οποίο έχει προκύψει πολύ πρόσφατα και δεν μπορεί να έχει αποτελέσει ακόμα σημαντικό παράγοντα εξελικτικής πίεσης. Λαμβάνοντας υπόψη ότι η γεωργία, που οδήγησε στη συγκέντρωση πληθυσμών, έχει ιστορία μόνο μερικών χιλιάδων ετών, γίνεται αντιληπτό ότι οι ανθρώπινες κοινωνίες (μικρές

νομαδικές ομάδες) ανέπτυξαν και διατήρησαν τα χαρακτηριστικά τους μέσα σε ένα σχετικά σταθερό φυσικό και κοινωνικό περιβάλλον με πολύ διαφορετικές απαιτήσεις από τις σημερινές. Έτσι, δεν έχει νόημα να εξετάζουμε τις νοητικές λειτουργίες ως προσαρμοσμένες στα προβλήματα της σημερινής κοινωνίας, καθώς δεν υπάρχει λόγος να θεωρούμε ότι ο νους είναι προσαρμοσμένος για να λειτουργεί στις συνθήκες που δημιούργησε. Όπως τα δόντια μας δεν είναι προσαρμοσμένα στις διατροφικές συνθήκες που ακολούθησαν την εξάρτηση από τη γεωργία, με αποτέλεσμα το διαδεδομένο πρόβλημα της τερηδόνας, έτσι και ο νους μας μπορεί κάλλιστα να μην είναι προσαρμοσμένος στη σχολική εκπαίδευση ή τη μαζική συμβίωση σε πόλεις χιλιάδων κατοίκων. Υπό αυτό το πρίσμα, η κατανόηση του νου δεν μπορεί να προκύψει από μελέτη του στις τεχνητές σημερινές συνθήκες αλλά πρέπει να προέλθει λαμβάνοντας υπόψη τις συνθήκες στις οποίες εξελίχθηκε και προσαρμόστηκε.

Μέσα σε αυτήν την εικόνα έρχεται η εξελικτική θεώρηση του νου να δικαιολογήσει σεναριακά τις προσαρμοστικές πιέσεις που οδήγησαν στην εξέλιξη καθεμιάς από τις εξειδικευμένες λειτουργικές μονάδες του νου. Οι συμπληρωματικές υποθέσεις περιλαμβάνουν τις γενικότερες απαιτήσεις εξελικτικού τύπου για επιβίωση, αναπαραγωγή, όπως και τις λιγότερο γενικές προδιαγραφές για εξεύρεση τροφής και καταφυγίου, αποφυγή θηρευτών, αναπαραγωγική επιλογή (συντρόφου), κοινωνική αλληλεπίδραση και συναλλαγή, αλληλοϋποστήριξη μεταξύ συγγενών και συμμάχων, διατήρηση της κοινωνικής ιεραρχίας κλπ. Εξειδικευμένες νοητικές λειτουργίες μπορούν να διαμεσολαβούνται από νοητικούς μηχανισμούς που υποστηρίζουν εξελικτικά ορθολογικές επιλογές, όπως ενδεχομένως είναι η έμφυτη ικανότητα ηθικής κρίσης (αναγνώριση καλού-κακού) και αισθητικής αποτίμησης (αναγνώριση ωραίου-άσχημου). Με αυτά τα στοιχεία αφειτηρίας, οι υποστηρικτές της εξελικτικής προσέγγισης επιχειρούν να παρουσιάσουν αποδεκτά σενάρια για τους λόγους που οδήγησαν στη δημιουργία και προσαρμογή συγκεκριμένων νοητικών μονάδων, στοχεύοντας κυρίως στη δικαιολόγηση της υπόθεσης της ύπαρξης των μονάδων αυτών. Η στρατηγική αυτή συνίσταται στην υπόθεση ότι αν μια νοητική λειτουργία έχει λόγο ύπαρξης με βάση τις υποτιθέμενες ανάγκες του ανθρώπου στο εξελικτικό του περιβάλλον, τότε ισχυροποιείται επιστημολογικά η θέση της στο μοντέλο της νόησης.

7.5 Διαψευσιμότητα

Ένα μειονέκτημα της εξελικτικής προσέγγισης είναι ότι ενώ δεν προσφέρει συγκεκριμένα εργαλεία κατασκευής νοητικών μοντέλων και κριτήρια ελέγχου τους, παρέχει ένα υπόβαθρο ιδεών στο οποίο μπορούν να αναπτυχθούν ετερόκλητες και ανεξέλεγκτες θεωρητικές προτάσεις. Η εξαιρετικά ελκυστική ιδέα της βιολογικής σκέψης στην δικαιολόγηση νοητικών λειτουργιών με αναφορά σε αρμοστικά χαρακτηριστικά τους διατρέχει τον κίνδυνο να οδηγήσει σε ακατάσχετη σεναριολογία, η οποία εξυπηρετεί εκ των προτέρων προτιμήσεις. Έτσι, υπό τη σκέπη της εξελικτικής ψυχολογίας βρίσκει κανείς αφενός εξονυχιστικά μελετημένες και τεκμηριωμένες προσεγγίσεις, όπως είναι η θεωρία της «ανίχνευσης α-

πατεώνων» (cheater detection)³² στο πρόβλημα της επιλογής του Wason, και αφετέρου διάφορες αβάσιμες εικασίες για τους λόγους που οι άνδρες επιλέγουν γυναίκες με κάποια χαρακτηριστικά (και αντίστροφα), που ενίοτε παραπέμπουν περισσότερο σε εκλαϊκευτικά στερεότυπα παρά σε επιστημονική ουδετερότητα. Η επιμεριστική προσαρμοστικότητα και η σεναριολογία μιας υποτιθέμενης παλαιο-ανθρωπολογικής οικολογίας έχουν συνεισφέρει στο κακό όνομα της κοινωνιοβιολογίας³³, σε συνδυασμό με πολιτικές τοποθετήσεις και άλλες μη επιστημονικές αντιρρήσεις που έχουν να κάνουν με ενδόμυχους φόβους των επιπτώσεων μιας άκριτης αποδοχής της ανθρώπινης φύσης. Ίσως το βασικότερο πρόβλημα με την ακραία έκφραση της εξελικτικής θεώρησης του νου, όπως και με την κοινωνιοβιολογία, να είναι όχι το αν είναι λανθασμένη, αλλά το ότι δεν είναι δυνατό να εξακριβώσουμε αν είναι λανθασμένη.

Μια εξελικτική θεώρηση για το νου είναι σε κάποιο βαθμό αυτονόητη και τετριμμένη. Βέβαια δεν είναι σαφές το κατά πόσον οι νοητικές δυνατότητες, είτε ως συνολικό σύστημα νόησης είτε ως επιμέρους λειτουργικά στοιχεία, αποτέλεσαν οι ίδιες στόχο αρμοστικών πιέσεων. Δεν είναι λοιπόν γνωστό με βεβαιότητα αν οι νοητικές λειτουργίες (ή η νόηση γενικότερα) μπορούν να θεωρηθούν αρμοστικά προϊόντα της εξέλιξης με την αυστηρότερη αυτή έννοια. Ο λόγος που δεν είναι γνωστό αυτό διαθέτει δύο επιμέρους συνιστώσες: Πρώτον, γνωρίζουμε από τη βιολογική μελέτη της εξέλιξης ότι υπάρχουν σημαντικοί μη προσαρμοστικοί μηχανισμοί που οδηγούν στην εμφάνιση, εξαφάνιση, τροποποίηση ή ριζικό μετασχηματισμό ενός χαρακτηριστικού, φυσικού ή λειτουργικού. Και δεύτερον, είναι αδύνατο να διαπιστωθεί είτε πειραματικά είτε παλαιοντολογικά αν μετρήσιμες διαφορές στο επίπεδο της νόησης (συγκεκριμένων νοητικών διεργασιών) αποτέλεσαν προσαρμοστικό πλεονέκτημα για τους κατόχους τους. Δεν είναι δηλαδή πρακτικά εφικτό να διαπιστωθεί αν οι προ-άνθρωποι ή πρωτο-άνθρωποι που διέθεταν «περισσότερη» ή «καλύτερη» νόηση (γενικά ή σε κάποια εξειδικευμένη λειτουργία) είχαν, εξαιτίας αυτής της διαφοράς στη νοητική λειτουργία, μεγαλύτερη πιθανότητα διατήρησης και εξάπλωσης του γενετικού τους υλικού (επιβίωσης και αναπαραγωγής). Πάντα θα μπορεί να εγερθεί η ενδεχόμενη αντίρρηση ότι η νόηση αποτελεί συνοδό χαρακτηριστικό κάποιου άλλου αρμοστικού γνωρίσματος, ή ότι προέκυψε ως παράπλευρη συνέπεια ενός συνδυασμού άλλων αρμοστικών χαρακτηριστικών που εδραιώθηκαν για διαφορετικούς λόγους. Συνεπώς ό,τι μπορεί να ειπωθεί τώρα για το θέμα είναι εκ των υστέρων δικαιολόγηση και όχι επιστημονική τεκμηρίωση.

Το πρόβλημα με τις εκ των υστέρων δικαιολογήσεις είναι ότι, ως μεθοδολογική στρατη-

³²Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163-228). Oxford University Press.

³³Κοινωνιοβιολογία ονομάζεται η μελέτη των βιολογικών βάσεων της κοινωνικής συμπεριφοράς σε όλα τα είδη, συμπεριλαμβανομένου του ανθρώπου. Δίνει ιδιαίτερη έμφαση στην εξελικτική οπτική γωνία που περισσότερο ενώνει παρά χωρίζει τα είδη, από τα μυρμήγκια μέχρι τον άνθρωπο. Η επιθυμία να αναχθούν όψεις της ανθρώπινης κοινωνικής συμπεριφοράς σε εγγενείς τάσεις και πληθυσμιακά χαρακτηριστικά έχει οδηγήσει σε ορισμένες εκατέρωθεν ακρότητες: αφενός εξαιρετικά ασθενείς θεωρίες, μάλλον εικασίες, και αφετέρου εξαιρετικά ισχυρές αντιρρήσεις πολιτικού και ηθικού παρά ουσιαστικού επιστημονικού περιεχομένου.

γική, δεν μπορούν να διαψευστούν, διότι πάντα μπορεί να παραχθεί μια νέα «διευκρίνιση» ή τροποποίηση, αν τυχόν προκύψουν νέα στοιχεία. Για ένα ζήτημα όπως η εξέλιξη της νόησης, με το οποίο ελάχιστα πραγματικά στοιχεία με φυσική υπόσταση έχουν ισχυρή σχέση και είναι μάλλον αμφίβολο αν ποτέ ανακαλυφθούν νέα και ισχυρότερα, μια εκ των υστέρων δικαιολόγηση είναι ουσιαστικά εκ γενετής άτρωτη. Από τη θετική πλευρά, πρόκειται για μια «αφήγηση» που διευκολύνει στη σύλληψη και νοηματοδότηση και που μπορεί να δώσει ερεθίσματα για περισσότερη και πιο βαθιά μελέτη των ζητημάτων. Από την αρνητική πλευρά, πρόκειται για ένα βλαβερό παραμύθι που αφενός διαστρεβλώνει την πραγματικότητα (διότι είναι αμφίβολο να έχει κάποιος μαντέψει σωστά με την πρώτη, δηλαδή χωρίς πειράματα, βάσει τόσο ελλιπών στοιχείων) και αφετέρου παραπλανά τους ενδιαφερόμενους (διότι δίνει την εντύπωση ότι διαθέτουμε εξήγηση για ένα γεγονός επιστημονικά ακατανόητο). Μια ενδιαμέση προσέγγιση, κάπως ουδέτερη, θα ήταν ίσως λιγότερο αυστηρή στους χαρακτηρισμούς της αλλά και πάλι θα όφειλε να διαχωρίσει την επιστημολογική θέση της δικαιολόγησης από εδραιωμένες επιστημονικές θεωρίες.

7.6 Περίληψη

Η εξελικτική προσέγγιση στο νου δεν περιορίζει το είδος των μοντέλων που μπορούμε να κατασκευάσουμε για τις νοητικές λειτουργίες. Εστιάζει στη φυλογονική εξέλιξη των νοητικών λειτουργιών, δηλαδή το πώς προέκυψαν οι σχετικοί νοητικοί μηχανισμοί κάτω από εξελικτικές πιέσεις μέσα από τη διαδοχή των γενεών. Στη φυλογονική χρονική κλίμακα αναδύονται νοητικές λειτουργίες που αποδεικνύονται αποδοτικές στην αντιμετώπιση πιεστικών προβλημάτων επιβίωσης, και εντέλει επιλέγονται εκείνες που οδηγούν σε αυξημένη αναπαραγωγική επιτυχία, δηλαδή αρμοστικότητα. Η εξελικτική θεώρηση δεν προσδιορίζει τι είδους μηχανή είναι και τι υπολογισμούς κάνει ο νους, ούτε αναφέρεται καν στην οντογονική διάσταση (χρονική κλίμακα της ανάπτυξης του ατόμου) ή τη δυναμική διάσταση (χρονική κλίμακα της εκτέλεσης μιας νοητικής διεργασίας). Η έμφαση στο περιβάλλον εξελικτικής προσαρμογής περιορίζει μόνο το είδος των προβλημάτων που οδήγησαν στην εμφάνιση και εξέλιξη των νοητικών διεργασιών, και μας κατευθύνει ως προς τα πεδία νοητικών διεργασιών που είναι πιθανότερο να αποτέλεσαν αρμοστικούς στόχους. Τέλος, η εξελικτική σκοπιά τονίζει την ανεξαρτησία μεταξύ επιμέρους νοητικών μηχανισμών, τους οποίους θεωρεί επιμέρους λύσεις σε διαφορετικά συγκεκριμένα προβλήματα, με αποτέλεσμα την εξειδίκευσή τους, σε πλήρη αντίθεση με τη θεώρηση ενός κεντρικού επεξεργαστή και ενός γενικού μηχανισμού σκέψης και λύσης προβλημάτων.

8 Σύνοψη των θεωρητικών προσεγγίσεων

Οι προαναφερθείσες θεωρητικές προσεγγίσεις βασίζονται σε διαφορετικές αφετηρίες και συλλαμβάνουν το νου με διαφορετικό τρόπο, τόσο ως προς τη σύστασή του όσο και ως προς το σκοπό του. Η συμβολική προσέγγιση βασίζεται στο μοντέλο του ηλεκτρονικού

υπολογιστή και των τυπικών μηχανών γενικότερα. Η συνδεδειστική προσέγγιση βασίζεται στα βιολογικά νευρωνικά δίκτυα και τη λειτουργία του εγκεφάλου γενικότερα. Η δυναμική προσέγγιση βασίζεται στα φυσικά δυναμικά συστήματα με τις τροχιές μέσα σε ένα χώρο φάσεων. Αυτές οι τρεις παρέχουν συγκεκριμένα εργαλεία για την προσομοίωση του νου και με αυτήν την έννοια είναι ταυτόχρονα και μεθοδολογικές προτάσεις για την κατασκευή υπολογιστικών μοντέλων: Η συμβολική προσέγγιση ορίζει πως ένα μοντέλο του νου αποτελείται από συμβολικές αναπαραστάσεις και συντακτικούς κανόνες, συγκροτώντας ένα συντακτικό επεξεργαστή συμβόλων. Η συνδεδειστική προσέγγιση ορίζει πως ένα μοντέλο του νου αποτελείται από κόμβους (αυτόνομες επεξεργαστικές μονάδες που χαρακτηρίζονται από ένα βαθμό ενεργοποίησης) και συνδέσεις (με ρυθμιζόμενα βάρη), συγκροτώντας ένα τεχνητό νευρωνικό δίκτυο. Η δυναμική προσέγγιση ορίζει πως ένα μοντέλο του νου είναι σύστημα με προσδιορισμένη χρονική εξέλιξη, σε διακριτό ή συνεχή χρόνο, μέσω εξισώσεων διαφορών ή διαφορικών εξισώσεων αντίστοιχα. Ένα δυναμικό μοντέλο μπορεί να υλοποιηθεί με εναλλακτικούς τρόπους, από συστήματα εξισώσεων μέχρι αναδρομικά συνδεδειστικά δίκτυα.

Οι τρεις τελευταίες προσεγγίσεις δεν προσδιορίζουν εργαλεία αλλά μόνο πεδία σύλληψης και έμφασης των νοητικών λειτουργιών. Ως εκ τούτου χρησιμεύουν περισσότερο στην επιλογή των νοητικών λειτουργιών που θα μελετήσουμε και το πλαίσιο στο οποίο θα τις μελετήσουμε. Η ενσώματη νόηση τονίζει τη συμμετοχή του σώματος και το οικολογικό πλαίσιο λειτουργίας στην εξέλιξη και εκτέλεση των νοητικών λειτουργιών, δίνοντας έμφαση στη στενή σύζευξη νου-περιβάλλοντος και στη δράση σε πραγματικό χρόνο, είτε υποτιμώντας την αφηρημένη νόηση είτε θεωρώντας την και αυτήν ένα είδος δράσης. Η πιθανοκρατική προσέγγιση δίνει έμφαση στην (εξ ορισμού) ορθολογικότητα της νοητικής λειτουργίας και εστιάζει στο προϊόν της πιθανοκρατικά βέλτιστης απόφασης, δίνοντας την ευκαιρία της μελέτης των παραγόντων που λαμβάνονται υπόψη στην παραγωγή της αλλά παραβλέποντας τους γνωσιακούς μηχανισμούς που επιτελούν τους υπολογισμούς. Τέλος η εξελικτική προσέγγιση τονίζει το ταίριασμα των επιμέρους νοητικών λειτουργιών στις συνθήκες και τα προβλήματα του εξελικτικού περιβάλλοντος, δίνοντας έμφαση στην ανάλυση των νοητικών λειτουργιών υπό το πρίσμα των εξελικτικών πιέσεων στις οποίες προσαρμόστηκαν, και αναδεικνύοντας την ακαταλληλότητα των σημερινών συνθηκών (και πολλών πειραματικών διατάξεων) για την ανακάλυψη των θεμελιωδών νοητικών ιδιοτήτων. Προς το παρόν δεν φαίνεται κάποια ξεκάθαρα πλεονεκτική ή μειονεκτική προσέγγιση, καθώς όλες έχουν τα πλεονεκτήματα και τα μειονεκτήματά τους, τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Δεν πρέπει να μας διαφευγει ότι μια πρακτικά εφαρμόσιμη θεωρία μπορεί να είναι πιο χρήσιμη από μια άλλη που φαίνεται θεωρητικά σωστότερη αλλά δεν είναι διαχειρίσιμη σε πειραματικό ή υπολογιστικό επίπεδο. Το μέλλον θα δείξει αν τελικά προκύψει γόνιμη σύνθεση των διαφορετικών προσεγγίσεων ή επιλογή κάποιας (ή κάποιων) ως πιο σωστής και πιο εποικοδομητικής για τη συνολική κατανόηση των νοητικών φαινομένων.

Μεθοδολογικές προσεγγίσεις

Η γνωσιακή επιστήμη είναι ένα ευρύτατο, κατεξοχήν διεπιστημονικό ερευνητικό πεδίο. Θα μπορούσε να υποστηρίξει κανείς ότι είναι ένα συνονθύλευμα θεωρητικών και μεθοδολογικών προσεγγίσεων στο ζήτημα της νόησης από διαφορετικούς κλάδους, άρα θα έπρεπε να μιλάμε για γνωσιακές επιστήμες και όχι γνωσιακή επιστήμη. Ιστορικά διαφαίνεται μια τάση συγκρότησης κλάδου, όχι μέσα από τη θεωρητική και μεθοδολογική ομογενοποίηση των προσεγγίσεων, αλλά εξαιτίας της εδραιωμένης πλέον συνεργασίας μεταξύ τους και της δεδομένης βούλησης και ικανότητας επιστημόνων από πολλούς κλάδους να συνεννοούνται μεταξύ τους. Πάντως παρότι αυξάνονται διεθνώς τα τμήματα και τα προγράμματα «γνωσιακής επιστήμης», το πρόσφατα αναδειχθέν ως κορυφαίο περιοδικό γενικών ανασκοπήσεων του κλάδου ονομάζεται σταθερά *Trends in Cognitive Sciences* (πληθυντικός αριθμός), θυμίζοντας πάντα ότι πρόκειται για ένα ετερόκλητο επιστημονικό σύνολο.

Η γνωσιακή επιστήμη, ως πολυσυλλεκτικό ερευνητικό πεδίο, αποφεύγει τα στεγανά. Όποιος έχει κάτι να προσφέρει είναι ευπρόσδεκτος – κι αν όχι από καθέναν ξεχωριστά, σίγουρα από την επιστημονική «κοινότητα» στο σύνολό της. Έτσι συνυπάρχουν και συνεργάζονται ψυχολόγοι, γλωσσολόγοι, μαθηματικοί, μηχανικοί, βιολόγοι, φιλόσοφοι, γιατροί, φυσικοί, κοινωνιολόγοι, ανθρωπολόγοι, ζωολόγοι κ.ά. Αν επιχειρήσουμε όμως μια γενική ταξινόμηση των μεθόδων που συνεισφέρουν τον κύριο όγκο δεδομένων στο έργο της μελέτης της νόησης, θα δούμε ότι διακρίνονται ξεκάθαρα τρεις κεντρικές κατευθύνσεις:

1. Η συλλογή εμπειρικών δεδομένων σχετικά με τη συμπεριφορά του ανθρώπου (και σε περιορισμένη κλίμακα άλλων ζώων) με πειραματικές μεθόδους και η θεωρητική συγκρότηση των αποτελεσμάτων σε μηχανισμούς και διεργασίες. Οι σημαντικότερες από τις παραγόμενες θεωρίες τείνουν να περικλείουν (να «εξηγούν») μεγάλο πλήθος πειραματικών στοιχείων αλλά οι τάσεις συγκερασμού μεταξύ των θεωριών στα επιμέρους υπο-πεδία είναι περιορισμένες. Η κατεύθυνση αυτή έχει τις ιστορικές ρίζες της στην ψυχολογία και πιο πρόσφατα ιδιαίτερα στον κλάδο της *γνωστικής ψυχολογίας*.
2. Η συλλογή εμπειρικών δεδομένων σχετικά με τη δομή και λειτουργία των νευρικών συστημάτων, σε άνθρωπο και ζώα, με ελεγχόμενη παρατήρηση και η συγκρότηση των ευρημάτων σε λειτουργικές περιγραφές σε επίπεδο συστήματος. Οι θεωρητικές κατασκευές είναι περιορισμένες και η έμφαση δίνεται στην περιγραφή σε πολλά επίπεδα (γενετικό, μοριακό, κυτταρικό, δικτυακό, συστημικό). Η κατεύθυνση αυτή έχει τις ιστορικές ρίζες της στη βιολογία και τη νευρολογία και πλέον εντοπίζεται ιδιαίτερα στους σύγχρονους κλάδους των *νευροεπιστημών*.
3. Η διατύπωση θεωριών και υποθέσεων σχετικά με τη δομή και την αρχιτεκτονική της νόησης με έμφαση σε θέματα όπως η σχέση εγκεφάλου-νου, το περιεχόμενο των εννοιών και οι μηχανισμοί (διεργασίες) σκέψης. Οι εμπειρικές παρατηρήσεις είναι περιορισμένου εύρους και ο χαρακτήρας τους είναι παραδειγματικός. Αντίθετα,

οι θεωρητικές κατασκευές διέπονται από μεγάλη αναλυτικότητα και επιχειρούν να αποτυπώσουν σε βάθος το περιεχόμενο και τις λειτουργίες της νόησης. Στην κατεύθυνση αυτή, που έχει τις ιστορικές ρίζες της στη *φιλοσοφία* και ιδιαίτερα τη λογική παράδοση, έχει παίξει κεντρικό ρόλο η *θεωρητική γλωσσολογία* και η σύλληψη της γλώσσας αφενός ως μαθηματικό κατασκεύασμα και αφετέρου ως όχημα των νοηματικών (εννοιολογικών, σημασιακών) αναπαραστάσεων που υποστηρίζουν και μη γλωσσικές νοητικές διεργασίες (γλώσσα της σκέψης).

Στις ενότητες που ακολουθούν περιγράφεται καθεμιά από τις τρεις προσεγγίσεις και παρουσιάζεται ο μεθοδολογικός χαρακτήρας που διέπει την παραγωγή γνώσης στα πλαίσιά της. Η έμφαση δίνεται στο επεξηγηματικό πεδίο και το είδος των στοιχείων που συνεισφέρει στο γενικότερο κι ευρύτερο πρόγραμμα της γνωσιακής επιστήμης.

Θα πρέπει να διευκρινιστεί ότι στόχος δεν είναι η παρουσίαση της σύγχρονης μορφής των πεδίων αυτών αλλά ούτε και η κριτική στην κάθε προσέγγιση. Σήμερα πλέον τα όρια μεταξύ κλάδων και επιστημονικών πεδίων στις γνωσιακές επιστήμες γενικά είναι ιδιαίτερα δυσδιάκριτα, και η σύγχυση ενισχύεται από τον έντονα διεπιστημονικό χαρακτήρα των περισσότερων πανεπιστημιακών τμημάτων, ερευνητικών ομάδων και προγραμμάτων σπουδών. Η συγχώνευση και συνεργασία κλάδων και επιστημονικών ειδικοτήτων αποτελεί ιδιαίτερα θετική εξέλιξη για τις πιθανότητες θετικής έκβασης στη μελέτη της νόησης, λόγω του ιδιαίτερα πολύπλοκου χαρακτήρα της και του εύρους των πεδίων που καλύπτει. Όμως είναι χρήσιμο να αναγνωριστούν, στα πλαίσια της γενικότερης συνεισφοράς, μια σειρά από θεμελιώδεις μεθοδολογικές καταβολές που συμπεριλαμβάνονται στη σύγχρονη γνωσιακή επιστήμη. Έτσι, στο μέρος αυτό παρουσιάζεται ο ιδιαίτερος πυρήνας κάθε μεθοδολογίας μέσα από μια καρικατούρα του κλάδου που τον κληροδότησε στη γνωσιακή επιστήμη. Αυτό σημαίνει ότι οι περιγραφές που ακολουθούν δεν εκφράζουν με λεπτομέρεια και ακρίβεια τη σημερινή κατάσταση στους επιμέρους κλάδους. Η επισήμανση των διαφορετικών μεθοδολογικών καταβολών έχει στόχο την ανάδειξη της συμπληρωματικότητας μεταξύ τους και της ολοκληρωμένης προσέγγισης που μπορεί να προκύψει από το συνδυασμό τους.

9 Γνωστική ψυχολογία

Η γνωστική ψυχολογία μπορεί να θεωρηθεί ως ένα επαναστατικό κίνημα μέσα στην ψυχολογία, ειδικότερα στην πειραματική, το οποίο έστρεψε τη θεωρητική προσοχή στο εσωτερικό του νου. Όμως ήδη από τη γέννηση της γνωστικής ψυχολογίας ως αυτόνομου κλάδου μελέτης υπήρχαν στον πυρήνα της σαφή τα στοιχεία που απέκτησαν γενικότερο ρόλο στη σύγχρονη γνωσιακή επιστήμη. Τα στοιχεία αυτά αφορούν στην έννοια της πληροφορίας και της επεξεργασίας και εντοπίζονται στη σχέση της ψυχολογίας με την πληροφορική και την υπολογιστική επιστήμη. Αρχικά ως κυβερνητική και στη συνέχεια ως τεχνητή νοημοσύνη, η μελέτη των συστημάτων επεξεργασίας και μετασχηματισμού πληροφορίας έπαιξε σημαντικό ρόλο στη θεωρητική διαμόρφωση της οντολογίας της γνωστικής ψυχολογίας.

Όπως φαίνεται από την ονομασία, γνωστική ψυχολογία είναι η ψυχολογία που αφορά στη γνώση. Η γνώση εδώ νοείται ως πληροφορία, δηλαδή ως διαφοροποίηση – από το τυχαίο, από το προηγούμενο, ή από το μέσο όρο. Η γνωστική ψυχολογία είναι ένας κλάδος που μελετά νοητικές διεργασίες εξετάζοντάς τες ως διεργασίες επεξεργασίας πληροφορίας, δηλαδή απόκτησης, διατήρησης, μετασχηματισμού και χρήσης της γνώσης. Η γνώση είναι στοιχεία πληροφορίας που επηρεάζουν τη λειτουργία του συστήματος της νόησης και επηρεάζονται από αυτή. Η γενικευμένη θεώρηση της γνώσης ως πληροφορίας και η έννοια που δίνεται στην πληροφορία μέσα από το μαθηματικό κλάδο της υπολογιστικής επιστήμης παρέχει στη γνωστική ψυχολογία τα εργαλεία για τη μελέτη ενός μεγάλου εύρους ψυχολογικών φαινομένων. Το πειραματικό μέρος της γνωστικής ψυχολογίας παρέμεινε κοντά σε φαινόμενα που εξετάζε και η προγενέστερή της συμπεριφοριστική προσέγγιση, όμως η θεωρητική έμφαση στην ερμηνεία των φαινομένων αυτών ήταν εντελώς διαφορετική. Η συμπεριφοριστική θεώρηση έδινε μεγάλο βάρος στη μάθηση, θεωρούσε όμως τη μάθηση ως αντιστοιχία αποκρίσεων σε ερεθίσματα. Στη γνωστική γλώσσα θα λέγαμε για αντιστοιχία εισροής-εκροής (ή εισόδου-εξόδου) σε σχέση με το νοητικό σύστημα. Η γνωστική ψυχολογία διατήρησε κάποιο ενδιαφέρον για τη μάθηση αλλά εστίασε τη μελέτη στο εσωτερικό του νοητικού συστήματος, όπου όχι μόνο συντελείται η μάθηση αλλά κυρίως όπου τα αποτελέσματα της μάθησης διαμορφώνουν και επηρεάζουν τη νοητική λειτουργία. Έτσι η μελέτη της μάθησης στράφηκε από παρατηρήσιμες αιτίες και παρατηρήσιμα αποτελέσματα προς μη παρατηρήσιμες οντότητες χειρισμού και δόμησης πληροφοριών.

9.1 Θεματικές ενότητες

Με δεδομένη τη θεωρητική βάση της πληροφορίας, ήταν φυσικό η γνωστική ψυχολογία να εστιάσει σε μηχανισμούς πρόσληψης πληροφοριών από το περιβάλλον προς το νου, μηχανισμούς συγκράτησης και επεξεργασίας πληροφοριών εντός του νου, και σε μηχανισμούς δράσης προς το περιβάλλον βάσει των πληροφοριών. Η θεμελιώδης υπόθεση της πληροφοριακής ροής παρέχει έτσι την κλασική θεματολογία της γνωστικής ψυχολογίας: αντίληψη, μνήμη, σκέψη, δράση. *Αντίληψη* είναι η διεργασία μετασχηματισμού ακατέργαστης αισθητηριακής πληροφορίας, όπως αυτή λαμβάνεται από το περιβάλλον, σε μορφές κατάλληλες για εσωτερική αναπαράσταση, επεξεργασία και αποθήκευση. Το προϊόν της αντιληπτικής διεργασίας ονομάζεται *αντίλημμα* και συνήθως ορίζεται βάσει των αναπαραστασιακών ιδιοτήτων του νου κατά περίπτωση. Η *μνήμη* αφορά στην ικανότητα διατήρησης πληροφοριών και περιλαμβάνει επιμέρους διεργασίες κωδικοποίησης και αποθήκευσης, συγκράτησης και διατήρησης, καθώς και ανάσυρσης. Η *σκέψη* περιλαμβάνει διεργασίες *λήψης αποφάσεων*, δηλαδή επιλογής μεταξύ εναλλακτικών δράσεων, *επίλυσης προβλημάτων*, δηλαδή προσέγγισης μιας κατάστασης-στόχου από μια αρχική κατάσταση εκκίνησης, και *συλλογισμού*, δηλαδή εξαγωγής συμπερασμάτων από ένα σύνολο προκειμένων, με λογικό (αναγωγικό ή επαγωγικό), αναλογικό ή ευρετικό τρόπο. Τις βασικές αυτές διεργασίες συμπληρώνει, εστιάζει και καθοδηγεί η *προσοχή*, η οποία δρα επιλεκτικά στη ροή των πληροφοριών και την ενεργοποίηση των διεργασιών εντός ή μεταξύ επιπέδων.

Εκτός από τις γενικές αυτές διεργασίες που αφορούν στη ροή και επεξεργασία των πληροφοριών, στην κλασική γνωστική ψυχολογία θα συναντήσει κανείς τουλάχιστον δύο ακόμα μεγάλα κεφάλαια: Ένα αφορά στις *έννοιες* και τις σημασίες τους, καθώς και στη γνωστική κατηγοριοποίηση κατά τις διάφορες διεργασίες αντίληψης και μνήμης, και το άλλο αφορά στη *γλώσσα* και τις γλωσσικές λειτουργίες. Τα δύο αυτά θέματα φέρνουν την έννοια της πληροφορίας από το υπολογιστικό στο σημασιακό πεδίο, σε επαφή με τη γενικότερη εννοιολογική ή σημασιολογική οργάνωση που θεμελιώνει και υποστηρίζει την ανθρώπινη σκέψη και επικοινωνία. Είναι και αυτά κεφαλαιώδους σημασίας διότι για την ακριβέστερη κατανόηση των νοητικών φαινομένων οφείλει κανείς να παραδεχτεί την ύπαρξη αυστηρής δομής στις εσωτερικές αναπαραστάσεις των πληροφοριών. Η έλλειψη της εσωτερικής δομής αυτής ήταν και ένα από τα μεγαλύτερα προβλήματα του συμπεριφορισμού. Πρόκειται για την οργάνωση των πληροφοριών αναφορικά με τις σημασίες τους και τη σχέση μεταξύ τους στο περιβάλλον του οργανισμού καθώς και στη σχέση του οργανισμού με το περιβάλλον του. Πρόκειται δηλαδή για το ρόλο του σημασιακού περιεχομένου των αναπαραστάσεων, άρα των εννοιών. Όμως οι έννοιες δεν αφορούν μόνο στην οργανωμένη αποθήκευση των πληροφοριών στα πλαίσια της μνήμης αλλά αποτελούν και το υπόβαθρο των γλωσσικών λειτουργιών, τουλάχιστον όσον αφορά στη σκέψη και την επικοινωνία. Παρότι πολλοί γλωσσολόγοι υποστηρίζουν ότι ορισμένες γλωσσικές διεργασίες διακρίνονται θεμελιωδώς από τις (μη γλωσσικές) γνωστικές λειτουργίες, στο ζήτημα των εννοιών γίνεται προφανές ότι τα όρια μεταξύ γνωστικού και γλωσσικού συστήματος, αν υπάρχουν, είναι ιδιαίτερα ασαφή.

Τα θέματα αυτά καθορίζουν τη δομή κάθε εισαγωγικού εγχειριδίου κλασικής γνωστικής ψυχολογίας. Συνεπώς μπορεί κανείς με σαφήνεια να οριοθετήσει έτσι το πεδίο που οι γνωστικοί ψυχολόγοι πιστεύουν ότι είναι της αρμοδιότητάς τους. Συχνά μπορεί κανείς να δει και ορισμένα άλλα θέματα να τραβούν την προσοχή στα πλαίσια της γνωστικής ψυχολογίας, όπως θέματα δράσης, ανάπτυξης, συναισθημάτων και κινήτρων κ.ά., χωρίς όμως να ξεφεύγουν από τη γενικότερη θεώρηση που διαγράφεται από την κεντρική αυτή θεματολογία. Τα τελευταία χρόνια, υπό το βάρος της διεπιστημονικότητας, η γνωστική ψυχολογία δείχνει μια τάση μετασχηματισμού σε *γνωσιακή νευροεπιστήμη*, με κύριο χαρακτηριστικό στην έμφαση της μελέτης εγκεφαλικών περιοχών και λειτουργιών τους. Έτσι, κεφάλαια σχετικά με τον εγκέφαλο κάνουν την εμφάνισή τους σε βιβλία γνωστικής κατά τα άλλα ψυχολογίας, δίνοντας το στίγμα της σύγχρονης μελέτης, χωρίς όμως να αλλοιώνουν το χαρακτήρα της κλασικής γνωστικής προσέγγισης, που είναι το συμπεριφορικό πείραμα και οι θεωρητικές νοητικές έννοιες που περιγράφουν τα αποτελέσματα.

9.2 Θεωρία και πείραμα

Η θεωρία της γνωστικής ψυχολογίας είναι εξαιρετικά ανεπτυγμένη και περιλαμβάνει πλήθος μη παρατηρήσιμων κατασκευασμάτων στα πλαίσια της θεώρησης αντίληψη-μνήμη-προσοχή που αναπτύχθηκε στην προηγούμενη ενότητα. Η μέθοδος της γνωστικής ψυ-

χολογίας είναι εμπειρική³⁴ και βασίζεται στο πείραμα και ιδιαίτερα στη μεθοδολογία της πειραματικής ψυχολογίας. Σε γενικές γραμμές, παράγει κανείς μια εμπειρική παρατήρηση ακολουθώντας μια πειραματική διαδικασία και στη συνέχεια περιγράφει και εξηγεί την παρατήρηση αυτή χρησιμοποιώντας όρους επεξεργασίας πληροφορίας. Βέβαια η θεώρηση της πληροφορίας είναι εκείνη που εξ αρχής οδηγεί στον πειραματικό σχεδιασμό και στην παραγωγή των εμπειρικών παρατηρήσεων, όπως συμβαίνει σε όλες τις εμπειρικές επιστημονικές μεθοδολογίες. Στόχος είναι η περιγραφή του νοητικού συστήματος με όσο το δυνατόν μεγαλύτερη λεπτομέρεια και ακρίβεια. Για την επίτευξη του στόχου αυτού αναπτύσσονται περίπλοκες θεωρητικές σχέσεις μεταξύ πληροφοριών και στη συνέχεια αυτές ελέγχονται πειραματικά ανάγοντας ένα επίπεδο πληροφορίας σε ερέθισμα και ένα άλλο σε απόκριση και χρησιμοποιώντας μια κατά βάση συμπεριφοριστική μεθοδολογία: παρουσιάζει δηλαδή κανείς στα άτομα που συμμετέχουν σε ένα πείραμα μια σειρά ερεθισμάτων με αυστηρά καθορισμένες ιδιότητες σε ένα αυστηρά καθορισμένο χρονοδιάγραμμα και καταγράφει με ακρίβεια τις αποκρίσεις τους δεδομένου του έργου που έχουν κληθεί να επιτελέσουν. Για παράδειγμα, μπορεί να παρουσιάζονται στοιχειώδη οπτικά ερεθίσματα όπως χρωματιστές βούλες στην οθόνη ενός υπολογιστή και το άτομο-«υποκείμενο»³⁵ πατά ένα πλήκτρο στο πληκτρολόγιο αμέσως μόλις αντιληφθεί την εμφάνιση ενός οποιουδήποτε ερεθίσματος. Σε μια άλλη πειραματική συνθήκη μπορεί το άτομο να πατά ένα πλήκτρο όταν αναγνωρίσει ερέθισμα συγκεκριμένου χρώματος, ή ένα πλήκτρο για ερέθισμα ενός σχήματος και διαφορετικό πλήκτρο για ερέθισμα άλλου σχήματος. Σε άλλου είδους πείραμα μπορεί να παρουσιάζονται ακουστικά λέξεις και το έργο να απαιτεί το πάτημα ενός πλήκτρου όταν μια λέξη αναγνωρίζεται ως τέτοια, ή όταν η σημασία της ανήκει σε κάποια σημασιολογική κατηγορία, κ.ά. Σε τέτοιες περιπτώσεις μπορεί κανείς να μετρήσει τον αριθμό των αποκρίσεων που προσδιορίζονται ως συναφείς: ορθές ή λανθασμένες απαντήσεις βάσει ενός προκαθορισμένου αντικειμενικού κριτηρίου, αποκρίσεις σε κάποιο συγκεκριμένο είδος ερεθίσματος ή αποκρίσεις κάποιου συγκεκριμένου τύπου. Μπορεί ακόμα να καταγραφεί με ακρίβεια ο χρόνος που μεσολάβησε από την παρουσίαση του ερεθίσματος μέχρι την απόκριση και από αυτό να συναχθούν συμπεράσματα όσον αφορά στη σχετική πολυπλοκότητα της εσωτερικής επεξεργασίας. Βέβαια υπάρχουν και άλλου είδους πειράματα, όπως αυτά στα οποία το άτομο που συμμετέχει καλείται να επιλύσει κάποιο πρόβλημα με μολύβι και χαρτί, να απομνημονεύσει ένα χάρτη ή μια ιστορία, να κατανοήσει και να εξηγήσει ένα θέμα κ.ά.

³⁴Λέγοντας «εμπειρικός» (empirical) εννοούμε ότι βασίζεται σε δεδομένα που προσλαμβάνουμε μέσω των αισθήσεων, δηλαδή σε παρατηρήσεις. Στις επιστήμες ο όρος αυτός πρακτικά ταυτίζεται με τον όρο «πειραματικός», δηλαδή οι παρατηρήσεις είναι ελεγχόμενες και διαρθρωμένες μετρήσεις στο πλαίσιο συστηματικού πειραματισμού. Αντιδιαστέλλεται με προσεγγίσεις βάσει λογικών ή θεωρητικών στοχασμών που δεν περιορίζονται από δημόσια παρατηρήσιμες ποσότητες, όπως λ.χ. στις ανθρωπιστικές σπουδές.

³⁵Ο όρος «υποκείμενο» (subject) χρησιμοποιείται στην ψυχολογία για να υποδηλώσει το άτομο που συμμετέχει σε ένα πείραμα και παράγει τις αποκρίσεις που αποτελούν τα πρωτογενή δεδομένα της εμπειρικής παρατήρησης. Σύμφωνα με τις οδηγίες της αμερικανικής ψυχολογικής εταιρείας ο όρος αυτός έχει πρόσφατα αντικατασταθεί από τη λέξη «συμμετέχων» (participant). Στα ελληνικά, που η λέξη «υποκείμενο» έχει και αρνητική χροιά, χρησιμοποιούμε συχνά και τον όρο «εθελοντής».

Γενικά, στη γνωστική ψυχολογία κατασκευάζουμε θεωρητικά σχήματα για να εξηγήσουμε τις παρατηρήσεις που καταγράφονται στα πειράματα. Τα θεωρητικά σχήματα περιλαμβάνουν μη παρατηρήσιμες οντότητες, οι οποίες ονομάζονται «εννοιολογικές κατασκευές», και σχέσεις αλληλεπίδρασης μεταξύ αυτών των οντοτήτων. Για παράδειγμα, λέγοντας «μνήμη» εννοούμε γενικά μια ιδιότητα χάρη στην οποία εξηγούμε την παρατήρηση ότι μπορούμε να διακρίνουμε μεταξύ κάποιας εικόνας που είδαμε στο παρελθόν και κάποιας άλλης που δεν είδαμε. Έτσι, η εννοιολογική κατασκευή της μνήμης, η οποία είναι μη παρατηρήσιμη, χρησιμοποιείται θεωρητικά για να δικαιολογήσει μια συμπεριφορική επίδοση, που είναι παρατηρήσιμη, δηλαδή τη σωστή επιλογή μεταξύ γνωστών και άγνωστων εικόνων. Οι διάφορες θεωρητικές κατασκευές συνδέονται με αιτιακές σχέσεις. Για παράδειγμα, διεργασίες προσοχής ελέγχουν την κατανομή επεξεργαστικών πόρων αντίληψης ερεθισμάτων έτσι ώστε κάποια ερεθίσματα να καταλήγουν σε αντιλήμματα. Η αναγνώριση γνωστών εικόνων προϋποθέτει κάποια διεργασία σύγκρισης αντιλημμάτων με αποθηκευμένες αναπαραστάσεις εικόνων, από την οποία μπορεί να προκύψει απόκριση. Έτσι, για τη σύνδεση μεταξύ των θεωρητικών επεξηγηματικών σχημάτων και των παρατηρήσιμων συμπεριφορών είναι απαραίτητο ένα ευρύτερο θεωρητικό σχήμα που θα συνδέει τις εννοιολογικές κατασκευές με παρατηρήσεις. Η εγκυρότητα της γνωστικής ψυχολογίας ως εμπειρικής επιστήμης βασίζεται στη στενή και συστηματική σύνδεση μεταξύ θεωρητικών οντοτήτων και αντικειμενικών μετρήσεων. Η θεωρητική ισχύς της απορρέει από τις διασυνδέσεις μεταξύ των εννοιολογικών κατασκευών και από τις πολλαπλές, πλεοναστικές διαδρομές πρόβλεψης παρατηρήσιμων βάσει αυτών των διασυνδέσεων.

9.3 Το γενικό γραμμικό μοντέλο και η μηδενική υπόθεση

Η γνωστική ψυχολογία καταγράφει ποσοτικές παρατηρήσεις οι οποίες προκύπτουν μέσα σε ελεγχόμενο πειραματικό περιβάλλον. Η γενική προσέγγιση στην επεξεργασία των πρωτογενών δεδομένων, δηλαδή των μετρήσεων, είναι η μαθηματική σύνδεση μέσα από ένα «γραμμικό μοντέλο», δηλαδή έναν τύπο που περιγράφει τις παρατηρούμενες τιμές (μετρήσεις) συναρτήσει των παραγόντων από τους οποίους εξαρτώνται. Για κάθε μέτρηση y (εξαρτημένη μεταβλητή) που υποθέτουμε πως εξαρτάται από κάποιον παράγοντα (ανεξάρτητη μεταβλητή) x , γράφουμε $y = \mu + x + \epsilon_x$, όπου συμβολίζουμε μ το μέσο όρο όλων των μετρήσεων, x τη μέση επίδραση του παράγοντα και ϵ_x το «υπόλοιπο» που είναι απρόβλεπτο. Αν είχαμε περισσότερους παράγοντες που θεωρητικά επηρεάζουν τη μέτρηση θα τους προσθέταμε και αυτούς αθροιστικά στην εξίσωση. Αυτό είναι το γενικό γραμμικό μοντέλο.³⁶ Αν ο παράγοντας (η ανεξάρτητη μεταβλητή) δεν είχε καμία επίδραση στη μέτρηση (την εξαρτημένη μεταβλητή), τότε ο τύπος θα ήταν απλώς $y = \mu + \epsilon_0$, δηλαδή το μόνο που θα μπορούσαμε να προβλέψουμε θα ήταν η μέση τιμή, ενώ κάθε διαφορά μέτρησης από τη μέση τιμή θα ήταν απρόβλεπτη και άρα μέρος του «σφάλματος» ϵ_0 . Αυτό σημαίνει πως ο

³⁶Λέγεται «γραμμικό» διότι η εξαρτημένη μεταβλητή είναι αθροιστική συνάρτηση των ανεξάρτητων μεταβλητών στην πρώτη δύναμη, δηλαδή δεν περιλαμβάνει μεταβλητές στο τετράγωνο, στον κύβο κλπ.

παράγοντας είναι προβλεπτικά χρήσιμος στο βαθμό που μειώνεται το απρόβλεπτο, δηλαδή όταν το ϵ_0 είναι μεγαλύτερο από το ϵ_x . Πόσο μεγαλύτερο; Η απάντηση σε αυτό το ερώτημα δίνεται από τη συνηθισμένη μέθοδο στατιστικής ανάλυσης της γνωστικής ψυχολογίας, που είναι η επαγωγική στατιστική με τις μεθόδους ανάλυσης διακύμανσης.

Το βασικό εργαλείο εξαγωγής συμπερασμάτων από το πείραμα είναι η ύπαρξη διαφορών μεταξύ πειραματικών συνθηκών. Λόγω της ανεξέλεγκτης πολυπλοκότητας, τόσο του νοητικού συστήματος όσο και των πειραματικών συνθηκών σε οποιαδήποτε περίπτωση, δεν είναι δυνατό να εξαχθεί συμπέρασμα από μια μεμονωμένη ομάδα μετρήσεων. Οι παράμετροι που μπορεί να επηρεάζουν την απόδοση ενός ατόμου σε ένα γνωστικό έργο είναι πάρα πολλές, συχνά αδύνατο να μετρηθούν, και πολλές πιθανόν εντελώς άγνωστες ακόμα. Ο μόνος τρόπος να καταλήξουμε σε αξιόπιστα συμπεράσματα είναι να κατασκευάσουμε τέτοιες πειραματικές συνθήκες ώστε ανά δύο να είναι απολύτως όμοιες και να διαφέρουν μόνο στη συγκεκριμένη παράμετρο την οποία μελετάμε. Η παράμετρος αυτή είναι «παράγοντας» (ανεξάρτητη μεταβλητή) στο γραμμικό μοντέλο. Έτσι, μπορεί τα άτομα που συμμετέχουν στο πείραμα να ακούν, για παράδειγμα, λέξεις που σημαίνουν ζώα σε μια συνθήκη και λέξεις που σημαίνουν εργαλεία σε άλλη συνθήκη, και να αποκρίνονται πατώντας ένα πλήκτρο όταν τις αναγνωρίζουν. Αν οι λέξεις είναι απολύτως όμοιες σε όλα εκτός από τη σημασία ζώου ή εργαλείου, δηλαδή έχουν τον ίδιο αριθμό συλλαβών, την ίδια συχνότητα εμφάνισης, την ίδια απεικονιστικότητα κλπ., τότε και μόνο μπορεί κανείς *συγκρίνοντας* τους χρόνους απόκρισης για τις δύο κατηγορίες λέξεων να συμπεράνει κάτι σχετικό με την κατηγορική αυτή διαφορά. Ο θεωρητικός παράγοντας «κατηγορία» έχει δύο τιμές (ζώα και εργαλεία) και εξετάζουμε αν επιδρά στην (αντικειμενική και παρατηρήσιμη) μέτρηση της ταχύτητας απόκρισης.

Δεδομένης της ανεξέλεγκτης διακύμανσης μεταξύ μετρήσεων, για πολλούς και διάφορους λόγους, συστηματικούς και μη, σε κάθε τέτοια πειραματική συνθήκη ο ερευνητής λαμβάνει μια κατανομή μετρήσεων. Έτσι για να μπορεί να αποφανθεί τελικά αν όντως υπάρχει διαφορά μεταξύ των πειραματικών συνθηκών είναι υποχρεωμένος να καταφύγει στην επαγωγική στατιστική. Υπολογίζει δηλαδή την πιθανότητα οι κατανομές μετρήσεων που έχει λάβει να μην προέρχονται από τη διαφορά στις πειραματικές παραμέτρους αλλά να αποτελούν προϊόν τυχαίας διακύμανσης. Με άλλα λόγια υπολογίζει την πιθανότητα οι όποιες διαφορές μεταξύ των πειραματικών συνθηκών να προέκυψαν κατά τύχη. Κατά σύμβαση, αν η πιθανότητα αυτή είναι μικρότερη από 5% ($p < .05$) τότε η διαφορά αναγνωρίζεται ως *στατιστικά σημαντική* και θεωρείται «πραγματική». Αν υπάρχει πιθανότητα μεγαλύτερη από 5% να προέκυψε κατά τύχη η παρατηρούμενη διαφορά στις μετρήσεις μας, τότε θεωρούμε ότι η διαφορά δεν οφείλεται στον πειραματικό παράγοντα αλλά στην τύχη, δηλαδή ότι δεν υπάρχει καμία διαφορά. Στην περίπτωση αυτή οι μέσοι όροι της μέτρησης στις διαφορετικές συνθήκες θεωρούνται στατιστικά μη διακρίσιμοι. Δεδομένου του αυθαίρετου κριτηρίου του 5%, είναι μαθηματικά αναγκαίο και εκ των προτέρων γνωστό ότι 5% όλων των μεμονωμένων θετικών αποτελεσμάτων θα είναι εσφαλμένα θετικό. Δηλαδή, ένα στα είκοσι στατιστικά σημαντικά πειραματικά ευρήματα θα είναι τυχαίο και

μη αξιόπιστο, εξ ου και η σημασία της επανάληψης μετρήσεων και πειραμάτων και της επιβεβαίωσης αποτελεσμάτων με νέες μετρήσεις σε άλλες συνθήκες. Δηλαδή, όταν βρούμε ένα ενδιαφέρον αποτέλεσμα πρέπει να επαναλάβουμε το πείραμα με διάφορες παραλλαγές μέχρις ότου βεβαιωθούμε ότι τα ευρήματα είναι αξιόπιστα και δεν πρόκειται για το αναπόφευκτο σφάλμα τυχαιότητας.³⁷

Κάθε στατιστικά σημαντική διαφορά μεταξύ πειραματικών συνθηκών ερμηνεύεται στα πλαίσια ενός μοντέλου επεξεργασίας πληροφορίας ως διαφορά διεργασιών ή διαφορά ιδιοτήτων των πληροφοριακών στοιχείων (ερεθισμάτων ή εσωτερικών αναπαραστάσεων). Η ύπαρξη διαφοράς, επειδή τεκμηριώνεται στατιστικά με την απόρριψη της τυχαιότητας, δεν μπορεί να είναι ποσοτική. Δηλαδή το μέγεθος της διαφοράς μεταξύ πειραματικών συνθηκών δεν αντιστοιχεί σε μέγεθος διαφοράς στο θεωρητικό μοντέλο αλλά σε βαθμό βεβαιότητας στο αν υπάρχει σε θεωρητικό επίπεδο κάποια διαφορά ή όχι. Συνεπώς στη συντριπτική πλειονότητα των περιπτώσεων οι εμπειρικές βάσεις των θεωριών της γνωστικής ψυχολογίας είναι ποιοτικές και όχι ποσοτικές (μπορούμε να ξέρουμε ότι δυο διεργασίες διαφέρουν αλλά όχι πόσο διαφέρουν). Επίσης δεν είναι δυνατόν να γνωρίζουμε με βεβαιότητα ότι δυο πειραματικές συνθήκες είναι απολύτως ισοδύναμες. Ίσως δεν μπορούμε να απορρίψουμε στατιστικά την πιθανότητα η οποιαδήποτε διαφορά μεταξύ τους να οφείλεται στην τύχη, αυτό όμως δεν μπορεί να αποδοθεί σε συγκεκριμένη θεωρητική ή μεθοδολογική ισοδυναμία και μπορεί να αντανakλά έλλειψη ευαισθησίας στις μετρήσεις μας. Άρα ο πειραματικός σχεδιασμός πρέπει να είναι τέτοιος ώστε να επιτρέπει σε διαφορές μεταξύ συνθηκών να ερμηνευτούν με τρόπο που να βγάζει νόημα στα πλαίσια μιας θεωρητικής κατασκευής.

Ανακεφαλαιώνοντας, τα πρωτογενή δεδομένα της γνωστικής ψυχολογίας είναι παρατηρήσεις (μετρήσεις) συμπεριφοράς όπως για παράδειγμα ακρίβεια απόκρισης (λ.χ. αριθμός σωστών απαντήσεων/επιλογών) ή ταχύτητα απόκρισης (χρόνος που μεσολαβεί από την παρουσίαση κάποιου ερεθίσματος μέχρι την καταγραφή της απόκρισης). Τα ευρήματα της γνωστικής ψυχολογίας είναι *στατιστικά σημαντικές διαφορές μεταξύ συνθηκών* σε κατάλληλα σχεδιασμένα πειράματα. Με βάση αυτά τα ευρήματα η γνωστική ψυχολογία συνεισφέρει στο οικοδόμημα της γνωσιακής επιστήμης θεωρητικά σχήματα για μη παρατηρήσιμες νοητικές οντότητες (αναπαραστάσεις και διεργασίες).

9.4 Γενίκευση και ο μέσος νους

Η γνωστική ψυχολογία εξετάζει ομάδες ατόμων, καταγράφοντας τις αποκρίσεις τους σε ερεθίσματα κάτω από ελεγχόμενες συνθήκες. Η στατιστική ανάλυση επιτρέπει τη γενίκευση των συμπερασμάτων, δηλαδή υποστηρίζει την υπόθεση ότι θα βρίσκαμε το ίδιο αποτέλεσμα (τις ίδιες στατιστικά σημαντικές διαφορές) αν κάναμε το ίδιο πείραμα σε άλλα άτομα.

³⁷Προσοχή! Η συνηθισμένη στατιστική ανάλυση δεν μας λέει την πιθανότητα επανάληψης του ευρήματος σε νέο δείγμα, παρά μόνο την πιθανότητα να οφείλεται στην τύχη το εύρημά μας. Η πιθανότητα επανάληψης είναι εντελώς διαφορετικό πράγμα, και μπορεί στατιστικά να είναι πολύ μικρότερη από 95%.

Σε ποια άτομα όμως; Από στατιστική σκοπιά δεν μπορούμε να γενικεύσουμε αδιακρίτως σε όλον τον ανθρώπινο πληθυσμό του πλανήτη, παρά μόνο σε εκείνον τον πληθυσμό για τον οποίο τα άτομα που έλαβαν μέρος στο πείραμα αποτελούν *αντιπροσωπευτικό* δείγμα. Δηλαδή στον πληθυσμό του οποίου τα χαρακτηριστικά έχουμε πλήρως αναπαραγάγει μέσα στο δείγμα του πειράματος. Στην πράξη, επειδή δεν είναι δυνατό να γνωρίζουμε όλα τα πιθανά γνωρίσματα των ανθρώπων που μπορεί να επηρεάσουν τη συμπεριφορά τους σε μια πειραματική διαδικασία, η αντιπροσωπευτικότητα επιδιώκεται μέσω της τυχαιότητας και της επάρκειας. Δηλαδή, θα έπρεπε να επιλέγουμε με απολύτως τυχαίο τρόπο μεμονωμένα άτομα από έναν πληθυσμό (π.χ. με κληρωτίδα, όπως στα τυχερά παιχνίδια) μέχρις ότου έχουμε τον αναγκαίο αριθμό ατόμων από τα δεδομένα των οποίων να προκύπτουν αξιόπιστες μετρήσεις. Στην περίπτωση αυτή τα αποτελέσματα μπορούν βάσιμα να γενικευτούν στον πληθυσμό δειγματοληψίας, δηλαδή στο σύνολο των ανθρώπων από το οποίο επιλέχτηκε τυχαία το δείγμα.

Στην πράξη αυτό δεν γίνεται σχεδόν ποτέ, αλλά χρησιμοποιούμε όποιο δείγμα είναι πρόχειρο, συνήθως φοιτητές, αν δεν συντρέχει ειδικός λόγος επιλογής άλλων ατόμων. Ο λόγος για τον οποίον η πρακτική αυτή είναι αποδεκτή είναι η υπόθεση του μέσου νου. Δηλαδή, υποθέτουμε πως όλοι οι άνθρωποι έχουν περίπου τις ίδιες νοητικές λειτουργίες. Αυτό δεν σημαίνει ότι είναι ακριβώς ίδιες ούτε ότι αναμένουμε ακριβώς τα ίδια αποτελέσματα από όλους - αναγνωρίζουμε την ύπαρξη τόσο ενδοατομικών όσο και τις διατομικών διαφορών, οι οποίες μπορεί να προέρχονται από τυχαιότητα, συγκυριακά ή συστηματικά αίτια. Όμως, θεωρούμε πως σε επίπεδο νοητικού μοντέλου, δηλαδή απλουστευμένης θεωρητικής περιγραφής, ισχύουν για όλους τους ανθρώπους οι ίδιες εννοιολογικές κατασκευές (π.χ. οπτική αντίληψη, οπτική μνήμη κλπ.) και οι ίδιες σχέσεις μεταξύ τους. Έτσι, αν φτιάχναμε μια θεωρία για τη νοητική λειτουργία κάθε ατόμου χωριστά, αυτή θα περιείχε τις ίδιες οντότητες και τις ίδιες αλληλεπιδράσεις τους, παρότι θα υπήρχαν ποσοτικές διαφοροποιήσεις μεταξύ ατόμων. Αυτή η υπόθεση, ότι όλοι οι άνθρωποι έχουν περίπου τον ίδιο νου (στο βαθμό λεπτομέρειας που μπορεί να προσεγγίσει μια θεωρητική περιγραφή), είναι η υπόθεση του μέσου ανθρώπου, ή εν προκειμένω του μέσου νου. Η υπόθεση αυτή είναι ο λόγος για τον οποίο νομιμοποιούμαστε να υπολογίζουμε το μέσο όρο αποκρίσεων διαφορετικών ατόμων (εφόσον υποθέτουμε ότι όλες οι αποκρίσεις προέρχονται από ίδιου τύπου νοητικό μηχανισμό), καθώς και ο λόγος για τον οποίο κατασκευάζουμε μια και μοναδική νοητική θεωρία για κάθε φαινόμενο, η οποία να ισχύει για όλους. Έτσι, στην πράξη γενικεύουμε πέρα από όσο επιτρέπει η στατιστική, όχι μόνο στον πληθυσμό δειγματοληψίας αλλά στο ανθρώπινο είδος.³⁸

Όπως και με άλλες υποθέσεις εργασίας, η ιδέα του μέσου νου διευκολύνει τη δουλειά μας και επιτρέπει καταρχήν τη μελέτη βολικών δειγμάτων, μένει όμως να αποδειχθεί κατά περίπτωση αν είναι βάσιμη η γενίκευση σε όλων των ειδών τις νοητικές λειτουργίες, ή

³⁸Όταν χρειάζεται, κάνουμε ειδικές εξαιρέσεις, π.χ. δεν συνυπολογίζουμε δεδομένα από άτομα με αισθητηριακή δυσλειτουργία ή νευρολογική-ψυχιατρική νόσο, διότι σε αυτές τις περιπτώσεις φαίνεται πως μπορεί να υπάρχουν διαφορετικές νοητικές λειτουργίες.

αν για κάποιες θα πρέπει να γίνει ειδική διαχείριση των ενδοατομικών και διατομικών διαφορών, που μπορεί να προέρχονται από γενετικούς και περιβαλλοντικούς παράγοντες.

9.5 Μοντέλα

Ως αποτέλεσμα των πειραματικών δεδομένων, η γνωστική ψυχολογία παράγει διαγράμματα ροής και μετασχηματισμού της πληροφορίας, στα οποία απεικονίζεται η διαδρομή και οι ιδιότητες της πληροφορίας σε κάθε περίπτωση και σε κάθε «στάδιο» επεξεργασίας. Η θεωρητική κατανόηση των διεργασιών είναι και αυτή διαγραμματική και πάντα σχετική (μεταξύ συνθηκών, μεταξύ παραμέτρων, ή μεταξύ κατηγοριών). Τα διαγραμματικά αυτά συστήματα με τις θεωρητικές κατασκευές της γνωστικής ψυχολογίας αποτελούν και τα μοντέλα που παράγει ο κλάδος για την περιγραφή και κατανόηση των γνωστικών λειτουργιών. Ένας βασικός μεθοδολογικός στόχος του κλάδου είναι ο διαχωρισμός και η κατηγοριοποίηση των λειτουργιών, όχι όμως για λόγους ταξινομικούς αλλά για την ακριβέστερη περιγραφή της ουσιαστικής νοητικής λειτουργίας. Για παράδειγμα, στη μελέτη μνημονικών διεργασιών παρατηρεί κανείς ότι κάποια έργα διακρίνονται από ορισμένα χαρακτηριστικά μνημονικής ικανότητας (χωρητικότητα, διάρκεια, είδος εσωτερικής κωδικοποίησης) ενώ άλλα έργα παρουσιάζουν διαφορετικό χαρακτήρα στα ίδια αυτά κριτήρια. Βάσει τέτοιων ομάδων παρατηρήσεων μπορεί κανείς να διακρίνει «είδη μνήμης» ως λειτουργικά υποσυστήματα, ώστε να εξηγήσει καλύτερα ένα μεγαλύτερο εύρος φαινομένων. Το ζητούμενο είναι να επιτυγχάνεται μεγάλη αύξηση στην επεξηγηματική ισχύ με μικρή αύξηση στη θεωρητική πολυπλοκότητα: αν αντί ένα γενικό μηχανισμό μνήμης θεωρήσουμε δύο είδη μνήμης, τη μακρόχρονη και τη βραχύχρονη, τότε πετυχαίνουμε καλύτερη περιγραφή εκατοντάδων πειραματικών ευρημάτων. Συνεπώς η συγκεκριμένη θεωρητική κατασκευή θεωρείται δικαιολογημένη και γίνεται αποδεκτή. Με τη λογική αυτή καταρτίζονται και τα διαγράμματα στις περισσότερες περιπτώσεις.

Ένα χαρακτηριστικό της κλασικής γνωστικής ψυχολογίας είναι ότι τα θεωρητικά μοντέλα που κατασκευάζονται τείνουν να χαρακτηρίζουν σχετικά απομονωμένες ομάδες πειραμάτων. Αυτό δε σημαίνει ότι για κάθε πέντε πειράματα φτιάχνει κανείς και νέο θεωρητικό μοντέλο αλλά ότι ο τρόπος που συλλαμβάνεται ένα ειδικό φαινόμενο είναι ανεξάρτητος από τον τρόπο που συλλαμβάνεται κάποιο άλλο. Ο βαθμός στον οποίο, για παράδειγμα, ένα μοντέλο ελέγχου των οφθαλμοκινήσεων κατά την ανάγνωση μπορεί να δεσμεύσει ένα μοντέλο οπτικής προσοχής για τον εντοπισμό σχημάτων, είναι αρκετά περιορισμένος, έως μηδενικός. Κάθε τέτοιου είδους θεωρητικό μοντέλο μπορεί να υποστηρίζεται από εκατοντάδες πειράματα και οι προβλέψεις του να βρίσκονται σε πολύ καλή συμφωνία με τα πειραματικά αποτελέσματα. Δεν υπάρχει όμως στην κλασική γνωστική ψυχολογία μια γενικευμένη τάση, μια κατεύθυνση προς κάποια κοινή αρχιτεκτονική του νου. Έτσι, ακόμα και διαχωρισμοί της μνήμης, όπως ρητή και άρρητη, παραμένουν αποκομμένοι ακόμα και από άλλους μνημονικούς διαχωρισμούς, όπως σημασιολογική και επεισοδίων, ή βραχύχρονη και μακρόχρονη, ενώ οι σχέσεις μεταξύ διαφορετικών «κεφαλαίων», όπως αντίληψη, μάθηση και προσοχή, είναι σε μεγάλο βαθμό ασαφείς. Αυτό βέβαια δε σημαίνει

σε καμία περίπτωση ότι τα θεωρητικά κατασκευάσματα της γνωστικής ψυχολογίας είναι ανυπόστατα, ούτε και ότι δε γίνονται προσπάθειες ενοποίησης και συλλογικής θεώρησης. Όμως οι προσπάθειες αυτές μάλλον εντάσσονται στα σύγχρονα ρεύματα της γνωστικής ψυχολογίας που πλέον είναι δύσκολο να διαχωριστούν από τη γνωσιακή επιστήμη και τη γνωσιακή νευροεπιστήμη, μια και η διεπιστημονικότητα της εποχής μας έχει σχεδόν γκρεμίσει τα όρια μεταξύ πολλών σχετιζόμενων πεδίων.

10 Νευροεπιστήμες

Ο δεύτερος άξονας της γνωσιακής επιστήμης προέρχεται από το χώρο της βιολογίας. Διαφέρει ριζικά από τον πρώτο άξονα, αυτόν της γνωστικής ψυχολογίας, σε όλα τα επίπεδα: θεωρητικά, μεθοδολογικά, στον παραδοσιακό του στόχο και στο είδος της γνώσης που παράγει. Οι νευροεπιστήμες συνεισφέρουν στη γνωσιακή επιστήμη ως μελέτη του υλικού υποστρώματος. Οι νευροεπιστήμες είναι λοιπόν χρήσιμες για τη μελέτη των νοητικών διεργασιών στο βαθμό που η δομή και λειτουργία του οργάνου που υλοποιεί το νου δίνει στοιχεία για τη λειτουργία του νου καθαυτού. Παρότι δεν είναι θεωρητικά εντελώς απαραίτητη η κατανόηση της λειτουργίας του εγκεφάλου για την περιγραφή της νόησης, εξαιτίας της πολυπλοκότητας της νόησης θεωρείται χρήσιμο από πολλούς να αντληθούν στοιχεία και από το βιολογικό χώρο, συμπληρωματικά των αμιγώς νοητικών, όποια κι αν είναι αυτά. Η συνεισφορά των νευροεπιστημών στη γνωσιακή επιστήμη βασίζεται θεμελιωδώς στην αναγωγιστική υπόθεση της αντιστοιχίας επιπέδου και στηρίζεται μεθοδολογικά στην ιδέα του λειτουργικού εντοπισμού.

10.1 Επίπεδα οργάνωσης και αναγωγισμός

Το νευρικό σύστημα οργανώνεται σε ένα τεράστιο εύρος κλιμάκων: από το κατώτερο, μοριακό επίπεδο, μέχρι το ανώτερο επίπεδο του ολόκληρου εγκεφάλου ως συστήματος επιμέρους δομών. Σε μια πρόχειρη απαρίθμηση μπορούν να αναφερθούν τα επίπεδα: μοριακό, υποκυτταρικό (μεμβράνες, οργανίδια), κυτταρικό (νευρώνες), πολυκυτταρικό (τοπικά νευρωνικά δίκτυα), συστημικό (δίκτυα μεγάλης κλίμακας), δομικό (πυρήνες και διακριτές εγκεφαλικές δομές και περιοχές), νευρικό (εγκέφαλος, μέρη του κεντρικού και περιφερικού νευρικού συστήματος). Σε κάθε επίπεδο τα προς μελέτη χαρακτηριστικά διαφέρουν ριζικά από τα χαρακτηριστικά των άλλων επιπέδων. Οι διαφορές αφορούν συνολικά στο κάθε επίπεδο: οι λειτουργικές μονάδες, οι ουσιώδεις ποσότητες-μεταβλητές, ο τρόπος και τα χαρακτηριστικά λειτουργίας και αλληλεπίδρασης, τα πάντα γύρω από κάθε επίπεδο ανήκουν μόνο στο επίπεδο αυτό και το προσδιορίζουν σε σχέση με τα άλλα.

Για ένα σύγχρονο επιστήμονα, κατά πάσα πιθανότητα μονιστή (υλιστή), είναι βέβαιο ότι η λειτουργία του νευρικού συστήματος καθορίζει και «εξηγεί» πλήρως τις νοητικές διεργασίες. Δηλαδή κάθε νοητικό γεγονός εκφράζει ή αποδίδεται αιτιακά σε κάποιο φυσικό γεγονός στο νευρικό σύστημα, χωρίς εξαιρέσεις. Η υπόθεση αυτή είναι τετριμμένη όσον

αφορά στο φυσικό (υλικό) της μέρος διότι ισοδυναμεί με την απόρριψη μη φυσικών «πνευματικών» ιδιοτήτων εκτός του υλικού κόσμου.³⁹ Όπως έχει αναφερθεί, αυτή η υπόθεση δεν επαρκεί για να στοιχειοθετήσει νοητικού τύπου εξήγηση, διότι μια τέτοια εξήγηση θα πρέπει να διατηρεί (ή να αντικαταστήσει κατάλληλα) τις ουσιώδεις νοητικές γενικεύσεις. Οι ουσιώδεις νοητικές γενικεύσεις αφορούν σε διεργασίες, ποσότητες και παραμέτρους ψυχολογικής σημασίας. Για παράδειγμα, στην οπτική αντίληψη θα πρέπει να εξηγηθεί η διεργασία κατασκευής ενός αντιλήμματος από ένα ερέθισμα. Θα πρέπει ακόμα να αποσαφηνιστούν οι ίδιες οι έννοιες «ερέθισμα» και «αντίλημμα», οι οποίες στο πεδίο της ψυχολογίας παραμένουν αρκετά ιμπρεσιονιστικές και σίγουρα αποτελούν θεωρητικά κατασκευάσματα ασαφούς σχέσης με τη φυσική πραγματικότητα.

Η ταυτόχρονη εξέταση και περιγραφή πολλών επιπέδων εντός του κλάδου των νευροεπιστημών δικαιολογείται και ενισχύεται στην πράξη από την πρόθεση και την επιτυχή έκβαση της αναγωγής μεταξύ επιπέδων. Λέγοντας «αναγωγή» εννοούμε την περιγραφή οντοτήτων, εννοιών και σχέσεων ενός επιπέδου με βάση οντότητες, έννοιες και σχέσεις άλλου επιπέδου. Πρόκειται για ένα πλέγμα αντιστοιχιών με το οποίο διατηρούνται ορθές περιγραφές και στα δύο επίπεδα. Για παράδειγμα, θεωρούμε ότι κάθε κυτταρική λειτουργία (εντός και μεταξύ κυττάρων) ανάγεται σε μια περιγραφή στο μοριακό επίπεδο, καθώς τα κύτταρα αποτελούνται από μόρια και οι χημικές και ηλεκτρικές σχέσεις μεταξύ μορίων επαρκούν πλήρως για να δικαιολογήσουν συγκεντρωτικά το φαινόμενο που παρατηρείται σε κυτταρικό επίπεδο. Για παράδειγμα, οι δυνάμεις που αναπτύσσονται μεταξύ μορίων συνιστούν «κυτταρική μεμβράνη» ελεγχόμενης διαπερατότητας. Με τον ίδιο τρόπο, κυτταρικές περιγραφές επαρκούν για να δικαιολογήσουν τα παρατηρούμενα φαινόμενα σε επίπεδο νευρικών δικτύων, καθώς η λειτουργία των δικτύων ανάγεται σε αλληλεπιδράσεις μεταξύ μεμονωμένων νευρώνων, κ.ο.κ. Αυτό δεν σημαίνει ότι καταργείται η ανάγκη περιγραφής σε διαφορετικά επίπεδα, διότι, για παράδειγμα, η έννοια της «κυτταρικής μεμβράνης» έχει επεξηγηματική ισχύ σε επίπεδο γενίκευσης πέρα και πάνω από το σύνολο των μορίων που απαρτίζουν τη μεμβράνη.

Με το ίδιο σκεπτικό θα μπορούσε κανείς να θεωρήσει ότι η σύνδεση μεταξύ νοητικών και εγκεφαλικών λειτουργιών αποτελεί μια διαθεωρητική αναγωγή, κατά την οποία τα νοητικά φαινόμενα θα δικαιολογούνται πλήρως από περιγραφή των αντίστοιχων βιολογικών λειτουργιών. Η επίτευξη ενός επεξηγηματικού στόχου σε νοητικό επίπεδο για τις νευροεπιστήμες προϋποθέτει την αντιστοίχιση μεταξύ ψυχολογικών και βιολογικών παραμέτρων και διεργασιών. Η διαθεωρητική αναγωγή, αυστηρά προσδιοριζόμενη, απαιτεί μια απόλυτη «ένα προς ένα» αντιστοιχία μεταξύ εννοιών και ποσοτήτων δύο θεωριών ώστε η μετάβαση από τη μία θεωρία στην άλλη να είναι αυτόματη, πλήρης, και επεξηγηματικά ισοδύναμη, διατηρώντας τις ειδικές περιπτώσεις και τις γενικεύσεις.⁴⁰ Αν πρόκειται για δύο θεωρίες επαρκώς προσδιορισμένες και επιβεβαιωμένες, τότε η διαθεωρητική αναγωγή δεν προ-

³⁹ Η υπόθεση περιλαμβάνει και την απόρριψη νοητικών λειτουργιών εκτός του νευρικού συστήματος, κάτι που είναι λιγότερο προφανές, αλλά τουλάχιστον με τα μέχρι στιγμής δεδομένα πρόκειται για μια αναμφισβήτητη αποδεκτή υπόθεση εργασίας.

⁴⁰ Αποκαλούμε «αναγωγισμό» την υπόθεση (ή πεποίθηση) ότι είναι εφικτή μια τέτοια αντιστοίχιση.

σθέτει επεξηγηματική ισχύ σε καμία από τις δύο, αλλά απλώς παρέχει γενίκευση μεταξύ πεδίων. Στην περίπτωση της γνωσιακής επιστήμης δεν ισχύουν οι προϋποθέσεις για κάτι τέτοιο, αλλά ούτε και ο στόχος της επαφής με τις αμιγώς βιολογικές νευροεπιστήμες είναι η αναγωγή. Στόχος της επαφής είναι η αλληλοϋποστήριξη και καθοδήγηση· πρόκειται δηλαδή για σχέση διεπαφής (interface), χάρη στην οποία θεωρητικές κατασκευές αφενός καθοδηγούν και αφετέρου αποσαφηνίζονται από βιολογικές παρατηρήσεις. Τελικός σκοπός είναι η περιγραφή των νοητικών λειτουργιών με νοητικούς όρους και η τεκμηρίωσή τους με εξηγήσεις στέρεα εδραιωμένες σε φυσικές παρατηρήσεις.

Η διεπαφή με τις νευροεπιστήμες απαιτεί μια μερική αναγωγή, κατά την οποία ορισμένες τουλάχιστον παράμετροι και λειτουργίες μεταξύ γνωσιακών θεωριών και βιολογικών παρατηρήσεων θα αντιστοιχιστούν μεταξύ τους. Στο σημείο αυτό είναι ιδιαίτερα σημαντική η υπόθεση της αντιστοιχίας επιπέδου, διότι για να πραγματοποιηθεί και να αποδώσει η επιθυμητή αναγωγή θα πρέπει να υπάρχει ένα επίπεδο περιγραφής της λειτουργίας του νευρικού συστήματος το οποίο να αντιστοιχεί καλά σε ένα νοητικό επίπεδο περιγραφής. Η σημασία των νευροεπιστημών για τη γνωσιακή επιστήμη βασίζεται στην ελπίδα ότι κάποιο από τα επίπεδα περιγραφής της λειτουργίας του νευρικού συστήματος θα αναφέρεται σε έννοιες και παραμέτρους που αντιστοιχούν σε ψυχολογικές έννοιες και παραμέτρους.

Αρχικά ίσως περίμενε κανείς ότι αναζητούνται από τις νευροεπιστήμες οι βιολογικές εκείνες έννοιες που αντιστοιχούν στις ήδη υπάρχουσες θεωρητικές κατασκευές της γνωστικής ψυχολογίας. Για παράδειγμα, να βρεθούν τα λειτουργικά αντίστοιχα της εικονικής και βραχύχρονης μνήμης, της οπτικής αναγνώρισης αντικειμένων κλπ. Η κατάρρευση της προσδοκίας αυτής ανοίγει το δρόμο για την ουσιαστική διεπαφή μεταξύ των κλάδων διότι δείχνει την αναγκαία κατεύθυνση τόσο για το επίπεδο βιολογικής περιγραφής όσο και για το επίπεδο ψυχολογικής θεωρίας. Από τα επίπεδα του νευρικού συστήματος που προαναφέρθηκαν, φαίνεται ότι η βασική υποψηφιότητα για διεπαφή τίθεται στο συστημικό, με τη δευτερεύουσα αλλά σημαντική συμβολή του πολυκυτταρικού και του δομικού. Δηλαδή το πιθανότερο σημείο επαφής μεταξύ νευροεπιστημών και γνωστικής ψυχολογίας είναι εκείνο όπου στοιχειώδεις ψυχολογικές διεργασίες αποδίδονται σε μεγάλης κλίμακας νευρωνικά δίκτυα στον εγκέφαλο. Η διαπίστωση αυτή οδηγεί σε αναθεώρηση της σύστασης των ψυχολογικών διεργασιών, ώστε να αναγνωριστούν ως στοιχειώδεις εκείνες που αφενός αντιστοιχούν σε συγκεκριμένα νευρωνικά δίκτυα και αφετέρου απαρτίζουν και συγκροτούν με ουσιαστικό τρόπο τις σύνθετες γνωστικές λειτουργίες. Η διαπίστωση αυτή οδηγεί επίσης την εστίαση της νευροεπιστημονικής έρευνας στο συστημικό επίπεδο. Με τη συμβολή και της κρίσιμης εντοπιστικής υπόθεσης, οι νευροεπιστήμες και η γνωστική ψυχολογία συγχωνεύονται στη γνωσιακή νευροεπιστήμη.

10.2 Εντοπιστική υπόθεση

Η βιολογία, μέσα από τη νευροεπιστήμη, συνεισφέρει στη γνωσιακή επιστήμη χάρη στη διεπαφή με αναγωγή στο κατάλληλο επίπεδο περιγραφής. Η επιτυχία του εγχειρήματος δεν εξαρτάται μόνο από την περιγραφή του νευρικού συστήματος στην επιθυμητή ανάλυ-

ση. Για να υπάρξει η αντιστοιχισή θα πρέπει κάθε μονάδα στη βιολογική περιγραφή να αποτελεί πράγματι και λειτουργική μονάδα για τις αντίστοιχες νοητικές διεργασίες. Θα πρέπει δηλαδή ένα νευρωνικό δίκτυο μεσαίας ή μεγάλης κλίμακας να βρίσκεται σε ένα σημείο ή μια περιοχή, την ίδια για κάθε εγκέφαλο, και να επιτελεί την ίδια διεργασία η οποία θα αντιστοιχεί πάντα στην ίδια θεωρητική κατασκευή της γνωστικής ψυχολογίας. Αυτό δεν είναι βέβαιο συμπέρασμα από την υπόθεση των επιπέδων. Ίσως η λειτουργία του νευρικού συστήματος να είναι όντως οργανωμένη στα προαναφερθέντα επίπεδα. Μπορεί όμως μια ψυχολογική γενίκευση να αποδίδεται σε δίκτυα μεγάλης κλίμακας που συγκροτούνται από διάσπαρτα, απομακρυσμένα τοπικά δίκτυα επεξεργασίας, και η κατανομή αυτή να παρουσιάζει μεγάλη διακύμανση μεταξύ ατόμων. Στην περίπτωση αυτή δεν υπάρχει ελπίδα διεπαφής, διότι ακόμα και αν μπορούσε να επιτευχθεί η αναγωγή θα ίσχυε για ένα μόνο άτομο. Αν, επιπλέον, η διασπορά δικτύων είναι αρκετά μεγάλη και αν τα τοπικά δίκτυα δεν αντιστοιχούν σε χρήσιμες λειτουργικές γενικεύσεις, τότε δεν μπορεί ούτε η αναγωγή να επιτευχθεί, ακόμα κι αν το συνολικό δίκτυο μεγάλης κλίμακας όντως αντιστοιχεί σε ψυχολογική διεργασία, διότι δεν θα μπορεί να αναγνωριστεί η θέση και ο ρόλος καθενός δικτύου στο σύστημα.

Ονομάζουμε «εντοπιστική» την υπόθεση εργασίας ότι (α) κάθε λειτουργική μονάδα του εγκεφάλου είναι χωρικά περιορισμένη και (β) κάθε στοιχειώδες τμήμα του εγκεφάλου επιτελεί μία μοναδική λειτουργία. Το πρώτο μέρος υπονοεί ότι υπάρχει μια περιγραφή των εγκεφαλικών (και άρα των αντίστοιχων νοητικών) λειτουργιών έτσι ώστε κάθε λειτουργία να επιτελείται από ένα τοπικό δίκτυο, δηλαδή από μια ομάδα νευρώνων, διασυνδεδεμένων μεταξύ τους, οι οποίοι βρίσκονται σε κοντινές θέσεις, συγκροτώντας ένα συμπαγές κομμάτι νευρικού ιστού. Το κομμάτι αυτό μπορεί να είναι ένας πυρήνας, μια περιοχή, ή μια φλοιική στήλη—η υπόθεση στη γενική της μορφή δεν δεσμεύεται ως προς το κατάλληλο μέγεθος που μπορεί να εξυπηρετήσει την εντοπισμένη περιγραφή. Το δεύτερο μέρος της υπόθεσης υπονοεί ότι υπάρχει ένα μέγεθος κόκκου, μια χωρική ανάλυση τέτοια ώστε αν κόψουμε κατάλληλα τον εγκέφαλο σε τέτοιου μεγέθους κομμάτια τότε κάθε κομμάτι θα κάνει πάντα μία και μόνο δουλειά και δεν θα περιλαμβάνει μέσα του στοιχεία που επιτελούν άλλες λειτουργίες. Αυτό σημαίνει ότι το νευρικό σύστημα γενικά, και ο εγκέφαλος ειδικότερα, αποτελούνται από δομικές και λειτουργικές μονάδες που ταυτίζονται μεταξύ τους, καθώς τα κύτταρα (νευρώνες) που επιτελούν διαφορετικές λειτουργίες δεν ανακατεύονται μεταξύ τους αλλά βρίσκονται σε διακριτές θέσεις ή περιοχές. Αν ισχύει η υπόθεση αυτή, τότε είναι βάσιμο να αναζητούμε την κατάλληλη *κλίμακα*, δηλαδή σε πόσο μεγάλα κομμάτια πρέπει να κόψουμε τον εγκέφαλο προκειμένου να απομονώσουμε στοιχειώδεις λειτουργίες, και την κατάλληλη *λειτουργική ανάλυση*, δηλαδή ποιες είναι οι λειτουργίες εκείνες που αποτελούν «στοιχειώδη τμήματα» των γενικότερων νοητικών λειτουργιών και επιτελούνται από μεμονωμένα εγκεφαλικά αρθρώματα, από ενιαίες δομές ή περιοχές.

Η υπόθεση του εντοπισμού λειτουργιών είναι πολύ παλιά και έχει περάσει από διάφορες μορφές και φάσεις στην ιστορία των νευροεπιστημών. Τελικά φαίνεται ότι κερδίζει διαρκώς έδαφος τις τελευταίες δεκαετίες, με την προϋπόθεση ότι θα επαναπροσδιοριστούν

οι στοιχειώδεις λειτουργίες. Πρόκειται για μια Πύρρειο νίκη των εντοπιστών, αφού αντί για «οπτική περιοχή» έχουμε ένα σωρό κουτάκια για την επεξεργασία διαφορετικών χαρακτηριστικών του οπτικού πεδίου, τα οποία δεν ξέρουμε πώς οδηγούν σε συνολικό αντίλημμα, ενώ αντί για «γλωσσική περιοχή» ή έστω «περιοχή παραγωγής λόγου» ή «περιοχή συντακτικής επεξεργασίας» δεν είναι ακόμα σαφές τι ακριβώς έχουμε. Υπάρχει σημαντική διατομική διακύμανση, η οποία διαφέρει μεταξύ λειτουργιών και συστημάτων. Αυτό σημαίνει πως ο εγκέφαλος κάθε ανθρώπου είναι σε κάποιο βαθμό διαφορετικός (όπως και κάθε άλλο όργανο, π.χ. η μύτη κάθε ανθρώπου), έτσι ώστε αφενός διακρίνονται γενικές ομοιότητες στο σχήμα και τη δομή, αφετέρου όμως υπάρχουν τεράστιες διαφορές στις λεπτομέρειες και δεν είναι προφανές πώς πρέπει να αντιστοιχιστεί ένας εγκέφαλος με κάποιον άλλον για να διαπιστωθεί αν ένας νευρώνας βρίσκεται στο ίδιο σημείο των δύο εγκεφάλων ή όχι. Δεν είναι δηλαδή καθόλου προφανές τι σημαίνει «ίδιο σημείο» μεταξύ διαφορετικών εγκεφάλων. Πέρα όμως από τις πρακτικές δυσχέρειες που συνεπάγεται η αποδοχή της τροποποιημένης εντοπιστικής υπόθεσης, αυτή παραμένει σημαντικότερη και ουσιαστικής σημασίας για την ένταξη των νευροεπιστημών στις γνωσιακές επιστήμες. Χωρίς τις υποθέσεις της (διαθεωρητικής) αναγωγής και του εντοπισμού (λειτουργιών), η γνωσιακή νευροεπιστήμη δεν έχει νόημα ως κλάδος, όπως δεν έχει νόημα γενικότερα και η ένταξη των νευροεπιστημών στη μελέτη του νου.

10.3 Μεθοδολογικά χαρακτηριστικά των νευροεπιστημών

Οι νευροεπιστήμες, ως βιολογικός κλάδος, είναι ελαφρύτες σε θεωρίες και βαρύτες σε παρατηρήσεις. Αντίθετα από τους ψυχολογικούς κλάδους, τα πάντα ή σχεδόν τα πάντα είναι ανοιχτά σε σχεδόν άμεση παρατήρηση. Εδώ όμως χρειάζονται κάποιες διευκρινίσεις. Όταν τα συμπεράσματα εξάγονται με βάση παρατηρήσεις παρουσία ή απουσία ενός αντιδραστηρίου σε εξαιρετικά περίπλοκες διαδικασίες («πρωτόκολλα»), μπορεί κανείς να αναρωτηθεί πόσο άμεση είναι τελικά η παρατήρηση. Η αμεσότητα γίνεται εμφανής σε αντιδιαστολή με τα (θεμελιωδώς μη παρατηρήσιμα) θεωρητικά κατασκευάσματα της γνωστικής ψυχολογίας. Για παράδειγμα, η έννοια της πρωτεΐνης και η σημασία της τρισδιάστατης (στερεοχημικής) δομής της για την ενζυματική της δράση μπορεί να πει κανείς ότι έχουν ξεφύγει από το στάδιο της αφηρημένης θεωρητικής κατασκευής και αφορούν σε συγκεκριμένα, δυνάμει χειροπιαστά, στοιχεία του φυσικού κόσμου. Ομοίως για το γενετικό υλικό: τα χρωμοσώματα είναι ορατά κατά την κυτταρική διαίρεση και οι υποθέσεις για τη φύση και τη λειτουργικότητα του γονιδιώματος επιβεβαιώνονται καθημερινά στη μοριακή κοπποραπτική της γενετικής μηχανικής.

Αντίθετα, η έννοια του γονιδίου παραμένει κυρίως θεωρητική, ειδικά μετά την ανακάλυψη της περίπλοκης σχέσης της με την αλληλουχία νουκλεοτιδικών βάσεων του γονιδιώματος. Η κατάσταση για τους υποδοχείς των νευροδιαβιβαστών και τους ιοντικούς διαύλους στην κυτταρική μεμβράνη είναι μάλλον κάπου ενδιάμεσα. Από την άποψη αυτή, το να μιλάμε για «άμεση παρατήρηση» είναι ίσως υπερβολικό και το να θεωρούμε τη βιολογία αθεωρητική επιστημονικά άδικο. Σε κάποιο βαθμό όμως, ακόμα και οι θεωρητι-

κές έννοιες της βιολογίας είναι διαθέσιμες προς παρατήρηση με μοναδικό περιορισμό τη διακριτική ικανότητα της τεχνολογίας. Με τη βελτίωση των τεχνικών μέσων και την εφευρετικότητα των ερευνητών καθίσταται δυνατή ολόένα και πιο λεπτομερειακή παρατήρηση. Οι πρωτεΐνες, το γενετικό υλικό και οι διάλυτοι ιόντων είναι όλα εκεί για να παρατηρηθούν όλο και πιο άμεσα, όσο επιτρέπει κάθε στιγμή η τεχνολογία. Έτσι και ο ίδιος ο στόχος της βιολογίας είναι βασικά παρατηρησιακός και αφορά σε φυσικές οντότητες. Η θεωρητική διάρθρωση της βιολογίας αφορά στη λειτουργία που επιτελούν τα παρατηρήσιμα καθώς και τον πιθανό σκοπό αυτής της λειτουργίας, υπό το πρίσμα της εξέλιξης.

Η αναφορά στην εξέλιξη είναι θεμελιακή για την κατανόηση των βιολογικών παρατηρήσεων αλλά και του σκεπτικού με το οποίο γίνονται και ερμηνεύονται οι παρατηρήσεις. Η αξιοποίηση της εξέλιξης ως θεμελιακού πλαισίου δεν είναι παρά η αναγνώριση ότι τα παρατηρούμενα χαρακτηριστικά των οργανισμών είναι κατά μέσο όρο εκείνα που συνεισέφεραν σε αρμοστικότητα στους προγόνους τους, μέσα στο συγκεκριμένο οικολογικό περιβάλλον τους, με αποτέλεσμα οι πρόγονοί τους να έχουν επιτυχημένη αναπαραγωγή (δηλαδή τους παρατηρούμενους σήμερα οργανισμούς) ενώ άλλοι οργανισμοί να αποτύχουν. Στην απλούστερη εκδοχή της, η παραδοχή αυτή είναι απλώς μια ταυτολογία: έχουν επιβιώσει και αφήσει απογόνους οι οργανισμοί εκείνοι που είχαν τα χαρακτηριστικά που προσδίδουν συγκριτικό πλεονέκτημα στην επιβίωση και αναπαραγωγή στις δεδομένες συνθήκες. Ως ερμηνευτικό πλαίσιο, η παραδοχή αυτή οδηγεί τον ερευνητή βιολόγο να εξετάζει δομές και λειτουργίες υπό το πρίσμα της χρησιμότητάς τους, δηλαδή να υποθέτει πως τα παρατηρούμενα φαινόμενα επιτελούν κάποια λειτουργία (έχουν κάποιο «σκοπό») που συντελεί, άμεσα ή έμμεσα, στην επιβίωση και αναπαραγωγή του οργανισμού. Αυτό δεν σημαίνει ότι οι οργανισμοί παρουσιάζουν βέλτιστη λειτουργικότητα ούτε ότι είναι άριστα προσαρμοσμένοι στο σημερινό περιβάλλον τους. Σημαίνει όμως πως τα επιμέρους χαρακτηριστικά τους δεν είναι εντελώς τυχαία αλλά αποτελούν απόρροια συγκεκριμένων εξελικτικών πιέσεων. Η υπόθεση αυτή μπορεί να μην είναι πάντοτε ακριβής (και ίσως δεν υπάρχει τρόπος να διακρίνουμε πότε είναι και πότε όχι), δίνει όμως μια σαφή υπόσταση και κατεύθυνση στη μελέτη των φαινομενικά χαοτικών αλληλεπιδράσεων στη βιολογική μικροκλίμακα.

Με άλλα λόγια, η βιολογία γενικά, και οι νευροεπιστήμες ειδικότερα, παράγουν παρατηρήσεις για φυσικά συστήματα. Η γνώση που προκύπτει δομείται για να περιγράψει τη φυσική υπόσταση και λειτουργία του νευρικού συστήματος. Με την έννοια αυτή η νευροεπιστημονική γνώση είναι *περιγραφική* ενώ η ψυχολογική γνώση είναι εμπειρική μεν, αλλά ουσιαστικά *θεωρητική*, διότι παράγει μοντέλα μη παρατηρήσιμων οντοτήτων. Αντίθετα, στη βιολογία ακόμα και τα μοντέλα αφορούν σε παρατηρήσιμα. Στην ψυχολογία, η ουσία του πειράματος είναι η παραγωγή πρωτογενών δεδομένων που ερμηνεύονται στη βάση θεωρητικών κατασκευών – αυτή η κατάσταση δεν μπορεί (και ούτε χρειάζεται) να αλλάξει στο πλαίσιο της ψυχολογίας. Αντίθετα, στις νευροεπιστήμες, η ουσία του πειράματος είναι η περιγραφή παρατηρήσιμων οντοτήτων, ακόμα και όταν παρεμβαίνουν μοντέλα και θεωρητικά κατασκευάσματα, διότι τότε η θεωρητική παρέμβαση θεωρείται θεμελιωδώς προσωρινή και όχι ουσιαστική.

10.4 Νευροαπεικόνιση

Η πιο συχνά χρησιμοποιούμενη μέθοδος παραγωγής βιολογικών στοιχείων που συνδέονται άμεσα με νοητικές διεργασίες είναι η νευροαπεικόνιση, δηλαδή η καταγραφή της λειτουργίας του νευρικού συστήματος κατά τη διάρκεια νοητικών έργων. Η καταγραφή μπορεί να γίνει με πολλούς τρόπους που εκμεταλλεύονται κυρίως τις ηλεκτρικές και μαγνητικές ιδιότητες της νευρωνικής δραστηριότητας καθώς και τις συνέπειες της δραστηριότητας στο μεταβολισμό και τη ροή του αίματος. Οι σημαντικότερες μέθοδοι νευροαπεικόνισης είναι:

Ηλεκτροδιακές καταγραφές Πρόκειται για άμεση αποτύπωση της ηλεκτρικής δραστηριότητας μεμονωμένων νευρώνων (single unit recordings), μιας μικρής ομάδας νευρώνων (multi-unit recordings), ή της συνολικής μέσης συγχρονισμένης δραστηριότητας (τοπικό πεδίο, local field potential) γύρω από το σημείο στο οποίο εισάγεται το ηλεκτρόδιο, ανάλογα με το μέγεθος του ηλεκτροδίου και τις συχνότητες καταγραφής. Με την ίδια διάταξη υπάρχει δυνατότητα και παροχής ηλεκτρικής διέγερσης, αντί καταγραφής της φυσικής διέγερσης, επηρεάζοντας έτσι τη νευρωνική λειτουργία στην περιοχή. Η μέθοδος αυτή δίνει ένα μέτρο του βαθμού νευρωνικής δραστηριότητας σε κάποιο σημείο του εγκεφάλου. Δεν χρησιμοποιείται στον άνθρωπο (παρά μόνο η διέγερση, για λόγους χαρτογράφησης, κατά τη διάρκεια χειρουργικών επεμβάσεων στον εγκέφαλο) λόγω δεοντολογικών περιορισμών. Χρησιμοποιείται πολύ σε ποντίκια, πιθήκους, και άλλα ζώα. Παρέχει πολύ υψηλή διακριτική ικανότητα τόσο στο χώρο όσο και στο χρόνο. Περιορίζεται μόνο στον αριθμό των ηλεκτροδίων που μπορούν να λειτουργούν συγχρόνως και συνεπώς στην έκταση του νευρικού ιστού που μπορεί να μελετηθεί. Κατά κανόνα οι καταγραφές της δραστηριότητας εξετάζονται σε συνάρτηση με κάποιο παρεχόμενο ερέθισμα ή με κάποια παρατηρούμενη συμπεριφορά του ζώου.

EEG (electroencephalography, ηλεκτροεγκεφαλογραφία). Πρόκειται για καταγραφή της ηλεκτρικής τάσης σε διάφορα σημεία του δέρματος πάνω στο κεφάλι, σε σχέση με κάποιο σημείο αναφοράς. Η ηλεκτρική τάση που καταγράφεται προέρχεται από τη *συντονισμένη* δραστηριότητα πολύ μεγάλων αριθμών νευρώνων σε εκτεταμένες περιοχές του εγκεφάλου. Η ασυντόνιστη δραστηριότητα, είτε σε χαμηλή είτε σε υψηλή διέγερση, παράγει αλληλοαναιρούμενες ηλεκτρικές τάσεις με αποτέλεσμα η καταγραφή να είναι πολύ μικρού πλάτους. Έτσι, ουσιαστικά το ΗΕΓ καταγράφει το βαθμό συγχρονισμού και όχι το βαθμό διέγερσης των νευρώνων. Η μέθοδος αυτή χρησιμοποιείται συχνά σε σχέση με ένα χρονικό σημείο αναφοράς το οποίο μπορεί να προκαλέσει συντονισμένη δραστηριότητα, όπως π.χ. ένα ερέθισμα. Τότε καταγράφονται *προκλητά δυναμικά* (evoked response potentials, ERP) με τα οποία παρακολουθούμε αδρά την πορεία και ακριβέστερα τη χρονική απόκριση του εγκεφάλου. Παρέχει πολύ υψηλή χρονική ανάλυση (χιλιοστά του δευτερολέπτου) αλλά εξαιρετικά φτωχή χωρική διακριτική ικανότητα διότι οι τάσεις που φτάνουν ως έξω στο δέρμα έχουν διέλθει από πολλά στρώματα διαφορετικών ηλεκτρικών ιδιοτήτων

και δεν μπορούν να αναλυθούν με ακρίβεια σε αρχικές συνιστώσες μέσα στον εγκέφαλο. Αυτό σημαίνει πως το ΗΕΓ αποτυπώνει πολύ αδρά την περιοχή όπου εστιάζεται ο αυξημένος συντονισμός που καταγράφεται, και χρειάζεται συμπληρωματικές καταγραφές με άλλες μεθόδους ώστε να βελτιωθεί ο εντοπισμός του σήματος.

MEG (magnetoencephalography, μαγνητοεγκεφαλογραφία· εναλλακτικά *μαγνητική απεικόνιση πηγής*, magnetic source imaging). Πρόκειται για καταγραφή των μαγνητικών πεδίων που παράγονται από τη συντονισμένη ηλεκτρική δραστηριότητα των νευρώνων. Ως προς την προέλευση του καταγραφόμενου σήματος μοιάζει πολύ με το ηλεκτροεγκεφαλογράφημα αλλά πλεονεκτεί στο ότι τα μαγνητικά πεδία επηρεάζονται λιγότερο από τα ηλεκτρικά πεδία κατά τη διέλευσή τους μέσα από τους ιστούς. Έτσι, χάρη σε ειδικές τεχνικές ανάλυσης των καταγραφών, η μέθοδος αυτή μπορεί να εντοπίσει πιθανές πηγές διέγερσης μέσα στον εγκέφαλο με πολύ υψηλότερη διακριτική ικανότητα από το ηλεκτροεγκεφαλογράφημα, ενώ διαθέτει εξίσου υψηλή χρονική ανάλυση. Μπορεί να χρησιμοποιηθεί σε σχέση με ερέθισμα και αντίστοιχη χρονική στιγμή αναφοράς ή σε σχεδίαση σύγκρισης μεταξύ συνθηκών. Δίνει σε κάθε χρονική στιγμή μια εκτίμηση για την ισχυρότερη πηγή ηλεκτρομαγνητικής διέγερσης μέσα στον εγκέφαλο, εντοπίζοντας έτσι μια υποτιθέμενη εστία συγχρονισμένης νευρωνικής δραστηριότητας. Συνδυάζεται με μαγνητική τομογραφία ώστε ο εντοπισμός του μαγνητικού συγχρονισμού να προσδιοριστεί ανατομικά σε συγκεκριμένη εγκεφαλική περιοχή. Η μέθοδος αυτή περιορίζεται από τις υποθέσεις που είναι απαραίτητες για την εκτέλεση των υπολογισμών καθώς και από την κατεύθυνση των αναδιπλώσεων του φλοιού (διαφορετική ισχύς σήματος από έλικες και αύλακες). Σε συνδυασμό με λειτουργική μαγνητική τομογραφία (fMRI, βλ. παρακάτω), μπορεί να δώσει λεπτομερειακή χρονική πληροφορία συμπληρωματικά προς τον ακριβέστερο χωρικό εντοπισμό της τελευταίας.

PET (positron emission tomography, τομογραφία εκπομπής ποζιτρονίων). Απαρχαιωμένη πλέον μέθοδος καταγραφής της τοπικής αιματικής ροής στον εγκέφαλο, η οποία εξαρτάται από το συνολικό επίπεδο νευρωνικής δραστηριότητας. Σε περιοχές έντονης νευρωνικής διέγερσης αυξάνεται η ροή του αίματος, το οποίο παρέχει το απαραίτητο οξυγόνο και τη γλυκόζη για τη λειτουργία των νευρώνων. Έτσι, αν στο αίμα υπάρχει ραδιενεργό οξυγόνο ή γλυκόζη μπορεί να εντοπιστεί η περιοχή αυξημένης συγκέντρωσής τους εξαιτίας της ακτινοβολίας που εκπέμπουν. Το ραδιενεργό υλικό χορηγείται ενδοφλεβίως πριν από την καταγραφή, και στη συνέχεια εκτελείται η πειραματική διαδικασία που θεωρούμε ότι εμπλέκει επιλεκτικά κάποιες εγκεφαλικές περιοχές. Επειδή καμία περιοχή του εγκεφάλου δεν είναι ποτέ εντελώς ανενεργή, για την απομόνωση των σχετικών περιοχών απαιτείται συνθήκη αναφοράς, κατά την οποία δεν εκτελείται το υπό μελέτη έργο. Έτσι, κατά την εκτέλεση του πειραματικού έργου αξιολογούνται οι *μεταβολές* στην τοπική αιματική ροή που προκαλούνται από αυτό, συγκρίνοντας στατιστικά τις καταγραφές μεταξύ συνθήκης αναφοράς και

πειραματικής. Η μέθοδος αυτή έχει μέτρια χωρική διακριτική ικανότητα και μηδενική χρονική ανάλυση, διότι χρειάζονται αρκετά λεπτά για να συγκεντρωθεί επαρκής ποσότητα ραδιενεργού ουσίας στις πιο ενεργές περιοχές. Επιπλέον, λόγω της ραδιενεργού ακτινοβολίας δεν είναι εφικτή η επανάληψη της καταγραφής, με αρνητικές επιπτώσεις στην αξιοπιστία των αποτελεσμάτων.

fMRI (functional magnetic resonance imaging, λειτουργική μαγνητική τομογραφία)⁴¹. Η πιο συχνή μέθοδος σήμερα στη γνωσιακή νευροεπιστήμη. Καταγράφει έμμεσα μεταβολές στην τοπική αιματική ροή, όπως και το PET, χωρίς όμως να απαιτείται ραδιενεργός ακτινοβολία διότι το fMRI εκμεταλλεύεται τις μαγνητικές ιδιότητες των μορίων που υπάρχουν στους ιστούς. Έτσι είναι ακίνδυνη μέθοδος και μπορεί να επαναληφθεί πολλές φορές και για πολύ χρόνο. Στην πραγματικότητα είναι ακόμα πιο έμμεση καταγραφή των αποτελεσμάτων της νευρωνικής λειτουργίας από τις προηγούμενες, καθώς το σήμα που φτάνει στο όργανο καταγραφής προέρχεται από την οξυγόνωση του αίματος, η οποία αυξάνεται όταν αυξάνεται η αιματική ροή, η οποία προκαλείται από αυξημένη νευρωνική δραστηριότητα μέσα σε χρονική περίοδο λίγων δευτερολέπτων. Όπως και άλλες μέθοδοι, απαιτεί και αυτή συνθήκη αναφοράς, ώστε να αξιολογηθούν τυχόν διαφορές στην τοπική αιματική ροή και να ερμηνευθούν ως προερχόμενες από διαφορετικής έντασης νευρωνική δραστηριότητα. Αυτό σημαίνει ότι τα αποτελέσματα περιορίζονται από δυσκολίες στην επιλογή της κατάλληλης συνθήκης. Συγκεκριμένα, είναι πολύ δύσκολο να σχεδιαστεί συνθήκη ελέγχου τέτοια ώστε να εμπλέκονται όλες οι λειτουργίες της πειραματικής συνθήκης εκτός από εκείνη που είναι υπό εξέταση, με τον ίδιο τρόπο όπως και στην πειραματική συνθήκη. Επιπλέον, η μέθοδος περιορίζεται από τη στατιστική διαδικασία της σύγκρισης, που είναι απαραίτητη προκειμένου να διαπιστωθεί αν η διαφορά στην καταγραφόμενη τοπική αιματική ροή είναι στατιστικά σημαντική. Όπως κάθε στατιστική σύγκριση, έτσι και αυτή εξαρτάται από το αυθαίρετο κριτήριο στατιστικής σημαντικότητας που θέτουν οι ερευνητές. Πάντως η λειτουργική μαγνητική τομογραφία έχει σχετικά καλή διακριτική ικανότητα στο χώρο, και με τους ολοένα ισχυρότερους μαγνήτες μπορεί σε ειδικές περιπτώσεις να φτάσει μέχρι το επίπεδο μεμονωμένων σιβάδων του εγκεφαλικού φλοιού, αν και η ανάλυσή της διαφέρει ανάλογα με την κλίση της επιφάνειας του φλοιού και το βάθος της καταγραφής. Η χρονική ανάλυση είναι μέτρια αλλά με ειδικές τεχνικές επεξεργασίας των καταγραφών μπορεί να εστιαστεί και σε μεμονωμένα ερεθίσματα, τα οποία πρέπει να απέχουν μερικά δευτερόλεπτα μεταξύ τους.

⁴¹ Δεν πρέπει να συγχέεται με τη *δομική* (structural) μαγνητική τομογραφία, που είναι αυτή που συνήθως χρησιμοποιείται ιατρικά για να απεικονιστούν οι ιστοί. Παρότι χρησιμοποιείται η ίδια τεχνολογία και αξιοποιούνται τα ίδια φυσικά φαινόμενα, διαφέρει ο τρόπος μαγνητικής διέγερσης και ως εκ τούτου η αντίθεση που καταγράφεται στη σύσταση των ιστών. Συγκεκριμένα, η δομική τομογραφία απεικονίζει την περιεκτικότητα σε νερό (καλύτερα, την αντίθεση νερού-λίπους), ενώ η λειτουργική τομογραφία απεικονίζει την οξυγόνωση του αίματος (καλύτερα, την αναλογία οξυγονωμένης αιμοσφαιρίνης).

TMS (transcranial magnetic stimulation, διακρανιακή μαγνητική διέγερση). Αντίθετα με τις προαναφερθείσες μεθόδους, η διακρανιακή μαγνητική διέγερση δεν καταγράφει νευρωνική δραστηριότητα ούτε κάποιες έμμεσες ενδείξεις συνδεδεμένες με αυτή. Πρόκειται για μια ενεργή μέθοδο, κατά την οποία εφαρμόζεται ένας ισχυρός μαγνητικός παλμός τοπικά, με έναν ηλεκτρομαγνήτη που τοποθετείται πάνω από την υπό μελέτη εγκεφαλική περιοχή. Ο παλμός διαταράσσει την ηλεκτρομαγνητική λειτουργία του εγκεφαλικού ιστού που βρίσκεται από κάτω του με αποτέλεσμα την προσωρινή απενεργοποίηση της περιοχής, κάτι σαν μικρό, τοπικό, ελεγχόμενο και αναστρέψιμο εγκεφαλικό επεισόδιο. Ο μαγνητικός παλμός πυροδοτείται σε συγκεκριμένη χρονική στιγμή, αναφορικά με κάποιο παρεχόμενο ερέθισμα ή με κάποια παρατηρήσιμη συμπεριφορά του ατόμου, ενώ καταγράφεται κάποια απόκριση. Η καταγραφή αυτή συγκρίνεται με συνθήκη αναφοράς κατά την οποία ο μαγνητικός παλμός εφαρμόζεται σε άλλη περιοχή ή παραμένει ανενεργός. Με τον τρόπο αυτό μελετάται η επίδραση της διαταραχής της εγκεφαλικής λειτουργίας σε συγκεκριμένες περιοχές, θεωρώντας ότι οι παρατηρούμενες διαφορές μεταξύ του χρονικού διαστήματος διέγερσης και του χρονικού διαστήματος αναφοράς αποκαλύπτουν το ρόλο της διεγερόμενης περιοχής στην υπό μελέτη διεργασία ή, τουλάχιστον, κατά πόσο η διεγερόμενη περιοχή εμπλέκεται ή όχι στην εκτέλεση της διεργασίας.

Οι τεχνικές αυτές αποτελούν σήμερα την αιχμή της ερευνητικής τεχνολογίας του υλικού υποστρώματος του νου. Πάντως θα πρέπει να γνωρίζουμε τους σημαντικούς περιορισμούς που τίθενται στην ερμηνεία των καταγραφών από τις ιδιότητες της κάθε τεχνολογίας. Επίσης είναι σημαντικό να μην ξεχνάμε πως η γνώση κάποιων «διεγερμένων» περιοχών δε συνεπάγεται άμεσα καμία χρήσιμη πληροφορία για τις νοητικές λειτουργίες. Ο μόνος τρόπος προσέγγισης του ζητήματος του νου μέσα από νευροεπικονιστικές τεχνικές περνά μέσα από πολυεπίπεδα πειράματα που με συγκλίνουσες ενδείξεις από διαφορετικές τεχνικές υποδεικνύουν τη φύση των επιμέρους εγκεφαλικών διεργασιών και έμμεσα τις νοητικές διεργασίες που αντιστοιχούν σε αυτές.

Ανακεφαλαιώνοντας, τα πρωτογενή δεδομένα των νευροεπισημών είναι έμμεσες παρατηρήσεις νευρωνικού συγχρονισμού ή αποτελεσμάτων της νευρωνικής διέγερσης. Τα ευρήματα των σχετικών πειραμάτων είναι στατιστικά σημαντικές διαφορές στο βαθμό συγχρονισμού ή στις επιπτώσεις της νευρωνικής διέγερσης μεταξύ συνθηκών σε κατάλληλα σχεδιασμένα πειράματα. Με βάση αυτά τα ευρήματα οι νευροεπισημές συνεισφέρουν στο οικοδόμημα της γνωσιακής επιστήμης δομικά και λειτουργικά στοιχεία για τις εγκεφαλικές περιοχές που συμμετέχουν στις υπό εξέταση συμπεριφορές και, κατ' επέκταση, στις αντίστοιχες υποθετικές νοητικές διεργασίες.

11 Φιλοσοφία και γλωσσολογία

Αντίθετα από τις εμπειρικές προσεγγίσεις της γνωστικής ψυχολογίας και των νευροεπισημών, ο τρίτος βασικός «κλάδος» της γνωσιακής επιστήμης δε συνεισφέρει περιγραφές

παρατηρήσεων ή μοντέλα πειραματικών αποτελεσμάτων, αλλά *a priori* επιχειρήματα για τη δομή και τη λειτουργία του νου. Η θεωρητική αυτή παράδοση βασίζεται στη *διαίσθηση* και το *παράδειγμα*: ένας φιλόσοφος ή ένας γλωσσολόγος παρατηρεί κάτι στη νοητική ή γλωσσική του συμπεριφορά και διατυπώνει ένα λογικό επιχείρημα για να αποδώσει τις ρίζες της συμπεριφοράς σε κάποια νοητική δομή ή λειτουργία. Απόρροια της φιλοσοφικής προσέγγισης στη θεώρηση του νου ως επεξεργαστή πληροφοριών αποτελεί όλη η σύλληψη της νόησης ως συμβολικής επεξεργασίας που παρουσιάστηκε στο πρώτο μέρος.

11.1 Γλωσσική ικανότητα

Η συμβολική φιλοσοφία του νου και η κλασική θεωρητική γλωσσολογία (κατά Chomsky) παρουσιάζουν πολλά κοινά χαρακτηριστικά. Και στις δύο περιπτώσεις θεωρείται *a priori* δεδομένο ότι το σύστημα-στόχος (νόηση και γλώσσα) αποτελεί ένα τυπικό (φορμαλιστικό) υπολογιστικό σύστημα. Οι όποιες διαφωνίες αφορούν στο ακριβές είδος του φορμαλιστικού συστήματος ή σε λεπτομέρειες της λειτουργίας του. Όταν ο Chomsky έκανε την επανάσταση στη γλωσσολογία με τις «συντακτικές δομές» του, δεν έθεσε το ζήτημα αν η γλώσσα είναι ένα υπολογιστικό σύστημα ή όχι, αλλά θεωρώντας αυτό δεδομένο απέδειξε ότι κάποιοι συγκεκριμένοι τύποι φορμαλισμού δεν επαρκούν για να περιγράψουν ορισμένες ιδιότητες της γλωσσικής ικανότητας. Μάλιστα πολλά από τα παραδείγματά του δεν προέρχονταν καν από κάποια πραγματική φυσική γλώσσα αλλά ήταν αφηρημένες προσεγγίσεις σε μαθηματικά συστήματα.

Κατά τον Chomsky και τους υποστηρικτές του, γλώσσα, ή γλωσσικό σύστημα, είναι το σύστημα που μπορεί να διακρίνει τις γραμματικές από τις μη γραμματικές προτάσεις. Είναι σημαντικό ότι εδώ το γλωσσικό σύστημα ταυτίζεται με τη *γραμματική* ικανότητα. Ας δούμε τι ακριβώς σημαίνει αυτό. Ένα γλωσσικό σύστημα διαθέτει ένα σύνολο συμβόλων· ας πούμε τις λέξεις.⁴² Τα σύμβολα αυτά διατάσσονται σε αλληλουχίες σχηματίζοντας φράσεις, οι οποίες χρησιμοποιούνται για το σχηματισμό και τη μετάδοση σύνθετων νοημάτων βασισμένων στο νόημα των λέξεων και τις σχέσεις διάταξης μεταξύ τους. Από όλες τις πιθανές αλληλουχίες συμβόλων, δηλαδή τις πιθανές φράσεις που μπορούν να σχηματιστούν από συνδυασμούς συμβόλων στη σειρά, μόνο μερικές είναι αποδεκτές. Για παράδειγμα, οι αλληλουχίες «το κορίτσι δαγκώνει το σκυλί» και «το σκυλί δαγκώνει το κορίτσι» είναι αποδεκτές αλληλουχίες των συμβόλων «το», «σκυλί», «δαγκώνει», «κορίτσι». Αντίθετα, οι αλληλουχίες «δαγκώνει το κορίτσι σκυλί το» και «το δαγκώνει σκυλί κορίτσι το» δεν είναι αποδεκτές, διότι παραβαίνουν τους εσωτερικούς κανόνες δομής της γραμματικής, δηλαδή του γλωσσικού συστήματος. Η περίφημη φράση του Chomsky «άχρωμες πράσινες ιδέες κοιμούνται μανιασμένα», παρότι δεν μπορεί να σημαίνει τίποτα, είναι παρόλα αυτά καλά σχηματισμένη από την άποψη της γραμματικής. Δείχνει, μεταξύ άλλων, ότι η γραμματική ικανότητα για την οποία μιλάμε είναι κάτι υπαρκτό, δομικό στοιχείο της γλώσσας, το

⁴²Για τις ανάγκες των παραδειγμάτων στην ενότητα αυτή δεν είναι απαραίτητο να αναφερθούν οι πολυπλοκότητες που προέρχονται από τη μορφολογική συνθετικότητα.

οποίο δεν εξαρτάται από τη νοηματική απόδοση. Σύμφωνα με τον ορισμό των κατά Chomsky θεωρητικών γλωσσολόγων, στόχος της γλωσσολογίας είναι να περιγραφεί ένα σύστημα το οποίο θα δέχεται ως εισροή μια αλληλουχία συμβόλων και θα αποδίδει ως εκροή την αποδοχή ή απόρριψη της αλληλουχίας. Όταν η λειτουργία του συστήματος αυτού θα συμφωνεί με τη διαισθητική κρίση ενός φυσικού ομιλητή της γλώσσας, τότε θεωρούμε ότι οι κανόνες βάσει των οποίων λειτουργεί το σύστημα θα περιγράφουν με ακρίβεια τη γλωσσική ικανότητα του ομιλητή. Ιδανικά, η προσέγγιση αυτή θα υποστήριζε πως οι ίδιοι κανόνες πραγματώνονται στο γλωσσικό σύστημα του ομιλητή και σε αυτούς οφείλεται η γλωσσική του ικανότητα.

Η προσέγγιση αυτή επιτρέπει στο γλωσσολόγο να εστιάσει στα πραγματικά και δυνητικά γλωσσικά δεδομένα, δηλαδή στις αλληλουχίες λέξεων που εμφανίζονται (παρατηρούνται) ή θα μπορούσαν να εμφανίζονται (είναι αποδεκτές) σε μια γλωσσική κοινότητα, και να μελετά τη διάρθρωση των προτάσεων ανεξάρτητα από τη χρήση και το ρόλο τους. Ο θεωρητικός γλωσσολόγος αυτής της κλασικής προσέγγισης, λοιπόν, δεν μελετά τον ομιλητή αλλά ένα σύνολο προτάσεων (πραγματικών ή δυνητικών). Κατασκευάζει ένα μαθηματικό σύστημα περιγραφής των προτάσεων, ως αυτόνομο «γλωσσικό σύστημα». Οι διαφορετικοί κλάδοι της θεωρητικής γλωσσολογίας εφαρμόζουν αυτή τη μέθοδο σε διαφορετικές ιδιότητες των γλωσσικών δεδομένων. Έτσι, ο κλάδος μελέτης της σύνταξης ασχολείται με τη διάταξη της αλληλουχίας των λέξεων ώστε να συγκροτούν γραμματικά αποδεκτές προτάσεις. Ο κλάδος της μορφολογίας ασχολείται με τους συνδυασμούς επιμέρους μορφημάτων που συγκροτούν κλιτές, παράγωγες και σύνθετες λέξεις, για παράδειγμα το συνδυασμό του θεματικού μορφήματος παιζ- με τις καταλήξεις -ω, -εις, -ει κλπ. Ο κλάδος της φωνολογίας μελετά τους συνδυασμούς των αφηρημένων στοιχειωδών μονάδων χωρίς σημασία που απαρτίζουν τις λέξεις. Οι μονάδες αυτές θεωρείται ότι σχετίζονται με τους φθόγγους, δηλαδή τους παρατηρούμενους ήχους της ομιλίας. Η φωνολογία θα πρέπει να διακρίνει, για παράδειγμα, την αποδεκτή λέξη /anθropos/ («άνθρωπος») από τη μη αποδεκτή /nθpaoσp/ (που περιλαμβάνει τους ίδιους φθόγγους). Τέλος, ο κλάδος της σημασιολογίας μελετά τον τρόπο που συνδυάζονται οι σημασίες των λέξεων μιας φράσης για να συγκροτήσουν το νόημα ολόκληρης της φράσης. Φυσικά, κι εδώ η σημασία και το νόημα περιγράφονται με τυπικό (συντακτικό) τρόπο, εστιάζοντας φορμαλιστικά στον τρόπο συνδυασμού και όχι στο εννοιολογικό περιεχόμενο καθαυτό.

Ο ορισμός της γλωσσικής ικανότητας ως τυπικό σύστημα παραβλέπει σκόπιμα τις *χρήσεις* της γλώσσας, τόσο για τη σκέψη και την εννοιολογική δόμηση του νου όσο και για την επικοινωνία, καθώς ασχολείται με «γλωσσικά δεδομένα» (λέξεις και προτάσεις) και όχι με ομιλητές και περιστάσεις. Καθώς όμως τα πραγματικά δεδομένα εξαρτώνται πάντα σε κάποιο βαθμό από ομιλητές και περιστάσεις, η προσέγγιση αυτή επιστρατεύει ένα θεωρητικό τέχνασμα, διαχωρίζοντας την ικανότητα από την επιτέλεση. Η *ικανότητα* (competence) αφορά σε έναν υποθετικό πυρήνα του συστήματος, ο οποίος είναι τέλεια φορμαλιστικός, ενώ η *επιτέλεση* (performance) έχει να κάνει με την εκτέλεση των λειτουργιών σε ένα πραγματικό φυσικό υπόβαθρο το οποίο μπορεί να υπόκειται σε διάφορους πρακτικούς

περιορισμούς. Για παράδειγμα, ένα τέλει υπολογιστικό σύστημα μπορεί να διαχειριστεί χωρίς δυσκολία μια πρόταση της μορφής «ο σκύλος που δάγκωσε τη γάτα που κυνήγησε το ποντίκι που έφαγε το τυρί. . . » όσο μεγάλη κι αν είναι, εφόσον δεν αλλάζει κάτι στη δομή των φράσεων και τις σχέσεις μεταξύ τους. Όμως σε ένα πραγματικό σύστημα, που δεν έχει απεριόριστη μνήμη, είναι αναμενόμενο ότι κάποια στιγμή θα δημιουργηθεί πρόβλημα στην επεξεργασία και την κατανόηση μιας πολύ μεγάλης τέτοιας πρότασης. Ο διαχωρισμός σε δύο υποσυστήματα και η εστίαση στη μελέτη του ενός επιτρέπει την απομόνωση του μαθηματικού πυρήνα από τις περιοριστικές δυσχέρειες του φυσικού κόσμου (και του πραγματικού εγκεφάλου στον οποίο λειτουργεί το σύστημα). Επιπλέον, ο διαχωρισμός αυτός επιτρέπει την ανάθεση στο «μη ενδιαφέρον» σύστημα εκτέλεσης, δηλαδή στην επιτέλεση, οποιασδήποτε απόκλισης από την τέλεια φορμαλιστική συμπεριφορά. Έτσι, μια ελλιπής πρόταση μπορεί να αποδοθεί σε ανεπαρκή γλωσσική μνήμη, μια ακατανόητη φράση σε ανεπαρκή επεξεργαστική χωρητικότητα κλπ.

Στη γλωσσολογία, η ικανότητα αναφέρεται στην άρρητη γνώση που έχει κάθε ομιλήτης για τη γλώσσα του. Η γνώση αυτή δεν είναι συνειδητή και ούτε μπορεί να εκφραστεί λεκτικά. Είναι η γνώση εκείνη που επιτρέπει στον ομιλητή να διακρίνει τις αποδεκτές από τις μη αποδεκτές προτάσεις, να επιτελεί γλωσσικές κρίσεις χάρη στη γλωσσική του διαίσθηση. Η ύπαρξη της γνώσης αυτής συνάγεται από την παρατηρούμενη ικανότητα του ομιλητή να κάνει αυτές τις γλωσσικές διακρίσεις, και δεν εξαρτάται από τη μόρφωση ή καν από γνώση του γραπτού λόγου (γραμματισμό). Αντίστοιχα, στη συμβολική προσέγγιση του νου, η χρήση μιας τυπικής «νοητικής γλώσσας» (ή γλώσσας της σκέψης, language of thought) συνάγεται από τις παρατηρούμενες νοητικές ικανότητες και κυρίως από τη συστηματικότητά τους. Προφανώς η νοητική γλώσσα δεν αποτελεί ρητή γνώση αλλά εργαλείο σκέψης. Ο τυπικός (φορμαλιστικός) χαρακτήρας της σημαίνει ότι οποιαδήποτε απόκλιση από την αφαιρετική απολυτότητα παρατηρείται στην πραγματική συμπεριφορά πρέπει να αποδοθεί σε περιφερειακά ελλείμματα και όχι σε αδυναμία του επεξεργαστή. Παρότι στη συμβολική προσέγγιση δεν χρησιμοποιείται συχνά η ίδια ορολογία, η έννοια της ικανότητας ως διακρίσιμης από την επιτέλεση είναι παρούσα, έστω και άδηλα.

11.2 Μεθοδολογία της διαίσθησης

Η σημασία της διαίσθησης είναι ιδιαίτερα σημαντική στη θεωρητική γλωσσολογία, όπου κατέχει κεντρική μεθοδολογική θέση έναντι των πειραματικών εμπειρικών στοιχείων άλλων κλάδων. Για ένα θεωρητικό γλωσσολόγο είναι αρκετή η προσωπική του διαίσθηση σχετικά με το αν μια φράση είναι αποδεκτή στη γλώσσα του ή όχι. Αν πρόκειται για γλώσσα που δεν μιλά ο ίδιος, τότε αρκεί η πρόσβαση σε έναν και μοναδικό «πληροφοριοδότη» (ή «πληροφορητή»), ο οποίος θα πρέπει απλώς να μιλά τη γλώσσα-στόχο ως μητρική, να είναι δηλαδή *φυσικός ομιλητής* (native speaker). Στην προσέγγιση αυτή, η επάρκεια της ατομικής διαίσθησης είναι κρίσιμη θεωρητική επιλογή και όχι δείγμα διανοητικής οκνηρίας ή μεθοδολογικής ανεπάρκειας. Συγκεκριμένα, κάθε άνθρωπος θεωρείται ότι διαθέτει ένα πλήρες γλωσσικό σύστημα για τη μητρική του γλώσσα. Αυτό το γλωσσικό σύστημα είναι ε-

παρκώς όμοιο μεταξύ των ομιλητών της ίδιας γλώσσας ώστε να είναι δυνατή η επικοινωνία, δηλαδή η μεταβίβαση πληροφορίας χωρίς απώλεια του νοήματος. Το γλωσσικό σύστημα είναι όμως και ατομικό, δηλαδή θεωρητικά μπορεί να μην υπάρχουν δύο άτομα που να έχουν πανομοιότυπα γλωσσικά συστήματα σε όλες τις λεπτομέρειες (το ατομικό γλωσσικό σύστημα ονομάζεται «ιδιόλεκτος» στη γλωσσολογία, κατά το ομαδικό «διάλεκτος»). Παρότι οι δύο αυτές παρατηρήσεις είναι κατά κάποιον τρόπο αντίθετες, βάσει οποιασδήποτε από τις δύο αρκεί ένας οποιοσδήποτε φυσικός ομιλητής για να παράσχει τις απαραίτητες πληροφορίες. Δεν είναι δηλαδή απαραίτητος (ή, από την άλλη πλευρά, δεν είναι εφικτός) ο έλεγχος αξιοπιστίας των γλωσσικών δεδομένων με τη συλλογή περισσότερων στοιχείων.⁴³

Λέγοντας «διαίσθηση» εννοούμε μια υποκειμενική κρίση η οποία βασίζεται σε άμεση αντίληψη και όχι σε τεκμηριωμένη αιτία. Οι διαισθητικές κρίσεις συνοδεύονται από την υποκειμενική εντύπωση ότι αυτές είναι βέβαιες και αναπόφευκτες. Όταν την ίδια διαισθητική κρίση τη δίνει το σύνολο, ή έστω η μεγάλη πλειονότητα, των ερωτώμενων, τότε είναι βάσιμο να θεωρήσουμε ότι η διατομική αξιοπιστία εκφράζει μια γενικότερη νοητική συστηματικότητα. Με την έννοια αυτή, οι κοινές διαισθήσεις αποτελούν ενδείξεις για την κοινή νοητική λειτουργία. Στη γλωσσολογία, η διαίσθηση λειτουργεί με το λεγόμενο «γλωσσικό κριτήριο» ή «γλωσσικό αίσθημα», την υποκειμενική αίσθηση που έχουμε για κάποια πρόταση ή λέξη ότι είναι αποδεκτή ή όχι. Δεν αναφερόμαστε σε δικαιολόγηση με ρητή αναφορά σε κανόνες της σχολικής (ή οποιασδήποτε άλλης) γραμματικής ούτε σε κανονιστικές αντιλήψεις των φιλολόγων για το πώς «πρέπει να μιλάμε σωστά τη γλώσσα μας». Το γλωσσικό κριτήριο το διαθέτουν πλήρως ανεπτυγμένο όλοι οι φυσικοί ομιλητές της γλώσσας. Αντίστοιχα, μπορεί να μιλά κανείς για ηθικό ή αισθητικό κριτήριο, αναφερόμενος σε παρόμοιας υποκειμενικής υφής άμεσες κρίσεις. Όπως οι γλωσσολόγοι χρησιμοποιούν παραδείγματα αποδεκτών και μη αποδεκτών προτάσεων, πυροδοτώντας τις κατάλληλες γλωσσικές διαισθήσεις, έτσι και οι φιλόσοφοι χρησιμοποιούν υποθετικές περιστάσεις και συγκεκριμένα παραδείγματα για να πυροδοτήσουν διαισθήσεις αναφορικά με τη φύση και τη λειτουργία του νου, κυρίως σε θέματα συνείδησης.

Στη φιλοσοφία του νου βρίσκουμε φημισμένα νοητικά (υποθετικά) πειράματα όπως το κινέζικο δωμάτιο (του Searle), τη δίδυμη γη (του Putnam) και την υπερνευροεπιστήμονα Μάιρη (του Jackson). Μια υποθετική κατάσταση κατασκευάζεται και χρησιμοποιείται για να δημιουργήσει μια εντύπωση και να οδηγήσει σε ένα «αυτονόητο» συμπέρασμα, δηλαδή σε μια διαίσθηση, για το τι είναι, τι μπορεί ή τι πρέπει να είναι, οι έννοιες, η σκέψη, οι προτασιακές στάσεις κλπ. Ο Dennett έχει ονομάσει τα προϊόντα της τεχνικής αυτής «αντλίες διαίσθησης» (intuition pumps). Ουσιαστικά και στην περίπτωση αυτή πρόκειται

⁴³Στην πράξη ανακύπτουν σημαντικά ζητήματα αξιοπιστίας των γλωσσικών διαισθήσεων για κάποιες ιδιόμορφες περιπτώσεις. Επίσης, ενώ η θεωρητική διάκριση γίνεται μεταξύ απολύτως αποδεκτών και απολύτως απαράδεκτων προτάσεων (ή λέξεων), στην πράξη υπάρχει μια διαβάθμιση, η οποία περιλαμβάνει και ενδιάμεσης αποδοχής ή αμβισθητούμενες περιπτώσεις. Παρότι η αυστηρή θεωρητική αντιμετώπιση αυτών των φαινομένων θα ήταν η ολοσχερής απόδοση στην επιτέλεση, λόγω της έκτασης του προβλήματος η γλωσσολογία αξιοποιεί τις ιδιομορφίες αυτές δανειζόμενη μεθοδολογικά στοιχεία από τον πειραματικό κλάδο της ψυχολογίας και μελετώντας τις αιτίες των ιδιαιτεροτήτων ή των διαβαθμίσεων.

για μεμονωμένα παραδείγματα, έστω κι αν δεν προκύπτουν από πραγματική ή δυνάμει παρατηρήσιμη συμπεριφορά όπως οι προτάσεις και λέξεις των γλωσσολόγων. Όπως στην κρίση των προτάσεων ως αποδεκτών ή μη οδηγούμαστε σε συμπεράσματα για το γλωσσικό σύστημα, έτσι και στην κρίση των συνεπειών των υποθετικών καταστάσεων καθοδηγούμαστε μέσω της διαίσθησης σε συμπεράσματα για τη νόηση. Για παράδειγμα, στο κινέζικο δωμάτιο κάποιος υποτίθεται πως ακολουθεί τυφλά κανόνες για να αντιστοιχίσει κάποια εισερχόμενα σχέδια με άλλα εξερχόμενα, μη γνωρίζοντας περί τίνος πρόκειται, ενώ οι Κινέζοι έξω απ' το δωμάτιο αναγνωρίζουν την ανταλλαγή σχεδίων ως γλωσσική επικοινωνία μέσω ιδεογραμμμάτων (γραπτού λόγου). Το διαισθητικά αβίαστο συμπέρασμα είναι ότι η ακολουθία εφαρμογής κανόνων δεν συνιστά κατανόηση, ένα επιχείρημα εναντίον της ισχυρής τεχνητής νοημοσύνης αλλά και της συμβολικής προσέγγισης στη νόηση.⁴⁴ Τα φιλοσοφικά και τα γλωσσολογικά επιχειρήματα μπορεί καμία φορά να φαίνονται τραβηγμένα ή αβάσιμα, αν εξαρτώνται από αμφίβολες υποθέσεις και σημεία εκκίνησης ή από ιδιαίτερα περιορισμένες συλλήψεις των συστημάτων-στόχων. Είναι όμως εξαιρετικά χρήσιμα διότι υποδεικνύουν το εύρος των φαινομένων τα οποία θα πρέπει να εξηγήσει επαρκώς μια αποδεκτή θεωρία και διότι καθιστούν προφανείς ορισμένες ανεπάρκειες, ασάφειες, και ιδιαίτερα αντιφάσεις σε όλα τα θεωρητικά συστήματα, φορμαλιστικά και μη, όσον αφορά στην ερμηνεία και περιγραφή των γλωσσικών και νοητικών συμπεριφορών.

Όπως είναι φανερό, η μεθοδολογία της διαίσθησης βασίζεται σε συγκεκριμένα παραδείγματα, είτε επιλεγμένα από την παρατήρηση, είτε ειδικά κατασκευασμένα για το σκοπό αυτό. Μια συγκεκριμένη πρόταση με κάποια δομή επαρκεί για να καταδείξει ότι η δομή αυτή είναι αποδεκτή (ή ότι δεν είναι). Μια συγκεκριμένη φανταστική ιστορία επαρκεί για να καταδείξει τις αναγκαίες συνέπειες κάποιας θεωρητικής υπόθεσης. Έτσι, σημαντικό μέρος της μεθοδολογίας των κλάδων αυτών συνίσταται στην επιλογή, συλλογή, ή κατασκευή των κατάλληλων παραδειγμάτων πάνω στα οποία ασκείται η κοινά αποδεκτή διαίσθηση για να επιφέρει το προσδοκώμενο αποτέλεσμα, δηλαδή τη θεωρητική πρόοδο. Όταν τα παραδείγματα είναι καλά επιλεγμένα ώστε να οδηγούν αβίαστα σε κοινά αποδεκτές διαισθήσεις τότε η επίδρασή τους είναι αναπόφευκτη. Με τον τρόπο αυτό η θεωρητική γλωσσολογία οικοδομεί θεωρητικά σχήματα για να εξηγήσει τις παρατηρούμενες γλωσσικές δομές, ενώ η φιλοσοφία διερευνά τις λογικές συνέπειες εναλλακτικών προκειμένων ώστε να αναδείξει συγκλίσεις, αντιφάσεις ή επιλογές. Υπάρχουν φυσικά πολλές σημαντικές διαφορές μεταξύ της γλωσσολογικής και της φιλοσοφικής προσέγγισης, αναφορικά όμως με τη μέθοδο ίσως είναι χρήσιμο να αναφέρουμε μία πολύ σημαντική από αυτές: Η γλωσσολογία βασίζεται σε κάποιο βαθμό σε παρατηρήσιμα (προτάσεις ή λέξεις), τα οποία μαζί με τις αντίστοιχες διαισθητικές κρίσεις αποτελούν τα πρωτογενή δεδομένα της. Αντίθετα, η φιλοσοφία δεν περιορίζεται υποχρεωτικά από τη φυσική πραγματικότητα αλλά εστιάζει στη λογική αναγκαιότητα, κι έτσι κατά κανόνα δεν βασίζεται στην αξιοποίηση δεδομένων. Όμως κι εδώ

⁴⁴Ας σημειωθεί όμως ότι η υποκειμενικότητα της διαίσθησης κρύβει τον κίνδυνο να αγνοήσουμε σοβαρά λογικά σφάλματα ή κενά, όπως για παράδειγμα στην περίπτωση του κινέζικου δωματίου όπου δεν προσδιορίζεται τι σημαίνει «κατανόηση» αλλά το επιχείρημα βασίζεται στη διαισθητική αντίληψη της λέξης.

υπάρχουν εξαιρέσεις. Στη γνωσιακή επιστήμη είναι ίσως ιδιαίτερα σημαντική η παράδοση της φυσικής φιλοσοφίας, όπου λαμβάνονται υπόψη ως αφετηρία ή αναφορά συστηματικές παρατηρήσεις από εμπειρικούς (πειραματικούς) κλάδους. Ειδικά για τη φιλοσοφία του νου αυτό σημαίνει παρατηρήσεις κυρίως από ψυχολογία και νευροεπιστήμες.

11.3 Μάθηση εναντίον έμφυτης γνώσης

Ένα πόρισμα της συμβολικής προσέγγισης στη γλώσσα και τη νόηση είναι ο έμφυτος χαρακτήρας τους. Όπως έχει υποστηρίξει ο Chomsky για τη γλώσσα και ο Fodor για τη νόηση, οι βασικές γλωσσικές και νοητικές δομές πρέπει να είναι έμφυτες διότι δεν είναι δυνατόν να μαθευτούν. Η «μαθευσιμότητα» (learnability) της γλώσσας και της σκέψης αποτελεί άλλο ένα κοινό στοιχείο της φορμαλιστικής προσέγγισης της θεωρητικής γλωσσολογίας και της συμβολικής φιλοσοφίας του νου.

Το επιχείρημα για τη γλώσσα έχει ως εξής: Ας αναλογιστούμε την περίπτωση ενός βρέφους που έρχεται στον κόσμο και αλληλεπιδρά σε ένα συγκεκριμένο γλωσσικό περιβάλλον. Ένα βασικό «έργο» του στα πρώτα χρόνια της ζωής του είναι να αποκομίσει, βάσει του γλωσσικού περιβάλλοντος, τα στοιχεία του γλωσσικού συστήματος που χρειάζονται για τη γλωσσική του επικοινωνία. Να γίνει δηλαδή φυσικός ομιλητής της γλώσσας του περιγύρου, κάτι που όλα τα παιδιά καταφέρνουν εκπληκτικά γρήγορα και εύκολα. Τα στοιχεία στα οποία μπορεί να βασιστεί είναι πεπερασμένα. Περιλαμβάνουν τα εκφωνήματα που ακούει γύρω του, όσα δηλαδή από αυτά του επιτρέπει η μνήμη του και η περιορισμένη γνωστική του ανάπτυξη να αντιληφθεί, να ερμηνεύσει σωστά, να καταχωρίσει και να θυμάται. Το γλωσσικό σύστημα το οποίο πρέπει να συναγάγει από τα στοιχεία αυτά είναι παραγωγικό. Δηλαδή, διαθέτει την ικανότητα παραγωγής απεριόριστου αριθμού φράσεων από το εκάστοτε πεπερασμένο λεξιλόγιο. Όμως, για οποιοδήποτε πεπερασμένο σύνολο στοιχείων, υπάρχουν άπειρα συστήματα ικανά να το «εξηγήσουν». Άρα κάθε στιγμή το παιδί που μαθαίνει τη γλώσσα έχει άπειρους τρόπους να ερμηνεύσει τις παρατηρήσεις του. Το σύνολο της γλωσσικής του εμπειρίας είναι θεμελιωδώς ανεπαρκές για να οδηγήσει επαγωγικά στην κατασκευή του απαραίτητου παραγωγικού συστήματος.

Για να κατανοήσουμε τη θεωρητική ανεπάρκεια αρκεί να σκεφτούμε τι απαιτείται να κάνει το αναπτυσσόμενο γλωσσικό σύστημα του παιδιού, προκειμένου να καταλήξει στο σωστό ενήλικο γλωσσικό σύστημα με βάση τα διαθέσιμα δεδομένα: Αφενός θα πρέπει να αποδέχεται ως «γραμματικές» τις φράσεις που προσέλαβε (που άκουσε και επεξεργάστηκε σωστά) το παιδί. Αφετέρου, από το απεριόριστο πλήθος φράσεων που δεν έχουν ποτέ προσληφθεί, θα πρέπει να απορρίπτει δυνητικά ορισμένες ως «μη γραμματικές» αλλά να αποδέχεται άλλες ως «γραμματικές» (που απλώς δεν έτυχε ποτέ να χρησιμοποιηθούν). Το πρόβλημα αυτό είναι μαθηματικά μη επιλύσιμο. Άρα, το γλωσσικό σύστημα, υπό αυτή την έννοια, είναι μη μαθεύσιμο. Συνεπώς ο θεωρητικός γλωσσολόγος που ασπάζεται τη φορμαλιστική προσέγγιση στη γλώσσα συμπεραίνει ότι το παιδί διαθέτει εκ των προτέρων (εγγενώς) τις απαραίτητες νοητικές δομές και διεργασίες που εξυπηρετούν το γλωσσικό σύστημα. Συμπεραίνει δηλαδή ότι η γλωσσική ικανότητα είναι κατά βάση έμφυτη. Ο

ρόλος της εμπειρίας, σε αυτήν τη θεωρητική προσέγγιση, είναι να καθοδηγήσει το παιδί να «ρυθμίσει» το έμφυτο γλωσσικό του σύστημα ώστε αυτό να ανταποκρίνεται στις ανάγκες της μητρικής του γλώσσας. Επειδή όλες οι ανθρώπινες γλώσσες θεωρείται ότι έχουν την ίδια δομή (αφού άλλωστε θεωρούνται έμφυτες, άρα γενετικά προκαθορισμένες), η υπόθεση της ρύθμισης «παραμέτρων» διαμορφώνει το έργο του γλωσσολόγου ως εξής: να περιγραφεί το αρχικό εκείνο γενικό γλωσσικό σύστημα και οι παράμετροί του, έτσι ώστε με κατάλληλες ρυθμίσεις των παραμέτρων του να προκύπτουν όλες οι γνωστές γλώσσες. Το σύστημα αυτό ονομάζεται Καθολική Γραμματική (Universal Grammar, ή UG).

Το προαναφερθέν επιχείρημα ονομάζεται «ένδεια του ερεθίσματος» (poverty of the stimulus, σε ελεύθερη απόδοση, επιχείρημα «ανεπάρκειας ερεθισμάτων») και δείχνει ότι δεν είναι δυνατόν να μαθευτεί ένα τυπικό (φορμαλιστικό) παραγωγικό σύστημα από ένα πεπερασμένο σύνολο δεδομένων. Το συμπέρασμα των γλωσσολόγων από αυτό είναι ότι η γλώσσα είναι μια έμφυτη ικανότητα.⁴⁵ Το αντίστοιχο επιχείρημα για τη νόηση, ιδωμένη ως τυπικό υπολογιστικό σύστημα, έχει ως εξής: Ας υποθέσουμε ότι το παιδί σε κάποιο στάδιο ανάπτυξης δεν διαθέτει ένα τελεστή που αντιστοιχεί σε επόμενο στάδιο ανάπτυξης. Για παράδειγμα, βρίσκεται στο «απλούστερο» στάδιο της προτασιακής λογικής και δεν μπορεί να συλλάβει σχέσεις και εκφράσεις που περιλαμβάνουν ποσοτικούς τελεστές (π.χ., $\forall xP(x)$) διότι αυτό απαιτεί ένα ιεραρχικά ανώτερο σύστημα. Για να συλλάβει και να «αποκτήσει» τους ποσοτικούς τελεστές πρέπει, υποθέτουμε, επαγωγικά να υποθέσει την ύπαρξή τους και εμπειρικά να την επιβεβαιώσει από τα στοιχεία που προσλαμβάνει. Όμως πώς θα διατυπώσει την πρόταση ώστε αυτή εν συνεχεία να επιβεβαιωθεί επαγωγικά; Η προς επιβεβαίωση πρόταση περιλαμβάνει αναγκαστικά τον ποσοτικό τελεστή, αφού αυτή είναι η υπόθεση μάθησης. Όμως η προς επιβεβαίωση πρόταση δεν μπορεί να περιλαμβάνει τον τελεστή αυτόν διότι αυτός δεν είναι γνωστός! Άρα το σύστημα δεν μπορεί να γίνει πιο περίπλοκο επεξεργαστικά μέσω της μάθησης διότι δεν μπορεί σε κάθε στάδιο να εκφράσει τις άγνωστες έννοιες και πράξεις και άρα δεν μπορεί να τις μάθει. Με το ίδιο επιχείρημα αποδεικνύεται ότι είναι αδύνατη η επαγωγική εννοιολογική μάθηση, διότι κατ'αντιστοιχία δεν είναι δυνατόν να διατυπωθεί προς επαγωγική συναγωγή μια πρόταση για την έννοια «σκύλος» αν δεν υπάρχει ήδη η έννοια αυτή.⁴⁶

11.4 Φύση των εννοιών

Η μελέτη της γλώσσας ειδικά και της νόησης γενικότερα απαιτεί την κατανόηση της φύσης των εννοιών. Με άλλα λόγια, είναι απαραίτητο να προσδιοριστεί ο μηχανισμός αντιστοιχίας μεταξύ των νοητικών (και γλωσσικών) αναπαραστάσεων και του εξωτερικού κόσμου, αυτό που ονομάζουμε σημασιακό περιεχόμενο των αναπαραστάσεων. Η μελέτη της γραμματικής ως συστήματος αποδοχής καλά σχηματισμένων φράσεων δεν περιλαμβάνει προβλη-

⁴⁵Το εναλλακτικό συμπέρασμα, είναι ότι η γλώσσα δεν είναι ένα τυπικό παραγωγικό σύστημα.

⁴⁶Και σε απόλυτη αντιστοιχία με την προηγούμενη περίπτωση υπάρχουν οι εναλλακτικές ερμηνείες ότι η εννοιολογική και η επεξεργαστική μάθηση δεν είναι επαγωγικές ή ότι το ίδιο το σύστημα επεξεργασίας δεν είναι τυπικό-λογικό.

ματισμούς σημασιολογικής φύσης. Όμως η ευρύτερη μελέτη του γλωσσικού συστήματος από τη γλωσσολογία σαφώς περιλαμβάνει το ζήτημα του σημασιακού προσδιορισμού του λεξιλογίου. Όσο για τη μελέτη της νόησης, εκεί το σημασιολογικό ζήτημα είναι ένα από τα κεντρικά προβλήματα. Τι είναι αυτό που παράγει εννοιολογικές αναπαραστάσεις και τους προσδίδει το σημασιακό τους περιεχόμενο; Τι είναι αυτό που κάνει την έννοια «σκύλος» να αντιστοιχεί σε μια συγκεκριμένη (αλλά όχι εύκολα περιγράψιμη) ομάδα τετράποδων τριχωτών ζώων που γαβγίζουν και κυνηγούν γάτες;

Υπάρχουν πολλές προτάσεις που αφορούν στη φύση των εννοιών. Ορισμένες έννοιες όπως «κόκκινο» ή «πόνος» θεωρούνται πρωταρχικές και εξαρτώμενες από την ίδια την υποκειμενική εμπειρία τους. Άλλες όμως, όπως οι προτασιακές στάσεις, δεν αντιστοιχούν σε σταθερή υποκειμενική εμπειρία και πρέπει να περιγραφούν σε σχέση με άλλες έννοιες, άρα η φύση τους προσδιορίζεται από το σύστημα συλλογισμών στο οποίο μετέχουν. Άλλες πάλι θεωρούνται σύνθετες διότι προσδιορίζονται από απλούστερες έννοιες υπό τη μορφή του ορισμού, όπως για παράδειγμα η έννοια «εργένης», η οποία αντιστοιχεί στο «ανύπαντρος άντρας». Έννοιες που αφορούν σε αντικείμενα (όπως καρέκλα γραφείου, καρέκλα, κάθισμα, έπιπλο) ή φυσικές ιεραρχίες (λυκόσκυλο, σκύλος, θηλαστικό, ζώο) μπορούν να θεωρηθούν ότι αποτελούνται από αντιληπτικά ή λειτουργικά πρότυπα και από τις αντίστοιχες ιδιότητες των γνωστικών κατηγοριών.

Ο Fodor έχει εκθέσει τις αδυναμίες όλων των συνηθισμένων προσεγγίσεων στη φύση των εννοιών και έχει υποστηρίξει ότι η συντριπτική πλειονότητα των εννοιών πρέπει να είναι «ατομικές», μοναδιαίες.⁴⁷ Δηλαδή ότι οι έννοιες δεν προκύπτουν από άλλες απλούστερες έννοιες μέσω ορισμών, δεν μπορούν να προκύψουν επαγωγικά και ούτε μπορούν να προσδιοριστούν από επιμέρους ιδιότητες, αλλά αντιστοιχούν σε αυτοτελείς εννοιολογικές οντότητες. Άρα «σκύλος» δεν είναι και δεν μπορεί να είναι «τετράποδο κατοικίδιο ζώο που γαβγίζει και κυνηγάει γάτες», ούτε ένα πρότυπο που προκύπτει από όλους τους σκύλους που περιλαμβάνονται στην αντιληπτική ιστορία ενός ατόμου, ούτε η υποκειμενική εμπειρία της αντίληψης ενός σκύλου, ούτε η σχέση των σκύλων με άλλες έννοιες κλπ. Το κύριο πρόβλημα που οδηγεί τον Fodor σε αυτήν την ακραία τοποθέτηση είναι ότι δεν υπάρχει τρόπος διάκρισης μεταξύ των επιμέρους εννοιών που απαρτίζουν μια σύνθετη έννοια και των υπόλοιπων εννοιών που δεν εμπλέκονται ως συστατικά αλλά μπορεί να σχετίζονται με αυτήν. Σε συνδυασμό με άλλα προβλήματα συνδυαστικότητας, αυτό σημαίνει ότι μια συνθετική έννοια δεν μπορεί να μαθευτεί. Ο σημασιακός ατομισμός του Fodor λέει ότι η έννοια του σκύλου είναι πρωτογενώς η «σκυλότητα» της αντίστοιχης αναπαραστάσης. Το αν αυτό δείχνει πρόοδο στο σημασιακό ζήτημα ή αναδεικνύει τις θεμελιώδεις αδυναμίες και τα αδιέξοδα της γνωσιακής επιστήμης είναι κάτι που κρίνει ο καθένας. Όμως θα πρέπει να σημειωθεί ότι για την ώρα δεν υπάρχουν πολύ καλύτερες προτάσεις που να έχουν φτάσει σε κάποιο συγκεκριμένο αποτέλεσμα και να έχουν δείξει μεγάλη επεξηγηματική ισχύ. Εδώ φαίνεται και η χρησιμότητα των φιλοσοφικών προβληματισμών στην αξιολόγηση των θεωρητικών συστημάτων και η έκθεση της ανεπάρκειας των υπαρχόντων επεξηγηματικών

⁴⁷Fodor, J. A. (1998). *Concepts*. Oxford University Press.

σχημάτων.

12 Συμπέρασμα: Συμπληρωματικότητα

Η γνωστική ψυχολογία συλλέγει εμπειρικά δεδομένα (συνήθως διαφορές ακρίβειας και χρόνου απόκρισης) μέσα από πειράματα για την κατασκευή μοντέλων με θεωρητικές οντότητες που αναφέρονται σε νοητικές λειτουργίες. Οι νευροεπιστήμες συλλέγουν παρατηρήσεις σε βιολογικά συστήματα και καταρτίζουν περιγραφές των φυσικών αυτών συστημάτων και των νευροχημικών λειτουργιών τους. Τέλος, οι φορμαλιστικοί κλάδοι παραθέτουν παραδείγματα και διαισθητικές κρίσεις ως «πρωτογενή δεδομένα» και χρησιμοποιούν λογικά επιχειρήματα για τη συγκρότηση γενικών θεωριών και μεταθεωριών σχετικά με την αρχιτεκτονική του νου και των λειτουργιών του. Οι τρεις αυτοί άξονες αλληλοσυμπληρώνονται και συγκροτούν τον κορμό της γνωσιακής επιστήμης, επικουρούμενοι από την πληροφορική για τη μελέτη των υπολογιστικών ιδιοτήτων των συστημάτων. Ο κοινός στόχος είναι η κατανόηση και περιγραφή της νόησης. Είναι δεδομένη η συλλογική δέσμευση από τη φυσική πραγματικότητα όπως αυτή εκφράζεται στα εμπειρικά στοιχεία των πειραμάτων, παρατηρήσεων και παραδειγμάτων. Η ουσιαστική πρόοδος θα φανεί μέσα από τη σύγκλιση των τριών αξόνων, όχι μεθοδολογικά αλλά στις θεωρητικές κατασκευές για την περιγραφή του νου και των νοητικών διεργασιών, όπως αυτές υλοποιούνται στον εγκέφαλο.

Υπολογιστικά μοντέλα νόησης

Η θεωρητική προσέγγιση της μελέτης του νου, που αναπτύχθηκε στο πρώτο μέρος των σημειώσεων αυτών, βρίσκει την εφαρμογή της σε πραγματικά μοντέλα, υλοποιημένα σε υπολογιστή, τα οποία προσομοιώνουν τις νοητικές λειτουργίες βασισμένα σε ορισμένες θεμελιώδεις υποθέσεις. Η δημιουργία μοντέλων σε υπολογιστή αποτελεί βασικό πλεονέκτημα της συνολικής γνωσιακής προσέγγισης έναντι των ιστορικά συστατικών κλάδων της διότι επιτρέπει τη μελέτη σε βάθος των νοητικών λειτουργιών και τη συσχέτιση μεταξύ τους. Επιχειρείται δηλαδή η περιγραφή μιας ποικιλίας διεργασιών και φαινομένων μέσα από γενικές αρχιτεκτονικές και ενιαία πλαίσια δομής και λειτουργίας. Η μελέτη της λειτουργίας των μοντέλων βοηθά στην κατανόηση των γενικών θεωρητικών αρχών, ενώ η μελέτη των αρχών είναι απαραίτητη για την κατανόηση και την αξιολόγηση των μοντέλων.

Από τις πολλές διαθέσιμες επιλογές μοντέλων που μπορεί κανείς να βρει στη βιβλιογραφία ξεχωρίζουν δύο, μία συμβολική και μία συνδεδετιστική. Πρόκειται για τα μοντέλα ACT-R και *tlearn*, τα οποία θα πρέπει μάλλον να χαρακτηριστούν ως γενικευμένα περιβάλλοντα πειραματισμού και προσομοίωσης νοητικών λειτουργιών παρά συγκεκριμένα μοντέλα. Πρόκειται δηλαδή για προτάσεις γενικευμένης αρχιτεκτονικής και δυνατότητες καθορισμού συγκεκριμένων σχέσεων αναπαραστάσεων και παρατηρήσεων. Οι δύο αυτές προτάσεις επιλέχθηκαν εξαιτίας της ευρείας χρήσης τους σε πλήθος νοητικών φαινομένων αλλά και τις διαθεσιμότητάς τους, δηλαδή της διάθεσης από τους κατασκευαστές ελεύθερου κώδικα λογισμικού και περιβάλλοντος αλληλεπίδρασης με το καθεαυτό λογισμικό, που διευκολύνουν σημαντικά τον πειραματισμό από άλλους μελετητές. Έτσι καθένας μπορεί να κατεβάσει και να εγκαταστήσει στον υπολογιστή του τα σχετικά προγράμματα και να έχει πρόσβαση στις ιδιαίτερες λειτουργίες και βασικές δομικές και λειτουργικές αρχές κάθε αρχιτεκτονικής για την προσαρμογή στα δικά του εμπειρικά δεδομένα και τις αντίστοιχες θεωρητικές υποθέσεις.

Στις ενότητες που ακολουθούν παρουσιάζονται συνοπτικά τα δύο αυτά περιβάλλοντα-μοντέλα νόησης. Και τα δύο έχουν χρησιμοποιηθεί από ερευνητές εκτός από τους κατασκευαστές τους και έχουν εφαρμοστεί σε πλήθος προβλημάτων. Η επιλογή των γνωστικών πεδίων που έχουν προσεγγιστεί με κάθε μοντέλο περιορίζεται σε κάποιο βαθμό από τις ιδιαιτερότητες των μοντέλων όσο και από τις προτιμήσεις των ερευνητών. Σίγουρα τα διαθέσιμα μοντέλα που βασίζονται σε κάθε αρχιτεκτονική διαφέρουν αρκετά όχι μόνο στην υλοποίησή τους αλλά και στο είδος των νοητικών φαινομένων στα οποία αφορούν. Για τη διευκόλυνση της κατανόησης των ιδιαίτερων χαρακτηριστικών κάθε προσέγγισης κρίνεται σκόπιμη η παρουσίαση μιας απόπειρας προσέγγισης του ίδιου φαινομένου και με τις δύο αρχιτεκτονικές. Έτσι μπορεί κανείς να διαπιστώσει τις ομοιότητες και τις διαφορές στις προσεγγίσεις καθώς και να εντοπίσει τα ουσιώδη και τα επουσιώδη σημεία όπου διαφέρουν. Για την εισαγωγή στη λειτουργία κάθε περιβάλλοντος χρησιμοποιείται το παράδειγμα της πρόσθεσης δυαδικών αριθμών που παρουσιάστηκε στην ενότητα 2.1 περί συμβόλων και υπολογισμών. Το παράδειγμα αυτό μπορεί να υλοποιηθεί με ένα πολύ

απλό μοντέλο σε κάθε περίπτωση, στο οποίο αποσαφηνίζονται οι βασικοί όροι λειτουργίας και χρήσης του περιβάλλοντος. Εν συνεχεία παρουσιάζεται η προσέγγιση, μέσα από κάθε αρχιτεκτονική, του προβλήματος του αορίστου των ρημάτων της αγγλικής γλώσσας, ένα πεδίο τριβής μεταξύ συμβολισμού και συνδεδεισμού εδώ και είκοσι πέντε περίπου χρόνια. Για την καλύτερη κατανόηση των ουσιαστικών διαφορών των δύο προσεγγίσεων προηγείται η ιστορική παρουσίαση και ανάλυση του προβλήματος του αορίστου και η ανάλυση των ουσιαστικών σημείων της διαμάχης που στρέφεται γύρω απ' αυτό.

13 Ο αόριστος των αγγλικών ρημάτων

13.1 Εισαγωγή στο πρόβλημα

Ο παρελθοντικός χρόνος της αγγλικής γλώσσας αντίστοιχος του ελληνικού αορίστου (past tense) προσφέρει ένα σπάνιο πεδίο αντιπαράθεσης κλασικών υπολογιστικών προτύπων και συνδεδεισμού διότι συνδυάζει κανονικότητες που προσφέρονται για τυπική (φορμαλιστική) περιγραφή, στα ομαλά ρήματα, με στοιχεία καθαρά συνειρμικά, στα ανώμαλα ρήματα (ή «εξαιρέσεις»). Συγκεκριμένα, για τη δημιουργία της παρελθοντικής μορφής ενός *ομαλού* ρήματος προστίθεται στην παροντική μορφή (του ενεστώτα) το επίθημα /d/, το οποίο μπορεί να πάρει την άφωνη (voiceless) μορφή [t] ή τη φωνούμενη (voiced) [d], αν ο τελικός φθόγγος της παροντικής μορφής είναι άφωνος ή φωνούμενος αντίστοιχα, ή τη μορφή [ɪd], αν ο φθόγγος αυτός είναι και ο ίδιος /t/ ή /d/.⁴⁸ Η αφομοίωση της φώνησης βρίσκεται σε απόλυτη αντιστοιχία με φωνολογικά φαινόμενα που παρατηρούνται σε άλλες περιστάσεις, όπως π.χ. και στο σχηματισμό του πληθυντικού αριθμού των ονομάτων (με αντίστοιχο μορφολογικό επίθημα το /z/). Η παρελθοντική μορφή ενός *ανώμαλου* ρήματος είναι εν πολλοίς αυθαίρετη αλλά αφενός βρίσκεται σχεδόν πάντα σε στενή φωνολογική συνάφεια με την παροντική μορφή, αφετέρου παρατηρούνται «οικογένειες» φωνολογικά συγγενών ρημάτων τα οποία σχηματίζουν με τον ίδιο ανώμαλο τρόπο τη παρελθοντική μορφή τους.

13.2 Η συμβολική προσέγγιση

Σύμφωνα με την κλασική θεώρηση, ο ομαλός σχηματισμός του παρελθοντικού χρόνου περιγράφεται από έναν κανόνα της μορφής $[s \rightarrow s + p]$ όπου η παροντική μορφή s μετατρέπεται στην παρελθοντική με τη *διατήρηση* του θέματος s και την προσθήκη ενός *σταθερού* επιθήματος p , το οποίο εν συνεχεία υφίσταται κάποια φωνολογική μετατροπή. Είναι πολύ σημαντικό να σημειωθεί ότι η φορμαλιστική λειτουργία του κανόνα αυτού συνεπάγεται την πλήρη ανεξαρτησία από τη φωνολογική μορφή του κάθε ρήματος. Ό,τι εμφανιστεί στο αριστερό σκέλος του κανόνα μεταγράφεται αυτόματα στη νέα μορφή με τον ίδιο τρόπο, την ίδια ταχύτητα, και ενδεχομένως τις ίδιες ατέλειες ή πιθανότητες σφάλματος, και παράγει παρελθοντική φωνολογική μορφή η οποία είναι υποχρεωτικά αποδεκτή και

⁴⁸Π.χ. fish→fished: /fɪʃ/→/fɪʃt/; fill→filled: /fɪl/→/fɪld/; fit→fitted: /fɪt/→/fɪtɪd/.

εξίσου φυσιολογική για οποιοδήποτε ρήμα. Επιπλέον, η φορμαλιστική φύση του κανόνα υπαγορεύει την απουσία απομνημόνευσης σε φωνολογικό επίπεδο. Δηλαδή τα ομαλά ρήματα δεν βρίσκονται πουθενά αποθηκευμένα ως τέτοια, αλλά η εφαρμογή του κανόνα γίνεται αυτομάτως για οποιοδήποτε ρήμα δεν έχει χαρακτηριστεί ανώμαλο. Συνεπώς ένα άγνωστο ρήμα (ή μια ψευδολέξη) που δεν μπορεί να έχει καταχωριστεί ως ανώμαλο, οφείλει να δεχτεί την εφαρμογή του κανόνα ομαλών ρημάτων όπως ακριβώς και κάθε γνωστό, σύνηθες ή σπάνιο, μη ανώμαλο ρήμα.

Η κατηγορία των ανώμαλων ρημάτων δεν υπόκειται σε κανόνα μεταγραφής για το σχηματισμό της παρελθοντικής μορφής. Τα ρήματα αυτά θεωρούνται αποθηκευμένα αυτούσια (ως λεξιλογικά λήμματα με τη φωνολογική μορφή τους, τη σημασιολογία τους κλπ.) σε μια συνειρμική μνήμη, συνδεδεμένα με την παρελθοντική μορφή τους, έτσι ώστε η αναζήτηση με βάση την παροντική μορφή ενός γνωστού ρήματος να παράγει την παρελθοντική μορφή. Παρά την εν γένει αυθαίρετη σχέση μεταξύ των δύο μορφών, τα περισσότερα ανώμαλα ρήματα μπορούν να ταξινομηθούν σε ολιγάριθμες κατηγορίες κοινών φαινομένων (συγγενούς φωνολογίας), όπως είναι η απουσία μεταβολής (π.χ. hit: /hit/→/hit/), η μεταβολή μέσου φωνήεντος (π.χ. sing: /siŋ/→/sæŋ/) ή τελικού φωνήεντος (π.χ. know: /no/→/nu/), κ.ά. Με την υπόθεση της απομνημόνευσης των ανώμαλων ρημάτων μπορούμε να εξηγήσουμε ορισμένα φαινόμενα που εξαρτώνται από τη φωνολογική μορφή τους. Για παράδειγμα, σπάνια ομαλά ρήματα (όπως π.χ. via) και κατασκευασμένα ψευδο-ρήματα (π.χ. spring) ενίοτε παράγονται (ή θεωρούνται αποδεκτά) με ανώμαλη παρελθοντική μορφή, κατά αναλογία με φωνολογικά παρόμοια ανώμαλα ρήματα (vought, sprang· πρβλ. buy, spring, αντίστοιχα). Ο συνειρμικός μηχανισμός μπορεί επίσης να δικαιολογήσει και φαινόμενα εξάρτησης από τη συχνότητα εμφάνισης των ρημάτων.

Με το διπλό αυτό μηχανισμό η κλασική προσέγγιση επιτυγχάνει την περιγραφή των δύο κατηγοριών ρημάτων χωρίς να θυσιάσει την ύπαρξη ενός φορμαλιστικού κανόνα στον οποίο να αποδίδεται η παρατηρούμενη κανονικότητα (των ομαλών ρημάτων). Επιπλέον, η υπόθεση του διπτού μηχανισμού είναι συνεπής και με αναπτυξιακές παρατηρήσεις στο σχηματισμό του αορίστου, οι οποίες παρουσιάζουν εξαιρετικό ενδιαφέρον. Συγκεκριμένα, η αναπτυξιακή καμπύλη σφαλμάτων στην παραγωγή του παρελθοντικού χρόνου από μικρά παιδιά περιλαμβάνει μια περίοδο εκ πρώτης όψεως αδικαιολόγητης «επιδείνωσης» της ικανότητας σχηματισμού του αορίστου ορισμένων ρημάτων.

Εν γένει, τα σφάλματα μπορούν να διακριθούν σε περιπτώσεις κατά τις οποίες ένα ανώμαλο ρήμα αποδίδεται στον αόριστο ως ομαλό και σε περιπτώσεις κατά τις οποίες ένα ομαλό ρήμα αποδίδεται με ανώμαλο σχηματισμό αντί της προσθήκης του ομαλού αοριστικού επιθήματος. Είθισται ο όρος *υπεργενίκευση* να χρησιμοποιείται για την πρώτη περίπτωση, μια και βάσει της κλασικής θεώρησης πρόκειται για κυριολεκτική υπεργενίκευση στην εφαρμογή ενός υφιστάμενου κανόνα, ενώ στη δεύτερη περίπτωση δεν ορίζεται αντίστοιχο αντικείμενο γενίκευσης. Και στις δύο περιπτώσεις μπορεί να θεωρηθεί ότι πρόκειται για εσφαλμένη εφαρμογή κάποιου στατιστικού προτύπου: είτε με εφαρμογή ενός γενικού «κανόνα» των ομαλών σε ένα ανώμαλο ρήμα, είτε με χειρισμό ενός ομαλού ρήματος

σύμφωνα με παρόμοια ανώμαλα που σχηματίζουν μια φωνολογική «οικογένεια».

13.3 Το ζήτημα της αναπτυξιακής παλινδρόμησης

Τα γλωσσικά δεδομένα που έχουν καταγραφεί σε μελέτες μικρών παιδιών δείχνουν ότι τα σφάλματα υπεργενίκευσης (δηλαδή σχηματισμού του αορίστου ενός ανώμαλου ρήματος με χρήση του «κανόνα» των ομαλών) παρουσιάζουν μια περίοδο αύξησης σε ένα ενδιάμεσο αναπτυξιακό στάδιο. Συγκεκριμένα, μετά από μια αρχική περίοδο κατά την οποία κατακτώνται οι μορφές και χρήσεις των αοριστικών μορφών ορισμένων συχνών (ομαλών και ανώμαλων) ρημάτων, ακολουθεί η φάση της μεγάλης ανάπτυξης του λεξιλογίου και της κατάκτησης της δυνατότητας σχηματισμού του ομαλού αορίστου σε αυθαίρετο ρηματικό θέμα. Στη δεύτερη αυτή φάση παρατηρούνται ορισμένες υπεργενικεύσεις σε ανώμαλα ρήματα των οποίων η ορθή αοριστική μορφή ήταν γνωστή κατά την πρώτη φάση, δηλαδή το παιδί έλεγε σωστά τον ανώμαλο αόριστο. Εκ πρώτης όψεως φαίνεται πως το παιδί «ξεχνά» κάτι που ήξερε – βέβαια τα πράγματα δεν είναι τόσο απόλυτα, μια και τέτοια λάθη συμβαίνουν με χαμηλή συχνότητα και όχι συστηματικά, αλλά το σημαντικό είναι ότι συμβαίνουν. Σε τρίτη φάση κατακτάται πλήρως το σύστημα των αορίστων, με τις σωστές αοριστικές μορφές να παράγονται για τα γνωστά ανώμαλα ρήματα και το ομαλό επίθημα να προστίθεται σε οποιαδήποτε άλλη, γνωστή ή άγνωστη, ενεστωτική ρηματική μορφή.

Η υπόθεση του διπτού μηχανισμού αποδίδει την αναπτυξιακή αυτή πορεία στο γεγονός της κατάκτησης του κανόνα για τα ομαλά ρήματα. Συγκεκριμένα, κατά την πρώτη φάση υποτίθεται ότι ενεστωτικές και αοριστικές μορφές απομνημονεύονται ως διακριτά λήμματα, ενδεχομένως με ανεξάρτητες σημασιολογικές διασυνδέσεις, τόσο για τα ομαλά όσο και για τα ανώμαλα ρήματα του πρώιμου γλωσσικού περιβάλλοντος. Η διαδικασία αυτή δικαιολογείται δεδομένου ότι μεγάλο μέρος των πιο συχνών ρημάτων της γλώσσας είναι ανώμαλα και συνεπώς δεν υφίσταται στατιστική βάση γενίκευσης κάποιας ομαλής αοριστικής μορφολογίας. Στη συνέχεια, με τη διεύρυνση του λεξιλογίου κατακτάται ο φορμαλισμός του σχηματισμού του αορίστου ως κανόνας, η ισχυροποίηση του οποίου (κάτω και από το ισχυρό στατιστικό βάρος του πλήθους των ομαλών ρημάτων) έχει ως παρενέργεια την ενίστε εσφαλμένη εφαρμογή του ακόμα και σε ανώμαλα ρήματα. Η δικαιολόγηση αυτής της φάσης συμπεριλαμβάνει και τη λημματική διάκριση μεταξύ ενεστωτικής και αοριστικής μορφής, διότι ενδέχεται η σύνδεση των λημμάτων να μην έχει ολοκληρωθεί ακόμα και έτσι να υφίσταται σημαντικό ανταγωνισμό από τον ισχυροποιούμενο κανόνα. Στην τρίτη και τελευταία φάση, τα αυθαίρετα ζεύγη ενεστωτικής και αοριστικής μορφής των ανώμαλων ρημάτων έχουν συνδεθεί ισχυρά ενώ έχει παγιωθεί και ο κανόνας για το σχηματισμό του ομαλού αορίστου προς εφαρμογή σε κάθε ρήμα που δεν είναι ρητά γνωστό ως ανώμαλο.

Η υπόθεση του διπτού μηχανισμού έχει ως κεντρικό στόχο τη διάσωση της φορμαλιστικής λύσης για τον ομαλό αόριστο. Αν τα ομαλά και τα ανώμαλα ρήματα μπορούν να σχηματίσουν τους αορίστους τους με τον ίδιο, έναν και μοναδικό, μηχανισμό τότε αυτός ο απλός μηχανισμός θα πρέπει να προτιμηθεί του διπτού βάσει της αρχής της απλότητας. Προτού όμως χρειαστεί να καταφύγουμε σε τέτοιες φιλοσοφικές μεθόδους για την επιλογή

της καλύτερης θεωρίας, είναι απαραίτητο να εξαντληθούν τα περιθώρια εμπειρικής μελέτης για τη διάκριση μεταξύ των υποψηφίων. Συγκεκριμένα, πρέπει να δειχθεί αν είναι καταρχήν δυνατό να αποδοθεί σε έναν απλό μηχανισμό όλο το φάσμα των αναπτυξιακών και τελικών παρατηρήσεων, και μάλιστα υπό φυσιολογικούς περιορισμούς π.χ. ως προς τη συχνότητα και κατανομή των διαθεσίμων ρηματικών μορφών και τις χρονικές (αναπτυξιακές) παραμέτρους. Επίσης πρέπει να αποκαλυφθούν, να αναλυθούν και μελετηθούν συστηματικά με εμπειρική παρατήρηση των φυσικών συστημάτων (δηλαδή των ομιλητών της αγγλικής, σε παιδική ηλικία και ως ενήλικες) τα σημεία διαφοροποίησης των προβλέψεων που προκύπτουν από την υπόθεση του διπτού και αυτήν του απλού συστήματος.

13.4 Συνδεδειστικά μοντέλα του αοριστικού συστήματος

Το 1986 η συλλογή με το γενικό τίτλο «παράλληλη κατανεμημένη επεξεργασία» (parallel distributed processing)⁴⁹ επανέφερε τα τεχνητά νευρωνικά δίκτυα στο προσκήνιο των μαθηματικών μεθόδων προσέγγισης των ψυχολογικών φαινομένων. Το δίτομο αυτό έργο μπορεί να θεωρηθεί το μανιφέστο του συνδεδεισμού, εκφράζοντας μια συγκεκριμένη θεωρητική τοποθέτηση και πλαισιώνοντάς την με πλήθος πρακτικών παραδειγμάτων. Ένα από τα κεφάλαια που έχουν τραβήξει περισσότερο την προσοχή αναφέρεται στο μορφολογικό σχηματισμό του παρελθοντικού χρόνου («αορίστου») της αγγλικής γλώσσας. Σε αυτό οι Rumelhart και McClelland διατυπώνουν την άποψη ότι αρκεί ένας μηχανισμός για το σχηματισμό του αορίστου τόσο των ομαλών όσο και των ανωμάλων ρημάτων και παρουσιάζουν ένα συνδεδειστικό σύστημα ως απόδειξη για τον ισχυρισμό αυτό. Το συγκεκριμένο σύστημα είναι ένα συνειρμικό δίκτυο δύο επιπέδων, στα οποία ενεργοποιούνται κατανεμημένες αναπαράστασεις του παροντικού και παρελθοντικού τύπου, και δύο υποστηρικτικών επιπέδων εισροής και εκροής με προκαθορισμένες συνδέσεις, τα οποία μετατρέπουν μια φωνολογική αναπαράσταση του ενεστώτα σε μια αντίστοιχη του αορίστου.

Το μοντέλο αυτό δέχτηκε ισχυρή και εξαντλητική κριτική⁵⁰ διότι ορισμένες από τις επιθυμητές ιδιότητές του ήταν απόρροια του προγράμματος «διδασκαλίας» του μοντέλου και όχι εγγενές χαρακτηριστικό της λειτουργίας του. Σε μια σειρά εργασιών που ακολούθησαν κατά τη δεκαετία του 1990, οι Plunkett και Marchman παρουσίασαν νέα συνδεδειστικά δίκτυα, αντιμετωπίζοντας τα προβλήματα της πρώτης απόπειρας των Rumelhart και McClelland, ενώ έκτοτε έχουν παρουσιαστεί και νεότερες βελτιώσεις από διάφορους ερευνητές, καθώς και δίκτυα κοινής επεξεργασίας του παρελθοντικού χρόνου και του πληθυντικού αριθμού. Παρότι η σχετική διαμάχη δεν έχει ακόμα λήξει, είναι πλέον σαφές πως το εύρος των σχετικών παρατηρήσεων μπορεί να περιγραφεί τόσο από ένα κλασικό μοντέλο διπτού μηχανισμού όσο και από ένα συνδεδειστικό μοντέλο στο οποίο δεν υπάρχουν φορμαλιστικοί κανόνες και όλα τα ρήματα υφίστανται ίδιου τύπου μεταχείριση.

⁴⁹Rumelhart, D. E., McClelland, J. L., & the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol 1 & 2). MIT Press.

⁵⁰Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73–193.

14 ACT-R

Το σύστημα ACT-R είναι ένα πλήρες περιβάλλον κατασκευής νοητικών μοντέλων το οποίο φιλοδοξεί να αποδώσει το «γενικό μοντέλο του νου». Περιλαμβάνει μια σειρά από περιοχές προσωρινής αποθήκευσης που ονομάζονται *ταμιευτές* (buffers), μακρόχρονη *αποφαντική* (ή «δηλωτική»-declarative) και *διαδικαστική* (procedural) μνήμη, καθώς και έναν κεντρικό εκτελεστικό τομέα με πρόσβαση στις περιοχές αποθήκευσης. Αποτελεί προϊόν πολύχρονης έρευνας και ανάπτυξης από την ομάδα του John Anderson στο Πανεπιστήμιο Carnegie Mellon.⁵¹ Η γενική αρχιτεκτονική του συστήματος είναι σταθερή και αποτελεί τη βάση πάνω στην οποία κατασκευάζονται τα διάφορα μοντέλα επιμέρους λειτουργιών. Επίσης σταθερές είναι και ορισμένες βασικές υποθέσεις και παράμετροι (π.χ. ο χρόνος εκτέλεσης κάθε βήματος). Έτσι το σύστημα χαρακτηρίζεται από τη γενικότητα που είναι απαραίτητη προκειμένου να υποστηριχθεί ότι εκφράζει τη μοναδική αρχιτεκτονική του νου.

Ο προσδιορισμός ενός συγκεκριμένου μοντέλου περιλαμβάνει την εισαγωγή δύο τύπων γνώσης (πληροφορίας) μέσα στο σύστημα: οι «αποφαντικές» πληροφορίες, που ονομάζονται *ψηγμάτα* (chunks), αποθηκεύονται στον αντίστοιχο μνημονικό χώρο ως συμβολικές παθητικές αναπαραστάσεις, ενώ οι «διαδικαστικές» πληροφορίες, οι οποίες είναι *παραγωγές*⁵² καθορίζουν τις δυνατές δράσεις του συστήματος. Οι συντακτικές ιδιότητες των ψηγμάτων-συμβόλων καθορίζονται από τη δήλωση του «τύπου» τους (chunk type), βάσει του οποίου μπορούν να συμμετέχουν ή όχι στις διάφορες παραγωγές. Η λειτουργία ενός μοντέλου μπορεί να ελέγχεται προγραμματίζοντας το περιβάλλον ACT-R σε γλώσσα Lisp ώστε να παρουσιάζονται ερεθίσματα και να καταγράφονται αποκρίσεις. Η δράση του μοντέλου κάθε στιγμή εξαρτάται από το «στόχο», ένα στοιχείο πληροφορίας που τοποθετείται σε ειδικό ταμιευτή (goal buffer), υποδηλώνοντας μια ισχυρή θεωρητική τοποθέτηση: ότι η λειτουργία του νου ανά πάσα στιγμή εξαρτάται και κατευθύνεται από συγκεκριμένους στόχους, οι οποίοι αποτελούν αποφαντικές πληροφορίες.

14.1 Απλή δυαδική πρόσθεση στο ACT-R

Η απλή δυαδική πρόσθεση περιλαμβάνει τρεις τύπους γνώσης (ψηγμάτων): τους στόχους, τα ψηφία, και τα αθροίσματα συγκεκριμένων ζευγαριών. Ενοείται πως οι ονομασίες των τύπων ψηγμάτων, καθώς και οι ονομασίες των επιμέρους πληροφοριών είναι απλές ετικέτες διάκρισης και το περιεχόμενό τους είναι εντελώς αδιάφορο για τη λειτουργία του μοντέλου, καθώς μιλάμε για συμβολική προσέγγιση. Για πρακτικούς λόγους φυσικά χρησιμοποιούμε ονομασίες που έχουν νόημα για μας στα πλαίσια της επιθυμητής λειτουργίας του μοντέλου. Τίποτα δεν θα άλλαζε στην εσωτερική αλληλουχία καταστάσεων του μοντέλου αν αντί για «ψηφίο» χαρακτηρίζαμε κάποια γνώση ως «καιρό» ή ως «1».

⁵¹<http://act-r.psy.cmu.edu/>

⁵²Παραγωγή ονομάζεται ένας κανόνας ο οποίος αποτελείται από δύο μέρη: μια *συνθήκη* και μια *ενέργεια*. Κάθε κανόνας είναι αυτόνομος, δεν αποτελεί μέρος κάποιας αλληλουχίας ούτε συνδέεται με άλλους. Κάθε κανόνας ελέγχει διαρκώς αν ικανοποιείται η συνθήκη του και όταν συμβεί αυτό εκτελεί την ενέργειά του.

Η δήλωση των τριών τύπων γίνεται ως εξής:

```
(chunk-type bin-digit)
(chunk-type binadd-fact bin1 bin2 sum carry)
(chunk-type binadd-goal bin1 bin2 answer)
```

Το σύστημα μπορεί να έχει μόνο στόχους πρόσθεσης ψηφίων, αφού τέτοιο είναι το μοντέλο. Οι στόχοι αυτοί χαρακτηρίζονται από την αναζήτηση της απάντησης (*answer*) που αντιστοιχεί σε δύο συγκεκριμένα ψηφία (*bin1* και *bin2*). Ειδικότερα, το σύστημα μπορεί να υποστηρίξει τους εξής στόχους, καθένας από τους οποίους αντιστοιχεί στη λύση ενός από τους τέσσερις συνδυασμούς ψηφίων:

```
(add-dm
 (do-binadd-00 isa binadd-goal bin1 zero bin2 zero answer nil)
 (do-binadd-01 isa binadd-goal bin1 zero bin2 one answer nil)
 (do-binadd-10 isa binadd-goal bin1 one bin2 zero answer nil)
 (do-binadd-11 isa binadd-goal bin1 one bin2 one answer nil)
)
```

Οι στόχοι αυτοί, ως στοιχεία πληροφορίας, προστίθενται στην αποφαντική μνήμη του συστήματος με την εντολή *add-dm*. Για τα δύο άλλα είδη γνώσης οι αντίστοιχες δηλώσεις στο σύστημα είναι:

```
(add-dm
 (zero isa bin-digit)
 (one isa bin-digit)
 (zero-zero isa binadd-fact bin1 zero bin2 zero sum zero carry zero)
 (zero-one isa binadd-fact bin1 zero bin2 one sum one carry zero)
 (one-zero isa binadd-fact bin1 one bin2 zero sum one carry zero)
 (one-one isa binadd-fact bin1 one bin2 one sum zero carry one)
)
```

Κάθε δήλωση που προστίθεται στην αποφαντική μνήμη περιλαμβάνει το είδος της γνώσης που αποτελεί (πεδίο *isa*) και την αντιστοίχιση των επιμέρους πεδίων για το συγκεκριμένο είδος γνώσης με τις κατάλληλες τιμές. Έτσι, για να ορίσουμε ένα σύμβολο που αντιστοιχεί σε γνώση του ψηφίου μηδέν, δηλώνουμε απλώς ένα ψήγμα τύπου *bin-digit* με όνομα *zero*, ενώ για να ορίσουμε ένα σύμβολο που να αντιστοιχεί στη γνώση πως $1 + 0 = 0$ δηλώνουμε ένα ψήγμα με όνομα *one-zero*, τύπου *binadd-fact*, με τις κατάλληλες τιμές στα πεδία προσθετέων (*bin1 one* και *bin2 zero*), αθροίσματος (*sum one*) και κρατούμενου (*carry zero*). Ομοίως ορίζονται και τα υπόλοιπα ψήγματα που είναι απαραίτητα.

Για τη λειτουργία του συστήματος είναι απαραίτητος κάποιος κανόνας παραγωγής, ο οποίος θα εκπληρώνει τους στόχους άθροισης όποτε αυτοί ενεργοποιούνται.

```
(p binadd-lookup
  =goal>
    isa      binadd-goal
    bin1     =bindigit1
    bin2     =bindigit2
    answer   nil
==>
  =goal>
    answer   retrieving
+retrieval>
  isa      binadd-fact
  bin1     =bindigit1
  bin2     =bindigit2
)
```

Κάθε παραγωγή αποτελείται από δύο μέρη, τα οποία χωρίζονται μεταξύ τους με το σύμβολο ==>. Το πρώτο μέρος είναι η συνθήκη ενεργοποίησης και το δεύτερο η ενέργεια που εκτελείται όταν εκπληρώνεται η συνθήκη. Εδώ, όταν διαπιστωθεί η ύπαρξη ενός στόχου τύπου άθροισης (binadd-goal), αναζητείται η κατάλληλη απάντηση στη μνήμη με απευθείας αναζήτηση της σχετικής πληροφορίας. Η αναζήτηση γίνεται με βάση τα στοιχεία του στόχου, δηλαδή τους προσθετέους, οι οποίοι στο συγκεκριμένο παράδειγμα αναπαριστώνται με τις ονομασίες bindigit1 και bindigit2. Παράλληλα, η παραγωγή ενημερώνει και την κατάσταση του στόχου, καθώς δεν βρίσκεται πια στο αρχικό βήμα αλλά στη φάση της αναζήτησης, και αυτό πρέπει να είναι ευρέως γνωστό στο σύστημα διότι η τρέχουσα κατάσταση καθορίζει ανά πάσα στιγμή τη λειτουργία. Όταν η ανάσυρση της πληροφορίας είναι επιτυχής τότε ενεργοποιείται μια δεύτερη παραγωγή για την εκτύπωση της απάντησης:

```
(p binadd-printans
  =goal>
    isa      binadd-goal
    answer   retrieving
=retrieval>
  isa      binadd-fact
  sum      =bindigit1
  carry    =bindigit2
==>
  =goal>
    answer   done
  !output!  (the sum is =bindigit2 =bindigit1)
)
```

Συγκεκριμένα, όταν αναγνωριστεί ότι υπάρχει ενεργός στόχος άθροισης στη φάση της αναζήτησης (retrieving) και έχει ανασυρθεί από τη μνήμη επιτυχώς η κατάλληλη πληροφορία,

παρουσιάζεται η απάντηση και απενεργοποιείται ως εκπληρωμένος ο στόχος.

Το συγκεκριμένο «μοντέλο» είναι φυσικά υπεραπλουστευμένο, εφόσον το μόνο που κάνει είναι να αναζητάει έτοιμες απαντήσεις από τη μνήμη, είναι όμως εύκολα κατανοητό δείχνοντας συγχρόνως όλα τα βασικά μέρη του προσδιορισμού και της λειτουργίας ενός μοντέλου στο ACT-R.

14.2 Ο αόριστος των ρημάτων στο ACT-R

Το πρόβλημα του αορίστου των ρημάτων στο ACT-R μπορεί να λυθεί με συνδυασμό αποθηκευμένης γνώσης, για τα ανώμαλα ρήματα, και κανόνων, για το σχηματισμό των ομαλών αορίστων. Αυτή είναι η πιο απλή αλλά λιγότερο ενδιαφέρουσα περίπτωση, διότι χάρη στους μηχανισμούς «διαδικασιοποίησης» (proceduralization) το ACT-R μπορεί να αποκομίσει νέους κανόνες από σύνολα γνώσεων, όταν αυτά έχουν κοινά στοιχεία που χρησιμοποιούνται συστηματικά. Έτσι οι σχετικές μελέτες έχουν επικεντρωθεί στη δυνατότητα σχηματισμού κανόνων για την καλύτερη περιγραφή των αναπτυξιακών δεδομένων.⁵³ Συγκεκριμένα, η επανειλημμένη έκθεση σε ζεύγη ενεστώτα-αορίστου, παράλληλα με απόπειρες παραγωγής του αορίστου με βάση τον ενεστώτα, μπορεί να οδηγήσει ένα σύστημα στο σχηματισμό του κατάλληλου κανόνα.

Βέβαια, το σύστημα δεν συμπεραίνει από μόνο του ότι πρόκειται για μορφολογική σχέση ούτε ότι το κρίσιμο στοιχείο είναι επίθημα. Η αρχική γνώση που παρέχεται στο μοντέλο περιλαμβάνει μια σειρά διαφορετικών στρατηγικών για την παραγωγή του αορίστου: Η πρώτη στρατηγική είναι η μηδενική στρατηγική. Δηλαδή, το σύστημα επιστρέφει ως αόριστο το ίδιο το ενεστωτικό θέμα. Η στρατηγική αυτή δίνει σωστό αποτέλεσμα μόνο για εκείνα τα ανώμαλα ρήματα που δεν αλλάζουν μορφή στον αόριστο. Η δεύτερη στρατηγική είναι με ανάκληση. Δηλαδή το σύστημα αναζητά στην αποφαντική μνήμη το ρήμα που καλείται να επεξεργαστεί, ώστε να επαναλάβει την ίδια απάντηση που είχε δώσει στο παρελθόν, καθώς στην περίπτωση αυτή θεωρείται ότι «το ξέρει».⁵⁴ Η τρίτη στρατηγική είναι με αναλογία. Δηλαδή το σύστημα αναζητά ένα οποιοδήποτε ρήμα στην αποφαντική μνήμη το οποίο να έχει σχηματιστεί με επίθημα, και χρησιμοποιεί το ίδιο επίθημα. Ουσιαστικά η στρατηγική αυτή κωδικοποιεί εξαρχής στο σύστημα την τάση να μάθει επιθήματα, χωρίς όμως να δίνεται κάποιο επίθημα ως γενικός κανόνας και χωρίς να προσδιορίζεται αν υπάρχουν περισσότερα από ένα πιθανά επιθήματα—αυτά μαθαίνονται στην πορεία.

Η υλοποίηση της στρατηγικής της αναλογίας γίνεται με δύο παραγωγές: Η μία παραγωγή αναζητά ψήγμα γνώσης με επίθημα και η δεύτερη εφαρμόζει το επίθημα που ανασύρθηκε. Η επανειλημμένη χρήση αυτής της στρατηγικής οδηγεί στην ένωση των δύο παραγωγών, διότι το ανασυρόμενο επίθημα είναι πάντοτε το ίδιο, εκείνο του ομαλού αο-

⁵³Taatgen, N. A., & Anderson, J. R. (2002). Why do children learn to say “Broke”? A model of learning the past tense without feedback. *Cognition*, 86, 123–155.

⁵⁴Στην πραγματικότητα η προηγούμενη απάντηση μπορεί να μην ήταν σωστή, εξαιτίας της ύπαρξης των άλλων στρατηγικών, όμως σταδιακά θα επικρατήσουν οι σωστές απαντήσεις λόγω της έκθεσης σε σωστά ζεύγη.

ρίστου. Έτσι, κάποια στιγμή το σύστημα σχηματίζει μια παραγωγή που αντί να αναζητά επίθημα εφαρμόζει απευθείας το επίθημα που συστηματικά ανασύρεται (το -ed). Αυτή η παραγωγή υλοποιεί τον κανόνα της ομαλής κλίσης. Σταδιακά η παραγωγή αυτή ισχυροποιείται διότι οδηγεί σε απάντηση ταχύτερα από την απόπειρα ανάκλησης του ρήματος και επομένως είναι πιο αποδοτική. Στη φάση αυτή παράγονται τα φαινόμενα υπεργενίκευσης.

Παράλληλα, για τα ανώμαλα ρήματα ισχυροποιείται σταδιακά η γνώση της αοριστικής μορφής καθώς και αυτή ανακαλείται συστηματικά όταν γίνεται αναζήτηση με το κάθε συγκεκριμένο ρήμα. Σε συνδυασμό με την αυξανόμενη διαθεσιμότητα των σωστών ζευγών (λόγω της αντιληπτικής έκθεσης σε αυτά, καθώς πρόκειται για ρήματα σχετικά μεγάλης συχνότητας), οι παραγωγές για τα ανώμαλα ρήματα ενισχύονται τόσο ώστε να ξεπεράσουν σε ισχύ την παραγωγή του γενικού κανόνα. Τότε εξαφανίζονται οι υπεργενικεύσεις και το σύστημα έχει πλέον φτάσει στη φάση ώριμης λειτουργίας.

Η προσέγγιση αυτή δείχνει ότι ένα συμβολικό σύστημα με διαβάθμιση της ισχύος των παραγωγών και της διαθεσιμότητας της αποφαντικής γνώσης μπορεί να σχηματίσει νέες παραγωγές βάσει των επανειλημμένων αλληλουχιών που χρησιμοποιεί, όταν παρατηρείται συστηματικότητα στα αποτελέσματά τους. Έτσι μια λειτουργία που αρχικά βασίζεται σε περισσότερα βήματα (δηλαδή σε εκτέλεση περισσότερων παραγωγών), με ενδιάμεση ανάκληση αποφαντικών πληροφοριών, σταδιακά διαδικασιοποιείται ώστε να εκτελείται άμεσα με εξειδικευμένες παραγωγές που συνδυάζουν το αποτέλεσμα των περισσότερων αρχικών βημάτων σε ένα, παρακάμπτοντας εντελώς την ανάκληση, άρα και την εμπλοκή της αποφαντικής μνήμης.

15 tlearn

Το tlearn δεν είναι ένα μοντέλο αλλά ούτε και μια γενικευμένη αρχιτεκτονική του νου όπως το ACT-R. Είναι απλώς ένα περιβάλλον περιγραφής και χειρισμού συνδεδεστικών δικτύων, δηλαδή ένα πρόγραμμα στο οποίο μπορούμε να περιγράψουμε σύνολα διασυνδεδεμένων κόμβων και τις δύο βασικές λειτουργίες της «εκπαίδευσης» και «δοκιμασία». Η *εκπαίδευση* συνίσταται στη ρύθμιση των βαρών των συνδέσεων μεταξύ των κόμβων με αλγόριθμους εποπτευόμενης διδασκαλίας (βλ. Ενότητα 3.3, σελ. 29–32). Η *δοκιμασία* ή *έλεγχος* (testing) του δικτύου επιτυγχάνεται με παρουσίαση κάποιου προτύπου στη σιβάδα εισροής και διάδοση της ενεργοποίησης μέσα από τις συνδέσεις ως τη σιβάδα εκροής, όπου μπορεί να ελεγχθεί η απόκλιση από την επιθυμητή απόκριση. Το περιβάλλον tlearn παρέχει πολλές δυνατότητες εξέτασης των δραστηριοτήτων των κόμβων και των βαρών των συνδέσεων, καθώς και δυνατότητες ανάλυσης των προτύπων ενεργοποίησης, ιδιαίτερα για τους κρυφούς κόμβους όπου θεωρούμε ότι σχηματίζονται οι αναπαραστάσεις επίλυσης των προβλημάτων.

Ο προσδιορισμός ενός μοντέλου στο tlearn περιλαμβάνει την αρχιτεκτονική του δικτύου (δηλαδή αριθμό κόμβων και συνδέσεις μεταξύ τους) καθώς και τα πρότυπα ενεργοποίησης που αντιστοιχούν στην εισροή και εκροή, τα οποία πρέπει να μάθει να αντιστοιχίζει

μεταξύ τους το δίκτυο. Το βασικότερο ζήτημα που έχει να επιλύσει κάθε ερευνητής για τον προσδιορισμό ενός συνδεδειστικού μοντέλου είναι ο τρόπος με τον οποίο θα μετατρεφεί το συγκεκριμένο νοητικό πρόβλημα που θέλει να μελετήσει σε αριθμητικές τιμές ενεργοποίησης κόμβων έτσι ώστε να διατηρούνται τα ουσιώδη χαρακτηριστικά του προβλήματος και οι κρίσιμες σχέσεις ομοιότητας μεταξύ των εισροών και των εκροών. Στόχος είναι να παρέχονται στο δίκτυο οι απαραίτητες πληροφορίες του περιβάλλοντος για να δούμε αν θα πετύχει την επιθυμητή γενίκευση, όχι όμως να δίνεται ήδη λυμένο το πρόβλημα.

15.1 Απλή δυαδική πρόσθεση στο tlearn

Η απλή δυαδική πρόσθεση δεν είναι παρά η αντιστοίχιση από δύο δυαδικά ψηφία (τους προσθετέους) σε δύο άλλα (άθροισμα και κρατούμενο). Εφόσον κάθε ψηφίο μπορεί να είναι μόνο μηδέν ή ένα, το πρόβλημα προσφέρεται για την απλή τοπική αναπαράσταση ενός κόμβου για κάθε ψηφίο. Χρειαζόμαστε λοιπόν δύο κόμβους εισροής και δύο κόμβους εκροής. Τα ζεύγη εισροής-εκροής που πρέπει να μάθει το σύστημα είναι τέσσερα, όσα και οι δυνατοί συνδυασμοί των δύο ψηφίων. Άρα ο αριθμός των κρυφών κόμβων θα πρέπει να είναι μικρότερος από τέσσερα, ενώ θα πρέπει να είναι τουλάχιστον δύο ώστε να μπορούν να υπάρξουν περισσότερες από μία διαφορετικές αποκρίσεις στην έξοδο. Η απλούστερη επιλογή, με δύο κρυφούς κόμβους, ορίζεται ως εξής:

NODES:

nodes = 4

inputs = 2

outputs = 2

output nodes are 3-4

CONNECTIONS:

groups = 0

1-2 from 0, i1-i2

3-4 from 1-2

Προσέξτε ότι οι κόμβοι εισροής δεν μετρούν ως κανονικοί κόμβοι, επειδή η ενεργοποίησή τους δεν υπολογίζεται από το σύστημα αλλά παρέχεται από τα πρότυπα του πειραματιστή. Έτσι ο αριθμός των κόμβων που δηλώνεται είναι 4, περιλαμβάνοντας μόνο τους 2 κρυφούς κόμβους (1 και 2) και τους 2 κόμβους εκροής (3 και 4). Οι συνδέσεις είναι πλήρεις μεταξύ των κόμβων ανά στιβάδα, άρα οι δύο κόμβοι εισροής (i1 και i2) συνδέονται με τους κρυφούς 1 και 2, και αυτοί με τη σειρά τους συνδέονται με τις εκροές 3 και 4. Υπάρχει και μια επιπλέον εντολή η οποία δηλώνει απουσία ομαδοποίησης κόμβων, ενώ χρησιμοποιείται και ο κόμβος πόλωσης (bias), με αριθμό 0, για υπολογιστικούς λόγους.

Η μόνη επιπλέον πληροφορία που χρειάζεται για την ολοκλήρωση του μοντέλου είναι τα ζεύγη εισροής-εκροής, τα οποία στην περίπτωσή μας αποτελούνται από τους συνδυασμούς προσθετέων και αθροίσματος-κρατούμενου. Σύμφωνα με τις απαιτήσεις του tlearn

δηλώνουμε δύο χωριστά αρχεία, ένα για τους κόμβους εισροής κι ένα αντίστοιχο, με την ίδια σειρά, για τους κόμβους εκροής. Το πρώτο (data) είναι:

```
distributed
4
0 0
0 1
1 0
1 1
```

Η λέξη distributed δηλώνει ότι αναφέρονται οι τιμές ενεργοποίησης για όλους τους κόμβους, και όχι μόνο για όσους είναι ενεργοί, ενώ ο αριθμός 4 ισούται με το πλήθος των προτύπων ενεργοποίησης που ακολουθούν. Σε κάθε σειρά υπάρχουν τόσα νούμερα (τιμές ενεργοποίησης) όσοι και οι κόμβοι εισροής που δηλώθηκαν με την αρχιτεκτονική του δικτύου. Τέλος, το αντίστοιχο αρχείο εκροής (teach) είναι:

```
distributed
4
0 0
0 1
0 1
1 0
```

Η λειτουργία του μοντέλου, εφόσον είναι συνδεδεστικό, δεν μπορεί παρά να είναι μετάδοση ενεργοποίησης μεταξύ των κόμβων, μέσω των συνδέσεων, άρα δεν υπάρχει κάτι άλλο να προσδιοριστεί. Το δίκτυο που προσδιορίζεται με τον τρόπο αυτό χρειάζεται πολλές επαναλήψεις για να μάθει την απλή αντιστοίχιση των τεσσάρων ζευγών, επειδή ο ένας κόμβος εκροής λειτουργεί ως αποκλειστική διάζευξη, δηλαδή είναι ανενεργός όταν οι δύο κόμβοι εισροής είναι συγχρόνως είτε ενεργοί είτε ανενεργοί. Το πρόβλημα απλουστεύεται με χρήση ενός επιπλέον κρυφού κόμβου αλλά εφόσον δεν υπάρχει κάποια ουσιαστική νοητική ερμηνεία του μοντέλου δεν υπάρχει λόγος βαθύτερης διερεύνησης.

15.2 Ο αόριστος των ρημάτων στο tlearn

Ο σχηματισμός του αορίστου των ρημάτων στο tlearn πρέπει να δηλωθεί ως πρόβλημα αντιστοίχισης σχεδίων ενεργοποίησης κόμβων. Αυτό σημαίνει ότι πρέπει με κάποιον τρόπο να μετατρέψουμε τις λέξεις σε δραστηριότητες ομάδων κόμβων. Θα μπορούσαμε να αντιστοιχίσουμε κάθε γράμμα σε έναν κόμβο ή κάθε ρήμα σε κάποιους συγκεκριμένους κόμβους, κλπ. Όμως η ουσία του προβλήματος βρίσκεται στην εξάρτηση των αορίστων από τους φθόγγους του θέματος και ειδικότερα από ορισμένα φωνητικά χαρακτηριστικά. Για παράδειγμα, η σωστή ομαλή κατάληξη εξαρτάται από το αν ο τελικός φθόγγος του θέματος είναι φωνούμενος, ενώ στα ανώμαλα υπάρχουν ομάδες οι οποίες χαρακτηρίζονται από ομοιότητες στην προφορά (π.χ. ring, sing κλπ.). Άρα η «σωστή» αναπαράσταση

που μπορεί να οδηγήσει σε λύση του προβλήματος θα πρέπει να βασίζεται στα φωνητικά χαρακτηριστικά των ρημάτων. Επιπλέον, μια φωνητική αναπαράσταση είναι πιθανό ότι σχετίζεται και με την πραγματική νοητική αναπαράσταση, εφόσον υποθέτουμε ότι οι λέξεις δεν είναι αυθαίρετα καταχωρισμένες στο νου αλλά βρίσκονται σε συστηματικά δομημένες σχέσεις με βάση τα κοινά χαρακτηριστικά τους.

Ένα πρόβλημα που προκύπτει είναι ότι τα ρήματα διαφέρουν σε μέγεθος (αριθμό φθόγγων) και άρα αν τα αναπαραστήσουμε άμεσα στο δίκτυο θα καταλήξουμε σε ορισμένους κόμβους των οποίων οι συνδέσεις θα λάβουν πολλή εκπαίδευση, επειδή θα ενεργοποιούνται σε όλες τις περιπτώσεις, ενώ οι συνδέσεις άλλων κόμβων θα εκπαιδευτούν ελάχιστα διότι θα ενεργοποιούνται μόνο στα μεγαλύτερου μήκους ρήματα. Όμως το μήκος των ρημάτων δεν είναι ουσιώδες χαρακτηριστικό του προβλήματος του αοριστού, άρα μπορούμε να απλοποιήσουμε το πρόβλημα χρησιμοποιώντας τεχνητά ρήματα σταθερού μήκους. Κατασκευάζουμε λοιπόν, για κάθε πραγματικό ρήμα, ένα άλλο με παρόμοια φωνητικά χαρακτηριστικά αλλά τρεις μόνο φθόγγους (σύμφωνο-φωνήεν-σύμφωνο). Διατηρώντας τις διαφορετικές κατηγορίες ρημάτων (βάσει του τρόπου παραγωγής αοριστού), τις σχετικές αναλογίες τους μέσα στο σύνολο, και τις σχετικές συχνότητές εμφάνισής τους, το παραγόμενο σύνολο αποτελεί ένα μοντέλο του προβλήματος, απαλλαγμένο από το πρόβλημα της διακύμανσης μήκους, στο οποίο όμως αναπαράγεται πιστά η ουσιώδης δομή του υπό εξέταση ζητήματος. Έτσι, παρακάμπτεται η τεχνική δυσκολία χωρίς να τίθεται σε κίνδυνο η εγκυρότητα της προσομοίωσης.

Το δεύτερο ζήτημα που προκύπτει στη συνδυαστική προσέγγιση είναι πώς θα αντιστοιχιστούν οι φωνητικές ιδιότητες σε κόμβους. Αναθέτοντας έξι κόμβους σε κάθε φθόγγο, μπορούμε να προσδιορίσουμε τους φθόγγους με μια κατανεμημένη αναπαράσταση των φωνητικών τους γνωρισμάτων: κάθε κόμβος κωδικοποιεί τοπικά ένα γνώρισμα (όπως σύμφωνο/φωνήεν, παρουσία ή απουσία φωνής, θέση ή τρόπος άρθρωσης κλπ.). Χρειαζόμαστε λοιπόν 18 κόμβους εισροής (3 φθόγγοι × 6 γνωρίσματα), ενώ στην έξοδο χρειαζόμαστε επιπλέον κόμβους εκροής για την κατάληξη, όταν υπάρχει. Εφόσον υπάρχουν 4 δυνατές καταλήξεις (οι 3 ομαλές και η ανώμαλη περίπτωση που δεν έχει κατάληξη) μπορούμε με προσθήκη 2 κόμβων να καλύψουμε όλες τις περιπτώσεις, ανεβάζοντας το συνολικό αριθμό των κόμβων εκροής σε 20.

Με την κωδικοποίηση αυτή το πρόβλημα του δικτύου ανάγεται σε δύο επιμέρους προβλήματα: Το ένα σκέλος αφορά τους 18 κόμβους εκροής που αντιστοιχούν στο αοριστικό θέμα και συνίσταται σε αναπαραγωγή των φωνητικών γνωρισμάτων της εισροής στην εκροή. Δηλαδή, το δίκτυο πρέπει να μάθει να ενεργοποιεί τους κόμβους εκροής σύμφωνα με τις αντίστοιχες ενεργοποιήσεις εισροής, έτσι ώστε να επαναλαμβάνεται το θέμα. Η αναπαραγωγή αυτή υπόκειται σε «εξαιρέσεις» καθώς ορισμένα θέματα (ανώμαλων ρημάτων) δεν πρέπει να αναπαραχθούν επακριβώς αλλά με κάποια διαφορά. Οι περιπτώσεις αυτές αποτελούν ιδιαιτερότητες που επιλύονται βάσει του φωνητικού πλαισίου. Για παράδειγμα, ο φθόγγος /Λ/ στην εισροή παράγει τον φθόγγο /æ/ στην εκροή όταν βρίσκεται μεταξύ των φθόγγων /r/ και /n/, ώστε το run να μετατραπεί σε ran. Τα βάρη από τους παρακεί-

μενους φθόγγους συνδυάζουν τις ενεργοποιήσεις του κρυφού επιπέδου όταν εμφανίζεται ο «ιδιαιτέρος» συνδυασμός φθόγγων έτσι ώστε να παραχθεί η διαφορετική εκροή αντί για τη συνηθισμένη. Με τον τρόπο αυτό το δίκτυο μαθαίνει να αντιστοιχίζει τόσο τις ομαλές όσο και τις ανώμαλες περιπτώσεις στις κατάλληλες εκροές. Επιπλέον, επειδή το έργο που έχει να επιτελέσει για τη συντριπτική πλειονότητα των ρημάτων είναι η ανακατασκευή του θέματος από τα φωνητικά γνωρίσματα των φθόγγων του, ουσιαστικά το δίκτυο έχει μάθει και μπορεί να αντιστοιχίσει οποιαδήποτε αλληλουχία θέματος στην εκροή, είτε αποτελεί γνωστό ρήμα είτε όχι. Έτσι, επιδεικνύει γενίκευση σε νέα ρηματικά θέματα.

Το δεύτερο σκέλος του προβλήματος αφορά τους δύο κόμβους εκροής που κωδικοποιούν την κατάληξη και συνίσταται σε ένα συνδυασμό αντιστοιχίσεων που ορίζονται από τα φωνητικά γνωρίσματα της εισροής. Για συγκεκριμένους συνδυασμούς φθόγγων, η εκροή σε αυτούς τους κόμβους μηδενίζεται, αντιστοιχώντας στα ανώμαλα ρήματα, στα οποία δεν προστίθεται αοριστικό επίθημα. Για όλους τους υπόλοιπους συνδυασμούς, οι κόμβοι αυτοί εξαρτώνται από συγκεκριμένα φωνητικά γνωρίσματα του τρίτου φθόγγου, άρα από τους συγκεκριμένους κόμβους εισροής που κωδικοποιούν αν ο τρίτος και τελευταίος φθόγγος του θέματος είναι φωνούμενος (για να επιλέξει μεταξύ των επιθημάτων /-d/ και /-t/) και αν είναι ο ίδιος /d/ ή /t/ (για να επιλέξει /id/). Από τη σκοπιά αυτή, και στα δύο «υποπροβλήματα» η σωστή απόκριση καθορίζεται από τη φωνολογική σύσταση του θέματος, και γι' αυτό το λόγο είναι κρίσιμο να έχει αυτή αναπαρασταθεί σωστά στην εισροή του δικτύου ώστε να διατηρήσει τις σχέσεις ομοιότητας στις οποίες βασίζεται η λύση του προβλήματος.

Το σύστημα που προκύπτει με τον τρόπο αυτό προσφέρεται για μελέτη των χαρακτηριστικών του σχηματισμού του αορίστου που σχετίζονται με τη στατιστική κατανομή των φωνητικών ιδιοτήτων των ρημάτων, αναπαραγωγή της αναπτυξιακής καμπύλης παλινδρόμησης, μελέτη των λαθών υπεργενίκευσης και ανταπόκρισης σε «νέα» ρήματα με τα οποία δεν εκπαιδεύτηκε το σύστημα. Η εκπαίδευση γίνεται με τον συνηθισμένο τρόπο, με χρήση ενός τυχαίου ζεύγους ενεστώτα-αορίστου σε κάθε δοκιμή, παρουσιάζοντας τη μορφή του ενεστώτα στην εισροή, υπολογίζοντας στην εκροή το σφάλμα βάσει της μορφής του αορίστου, και μεταβάλλοντας τα βάρη με αντίστροφη διάδοση του σφάλματος. Για να συγκλίνει ένα τέτοιο σύστημα σε αποδεκτή λύση, εφόσον διαθέτει επαρκή αριθμό κρυφών κόμβων, πρέπει να χρησιμοποιηθεί χαμηλός ρυθμός μάθησης, δηλαδή κάθε φορά η μεταβολή των βαρών να είναι ένα μικρό μέρος εκείνης που θα μηδένιζε το σφάλμα, έτσι ώστε σταδιακά τα βάρη να συμπεριλάβουν όλες τις περιπτώσεις, ομαλές και ανώμαλες.

Για να μπορέσει το δίκτυο να μάθει αποτελεσματικά όχι μόνο τις ομαλές αλλά και τις ανώμαλες περιπτώσεις, είναι σημαντικό αυτές να παρουσιαστούν με την αυξημένη συχνότητα με την οποία εμφανίζονται και στους ομιλητές, ιδιαίτερα για τα ρήματα που αλλάζουν εντελώς φωνολογική μορφή (όπως το go-went). Διαφορετικά το δίκτυο δεν μπορεί να μάθει να υπερβαίνει την τάση του να αναπαράγει πιστά το θέμα εισροής στην εκροή, διότι υπερισχύει η εκπαίδευση από τα (πολυπληθέστερα) ομαλά ρήματα. Αυτό σημαίνει ότι στον προσδιορισμό του συνόλου εκπαίδευσης, τα ζεύγη εισροής-εκροής που αντιστοιχούν

σε κάθε ρήμα πρέπει να επαναλαμβάνονται τόσες φορές όσες αντιστοιχούν στη συχνότητα εμφάνισης του συγκεκριμένου ρήματος.

Σε ένα τέτοιο σύστημα τα φαινόμενα υπεργενίκευσης εμφανίζονται επειδή τα ομαλά ρήματα υπερτερούν αριθμητικά. Συνεπώς το σφάλμα εκροής μειώνεται γρηγορότερα όταν τα βάρη αρχικά τροποποιούνται έτσι ώστε να παράγουν τις ομαλές καταλήξεις στην εκροή. Αφού το δίκτυο μάθει να αναπαράγει το θέμα εισροής και να παράγει την κατάλληλη ομαλή εκροή, μειώνοντας έτσι κατά πολύ το μέσο σφάλμα, τότε μπορούν να γίνουν και οι περαιτέρω απαραίτητες τροποποιήσεις στα βάρη ώστε να παράγονται σωστά και όσα ανώμαλα ρήματα απομένουν. Επειδή οι τροποποιήσεις αυτές επιδρούν σε λιγότερα ρήματα, η μείωση του σφάλματος που επιφέρουν είναι συγκριτικά μικρότερη και άρα η οριστική ρύθμιση των βαρών παίρνει περισσότερο χρόνο. Οπότε, στο μεσοδιάστημα είναι πιθανό να εμφανιστούν υπεργενικεύσεις. Όταν πια το δίκτυο έχει μάθει την αντιστοιχία των θεμάτων καλά, τότε μπορούν να εμφανιστούν και λανθασμένες γενικεύσεις ανώμαλου σχηματισμού αορίστου σε περιπτώσεις που ένα νέο ρήμα είναι φωνολογικά όμοιο προς ένα σύνολο ανώμαλων ρημάτων με παρόμοια συμπεριφορά. Αυτό δικαιολογείται πλήρως από τη φωνολογική φύση της αναπαράστασης του προβλήματος.

Παράρτημα

16 Συμπληρωματική βιβλιογραφία

Εισαγωγικά στη γνωσιακή επιστήμη

- Bermúdez, J. L. (2014). *Cognitive science: An introduction to the science of the mind*, 2nd ed. Cambridge University Press. (Βατή, πλήρης και σύγχρονη γενική εισαγωγή)
- Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. Oxford University Press. (Εύπεπτη συνοπτική εισαγωγή στις θεωρητικές προσεγγίσεις)
- Thagard, P. (2005). *Mind* (2nd ed.). MIT Press. (Πολύ εισαγωγικό, κάπως επιφανειακό)

Εκλαϊκευτικά-γενικού ενδιαφέροντος

- Kahneman, D. (2011). *Thinking, fast and slow*. Penguin. (Οι δύο όψεις—και ίσως δύο μηχανισμοί—της σκέψης)
- Corkin, S. (2013). *Permanent present tense: The man with no memory, and what he taught the world*. Penguin. (Διδάγματα για τη μνήμη, τα είδη και τις ιδιότητές της, από τη μελέτη ενός ασθενή με αμνησία)
- Chabris, C. F., & Simons, D. J. (2010). *The invisible gorilla, and other ways our intuitions deceive us*. Crown (Η πειραματική διερεύνηση των περιορισμών της αντίληψης και της σκέψης μας)
- Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. Penguin. (Πώς η λειτουργία του εγκεφάλου προσαρμόζεται στις απαιτήσεις του γραπτού λόγου)
- Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. Penguin. (Πώς μπορούν να μελετηθούν πειραματικά οι εγκεφαλικοί μηχανισμοί της συνείδησης)

Συλλογές και έργα αναφοράς

- Frankish, K., & Ramsey, W. M. (Eds.) (2012). *The Cambridge handbook of cognitive science*. Cambridge University Press. (Συνοπτική και εύπεπτη περιήγηση στα πεδία της γνωσιακής επιστήμης)
- Haugeland, J. (1997). *Mind design II*. MIT Press. (Συλλογή κλασικών κειμένων θεμελίωσης της γνωσιακής επιστήμης)
- Lamberts, K., & Goldstone, R. (Eds.) (2005). *Handbook of cognition*. Sage.
- Wilson, R. A., & Keil, F. C. (1999). *The MIT encyclopedia of the cognitive sciences*. MIT Press. (Εγκυκλοπαίδεια όρων της γνωσιακής επιστήμης)

Συμβολισμός

- Anderson, J.R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press. (Σύγχρονη υβριδική προσέγγιση με σύστημα παραγωγής)

- Fodor (1975). *The language of thought*. Harvard University Press. (Κλασικό έργο για την υπολογιστική-αναπαραστασιακή προσέγγιση ως συμβολική γλώσσα του νου)
- Pinker, S., & Mehler, J. (1988). *Connections and symbols*. MIT Press. (Κλασική κριτική στο συνδετισμό από συμβολική σκοπιά)
- Pylyshyn, Z. (1984). *Computation and cognition*. MIT Press. (Κλασικό έργο αναφοράς, θεμελίωση της συμβολικής προσέγγισης)

Συνδετισμός

- Elman, J. L., et al. (1996). *Rethinking Innateness*. MIT Press. (Πολύ σημαντικό έργο, συνδυάζει συνδετιστική και δυναμική προσέγγιση στην εμφυτότητα και τη μάθηση)
- Plunkett, K., & Elman, J. L. (1997). *Exercises in rethinking innateness*. MIT Press. (Συνοδευτικό του προηγούμενου, εγχειρίδιο πρακτικών ασκήσεων με το tlearn)
- Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986). *Parallel distributed processing, Vol. 1 & 2*. MIT Press. (Δίτομο κλασικό έργο αναφοράς, θεμελίωση του σύγχρονου συνδετισμού)
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition*. MIT Press. (Συνδετιστική προσέγγιση σε αναπαραστάσεις και σημασίες)

Δυναμική προσέγγιση

- Spivey, M. (2007). *The continuity of mind*. Oxford University Press. (Σύγχρονη, κάπως ιδιόμορφη αλλά κατανοητή, εισαγωγική παρουσίαση της δυναμικής προσέγγισης)
- Port, R. F., & van Gelder, T. (1995). *Mind as motion*. MIT Press. (Συλλογή εργασιών δυναμικής προσέγγισης σε διάφορες νοητικές λειτουργίες)

Ενσώματη νόηση

- Calvo, P., & Gomila, T. (Eds.) (2008). *Handbook of cognitive science: An embodied approach*. Elsevier. (Συλλογή εργασιών ενσώματης προσέγγισης)
- Gallagher, S. (2005). *How the body shapes the mind*. Oxford University Press. (Ενσώματη προσέγγιση με έμφαση στις αναπαραστάσεις του σώματος για τη νόηση)
- Gibbs, R. W. (2005). *Embodiment and cognitive science*. Cambridge University Press. (Εισαγωγικό εγχειρίδιο για την ενσώματη προσέγγιση στις διάφορες εκφάνσεις της)
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh*. Basic books. (Μανιφέστο της ενσώματης προσέγγισης με έμφαση στην εννοιολογική μεταφορά)

Εξέλιξη, εμφυτότητα, ανάπτυξη (βιολογία)

- Alcock, J. (2005). *Animal behavior* (8th edition). Sinauer. (Δημοφιλές εισαγωγικό εγχειρίδιο εξελικτικής προσέγγισης στη συμπεριφορά)
- Carrol, S. B. (2005). *Endless forms most beautiful*. Weidenfeld & Nicolson. (Εκλαϊκευμένη παρουσίαση της evo-devo)

Futuyma, D. J. (1998). *Evolutionary biology* (3rd edition). Sinauer. (Δημοφιλές εισαγωγικό εγχειρίδιο εξελικτικής βιολογίας)

Gilbert, S. F. (2006). *Developmental biology* (8th edition). Sinauer. (Δημοφιλές εισαγωγικό εγχειρίδιο αναπτυξιακής βιολογίας)

Εξελικτική ψυχολογία

Barkow, J. H., Cosmides L., & Tooby, J. (Eds.) (1992). *The adapted mind: Evolutionary psychology and the generation of culture* Oxford University Press. (Παλαιότερη συλλογή καλών κειμένων εξελικτικής προσέγγισης στην ανθρώπινη συμπεριφορά)

Buss, D. M. (Ed.) (2005). *The handbook of evolutionary psychology*. Wiley. (Πρόσφατη και πλήρης συλλογή κειμένων εξελικτικής προσέγγισης στην ψυχολογία)

Dawkins, R. (1989). *The selfish gene* (new edition). Oxford University press. (Εξαιρετική εκλαϊκευμένη εισαγωγή στην εξελικτική και κοινωνιοβιολογική σκέψη.)

Μπεϊζιανή προσέγγιση

Oaksford, M., & Chater, N. (Eds.) (1998). *Rational models of cognition*. Oxford University Press. (Συλλογή κειμένων αναφοράς για την πιθανοκρατική-ορθολογική προσέγγιση)

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press. (Μανιφέστο της μπεϊζιανής προσέγγισης)

Γνωστική ψυχολογία

Eysenck, M. W., & Keane, M. T. (2000). *Cognitive psychology: A student's handbook* (4th ed.). Psychology Press. (Τυπικό εισαγωγικό εγχειρίδιο γνωστικής ψυχολογίας)

Levitt, D. J. (Ed.) (2002). *Foundations of cognitive psychology*. MIT Press. (Ανθολογία κειμένων βασικής γνωστικής ψυχολογίας και συναφών θεμάτων)

Reisberg, D. (2012). *Cognition: Exploring the science of the mind, 5th ed.* W. W. Norton & Co. (Σύγχρονο εγχειρίδιο γνωστικής ψυχολογίας)

Νευροεπιστήμες

Gazzaniga, M. S. (Ed.) (2000). *The new cognitive neurosciences* (2nd ed.). MIT Press. (Συλλογή κειμένων σε θέματα σύγχρονης γνωσιακής νευροεπιστήμης.)

Huettel, S. A., Song, A. W., & McCarthy, G. (2004). *Functional magnetic resonance imaging*. Sinauer. (Πλήρης εισαγωγή στη λειτουργική μαγνητική τομογραφία)

Kandel, E., Schwartz, J., & Jessell, T. (2000). *Principles of neural science* (4th edition). McGraw-Hill. (Δημοφιλές εγχειρίδιο νευροεπιστημών)

Purves et al. (Eds.) (2004). *Neuroscience* (3rd ed.). Sinauer. (Τυπικό εισαγωγικό εγχειρίδιο νευροεπιστημών)

Γλωσσολογία

- Baker, M. C. (2001). *The atoms of language*. Basic Books. (Κατανοητή εισαγωγή στη θεωρητική γλωσσολογία της τσομσκιανής σχολής)
- Fromkin, V., Rodman, R., & Hyams, N. (2003). *An introduction to language* (7th ed.). Wadsworth. (Εισαγωγικό εγχειρίδιο γενικής γλωσσολογίας)
- Pinker, S. (1994). *The language instinct*. New York: W. Morrow & Co. (Απλή και διασκεδαστική παρουσίαση των θεμάτων της γλωσσολογίας, μονόπλευρη ανάλυση)

Φιλοσοφία

- Churchland, P. M. (1988). *Matter and consciousness* (revised edition). MIT Press. (Απλή και συνοπτική εισαγωγή στα βασικά της φιλοσοφίας του νου)
- Kim, J. (1996). *Philosophy of mind*. Westview Press. (Σύγχρονο και δημοφιλές εισαγωγικό εγχειρίδιο στη φιλοσοφία του νου)
- Rosenthal, D. M. (1991) (Ed.). *The nature of mind*. Oxford University Press. (Εξαιρετική συλλογή κλασικών κειμένων της φιλοσοφίας του νου)
- Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive neuroscience*. Psychology Press. (Σύγχρονη φιλοσοφία της γνωσιακής νευροεπιστήμης)
- Thagard, P. (Ed.) (2006). *Philosophy of psychology and cognitive science*. North Holland. Συλλογή κειμένων φιλοσοφίας της ψυχολογίας και της γνωσιακής επιστήμης.

Υπολογιστικά μοντέλα

- Houghton, G. (Ed.) (2005). *Connectionist models in cognitive psychology*. Psychology Press. (Κλασικός συνδεδετισμός)
- O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press. (Νευρωνικά μοντέλα από τη φυσιολογία ως τη νόηση)
- Polk, T. A., & Seifert, C. (2002). *Cognitive modeling*. MIT Press. (Συλλογή άρθρων υπολογιστικών μοντέλων της νόησης, καλύπτει ευρύτατο φάσμα προσεγγίσεων)
- Richards, W. (1988). *Natural computation*. MIT Press. (Παλιομοδίτικος συμβολισμός)
- Sun, R. (Ed.) (2008). *Cambridge handbook of computational psychology*. Cambridge University Press. (Σύγχρονη, απαιτητική εισαγωγή στις διάφορες υπολογιστικές προσεγγίσεις)