



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ

Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

— ΙΔΡΥΘΕΝ ΤΟ 1837 —

ΦΙΛΟΣΟΦΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΜΟΥΣΙΚΩΝ ΣΠΟΥΔΩΝ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**Γνωσιακή μουσικολογία και το υπολογιστικό σύστημα
IDyOT**

Γιώργος Βασιλείου Βελισσαρίδης

Επιβλέπουσα:

Χριστίνα Αναγνωστοπούλου, Αναπληρώτρια καθηγήτρια

ΑΘΗΝΑ

Σεπτέμβριος 2021

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Γνωσιακή μουσικολογία και το υπολογιστικό σύστημα IDyOT

Γιώργος Β. Βελισσαρίδης

A.M.: 1569201600003

Τριμελής Επιτροπή:

**Χριστίνα Αναγνωστοπούλου, Αναπληρώτρια καθηγήτρια
Αρετή Ανδρεοπούλου, Επίκουρη καθηγήτρια
Αναστασία Γεωργάκη, Καθηγήτρια**

Σημείωμα του συγγραφέα

Το δοκίμιο αυτό αποτελεί πτυχιακή εργασία η οποία συντάχθηκε για το Τμήμα Μουσικών Σπουδών του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών και υποβλήθηκε προς εξέταση τον Σεπτέμβριο του 2021. Ο συγγραφέας βεβαιώνει ότι το περιεχόμενο του παρόντος έργου είναι αποτέλεσμα προσωπικής εργασίας και ότι έχει γίνει η κατάλληλη αναφορά στην εργασία τρίτων, όπου κάτι τέτοιο ήταν απαραίτητο, σύμφωνα με τους κανόνες της ακαδημαϊκής δεοντολογίας.

Οι απόψεις που παρουσιάζονται στην παρούσα εργασία εκφράζουν αποκλειστικά τον συγγραφέα και όχι την επιβλέπουσα Καθηγήτρια.

Περιεχόμενα

Εισαγωγή.....	6
Abstract.....	6
Α' Μέρος: Θεωρητική μελέτη.....	7
1. Γνωσιακή Μουσικολογία.....	7
1.1. Γενικές πληροφορίες.....	7
1.2. Ιδιαιτερότητες της μοντελοποίησης της μουσικής.....	8
1.3. Μουσική ομοιότητα και μνήμη.....	9
1.4. Η νοητική αναπαράσταση του τόνου.....	11
1.5. Μουσική και αντίληψη αλληλουχίας.....	13
1.6. Μουσική και προσδοκία.....	14
1.7. Μουσική και συναίσθημα.....	15
2. Υπολογιστική Δημιουργικότητα.....	16
2.1. Κλάδος και αντικείμενο.....	16
2.2. Ορισμοί της δημιουργικότητας.....	16
2.3. Εννοιολογικοί χώροι και δημιουργικότητα.....	19
3. Θεωρία πληροφορίας.....	21
3.1. Ιστορικό Πλαίσιο.....	21
3.2. Εφαρμογή σε διαδοχικές δομές και κατηγοριοποίηση.....	21
4. Εννοιολογικοί χώροι.....	23
5. Γνωσιακή Αρχιτεκτονική.....	24
5.1. Στόχοι και μεθοδολογία της έρευνας στον τομέα της γνωσιακής αρχιτεκτονικής.....	24
5.2. Χαρακτηριστικά γνωσιακών αρχιτεκτονικών.....	26
Β' Μέρος: Η θεωρία του συστήματος IDyOT.....	28
6. Εισαγωγικά στοιχεία και επιστημονική προέλευση της θεωρίας IDyOT.....	28
6.1. Στόχοι του IDyOT και υπολογιστική μοντελοποίηση.....	28

6.2. Προσδοκία και αναπαράσταση στο IDyOM.....	28
6.3. Global Workspace Theory.....	29
7. Οπερατική χωροδία του νου και οι γεννήτριες του IDyOT.....	30
7.1. Η αναλογία της οπερατικής χωροδίας.....	30
7.2. Οι γεννήτριες και η εξελικτική τους υποστήριξη.....	31
7.3. Πρόβλεψη, στατιστική συχνότητα και απροσδοκία.....	32
7.4. Γεννήτριες και Οικουμενικός χώρος εργασίας.....	33
7.5. Η αίσθηση της αβεβαιότητας.....	33
7.6. Το κατώφλι εισαγωγής στον Οικουμενικό χώρο εργασίας.....	34
7.7. Σύνθεση.....	34
8. Υψηλότερα χαρακτηριστικά της αρχιτεκτονικής.....	35
8.1. Αναπαράσταση και μνήμη.....	35
8.2. Φασματική φύση των εννοιολογικών χώρων.....	37
8.3. Μνημονική παγίωση.....	38
8.4. Πρόβλεψη.....	40
8.5. Γνωσιακός κύκλος.....	41
9. Γνωσιακή υποστήριξη και εφαρμογές του IDyOT.....	42
9.1. Γνωσιακά στοιχεία για την εγκυρότητα της θεωρίας.....	42
9.2. Τεμαχιοποίηση γλώσσας και μουσικής.....	42
9.3. Λεξιλογική αμφισημία.....	44
9.4. Συγχρονισμός (entrainment).....	46
9.5. Δημιουργικότητα.....	47
Γ' Μέρος: Κινητοποίηση και συναίσθημα στην αρχιτεκτονική MicroPSI και σύνδεση με το IDyOT	49
10. Η θεωρία PSI και η γνωσιακή αρχιτεκτονική MicroPSI.....	49
10.1. MicroPSI.....	49
10.2. Συναίσθημα και νόηση.....	50

10.3. Η λειτουργία της κινητοποίησης.....	52
10.4. Είδη αναγκών και γνωστικές ανάγκες.....	54
10.5. Γνωσιακός κύκλος.....	57
10.6. Διαρρύθμιση της νόησης και συναίσθημα.....	58
11. Πιθανές συνδέσεις των δύο μοντέλων.....	62
Επίλογος.....	66
Βιβλιογραφία.....	67

ΕΙΣΑΓΩΓΗ

Αρχικά, αναλύονται ζητήματα της έρευνας στη γνωσιακή μουσικολογία, όπως η μουσική ομοιότητα, η νοητικές αναπαραστάσεις της μουσικής, η αντίληψη αλληλουχιών και η μουσική προσδοκία. Γίνεται εισαγωγή στην υπολογιστική δημιουργικότητα, τη θεωρία πληροφορίας, του εννοιολογικούς χώρους και τη γνωσιακή αρχιτεκτονική. Στη συνέχεια, εκτίθεται ενδελεχώς η γνωσιακή αρχιτεκτονική IDyOT, από την επιστημονική προέλευσή της και το χαμηλότερο επίπεδο λειτουργίας των γεννητριών της, μέχρι θέματα αναπαράστασης, μνήμης, πρόβλεψης, μνημονικής παγίωσης αλλά και εφαρμογές της σε τεμαχιοποίηση, συγχρονισμό και δημιουργικότητα. Τέλος, παρουσιάζεται η γνωσιακή αρχιτεκτονική MicroPSI, που μοντελοποιεί την κινητοποίηση και το συναίσθημα και εξερευνάται ένας πιθανός τρόπος σύνδεσής της με το IDyOT.

ABSTRACT

Initially, matters of cognitive musicology research are analyzed, such as musical similarity, musical mental representations, sequence perception and musical expectancy. Computational creativity, information theory, conceptual spaces and cognitive architecture are introduced. Consequently, the IDyOT cognitive architecture is thoroughly presented, from its scientific origins and the lower level of function of its generators, up to matters of representation, memory, prediction, memory consolidation but also its applications in chunking, entrainment and creativity. Finally, the MicroPSI cognitive architecture is presented, which models motivation and affect, and a possibility of connection with IDyOT is explored.

Α΄ ΜΕΡΟΣ: ΘΕΩΡΗΤΙΚΗ ΜΕΛΕΤΗ

1. Γνωσιακή Μουσικολογία

1.1. Γενικές πληροφορίες

Ο τομέας της γνωσιακής μουσικολογίας, ή μουσικής νόησης, αποτελεί ένα τμήμα της γνωσιακής επιστήμης, και επιχειρεί να εξηγήσει τα φαινόμενα σχετικά με τη μουσικότητα, την εγγενή μουσική τάση δηλαδή των ανθρώπων και την εξελικτική τους καταγωγή (Honing, 2018), αλλά και τα γνωστικά φαινόμενα σχετικά με τη μουσική πράξη σε όλα τα επίπεδα (αντίληψη, εκτέλεση, σύνθεση). Δίνεται, δηλαδή, μεγαλύτερη έμφαση στις νοητικές διεργασίες που προκαλούνται από τη μουσική, παρά στα ίδια τα χαρακτηριστικά της μουσικής. Ο κλάδος απαρτίζεται από γνωστικούς ψυχολόγους, νευροβιολόγους, μουσικολόγους και επιστήμονες υπολογιστών, καθένας από τους οποίους προσφέρει μία διαφορετική οπτική στη μελέτη της μουσικής νόησης, οι οποίες συχνά συνδυάζονται και αλληλοσυμπληρώνονται.

Ο όρος “γνωσιακή μουσικολογία” (computational musicology) χρησιμοποιήθηκε για πρώτη φορά το 1988 από τον Otto Laske, ο οποίος πραγματεύτηκε τη μουσική νόηση κυρίως από υπολογιστική σκοπιά, αλλά και με διεπιστημονικό χαρακτήρα (Volk et al., 2011). Η επιστημονική μελέτη, ωστόσο, της μουσικής νόησης ξεκίνησε πολύ νωρίτερα, κατά τις δεκαετίες του ‘70 και ‘80, με την πρωτοποριακή μελέτη των Longuet-Higgins (στον οποίο οφείλεται ο όρος “γνωσιακή επιστήμη/cognitive science”), Steedman, Krumhansl, Deutsch, Lerdaahl & Jackendoff και Cross, μεταξύ άλλων. Κατά τις επόμενες δεκαετίες, ο κλάδος συμπεριέλαβε έρευνες στους τομείς της αναπτυξιακής ψυχολογίας, της γλωσσολογίας, της νευροεπιστήμης, της μουσικής εκπαίδευσης, της υπολογιστικής μοντελοποίησης και της πειραματικής ψυχολογίας. Παρ’ όλα αυτά, η μουσική νόηση δεν θεωρείται ακόμη από πολλούς επιστήμονες της γνωσιακής επιστήμης εξίσου σημαντικό αντικείμενο μελέτης της γνωσιακής επιστήμης, όσο η μελέτη της γλώσσας, της όρασης, του προγραμματισμού, της συλλογιστικής ή της επίλυσης προβλημάτων, αν και στη βιβλιογραφία της γνωσιακής μουσικολογίας προτείνονται πολλοί λόγοι για τους οποίους η μουσική αποτελεί θεμελιώδες χαρακτηριστικό της ανθρώπινης νόησης (M. Pearce & Rohrmeier, 2012).

Ένα ισχυρό κριτήριο, το οποίο καταδεικνύει την καθολικότητα της μουσικής στον άνθρωπο αποτελεί το γεγονός ότι τη συναντάμε σε όλους τους μουσικούς πολιτισμούς (συνήθως η έννοια της είναι συνδεδεμένη με την αυτή του χορού), όπου συνοδεύει συνήθως πολλές θρησκευτικές και κοσμικές τελετές. Αλλά υπάρχουν κι άλλοι λόγοι για τους οποίους η μουσική κατέχει μοναδική θέση στην γνωσιακή επιστήμη (Wiggins, 2020):

- Εφήμερη: η μουσική δεν υπάρχει ως φυσικό αντικείμενο στον κόσμο (η αναπαράσταση της μουσικής σε μουσική σημειογραφία ή σε ηχογράφηση δεν είναι η ίδια η μουσική), επομένως εξαρτάται ολοκληρωτικά στη λειτουργία της μνήμης για τη γνωσιακή της επίδραση.
- Ανεπιστημονική (apéristemic): η μουσική δεν μπορεί να εκφράσει προτάσεις με Μπουλιανές ιδιότητες, δηλαδή αδυνατεί να δημιουργήσει αληθείς ή ψευδείς δηλώσεις. Εξάιρεση αποτελούν το φαινόμενο της ονοματοποιίας και της σύνδεσης της μουσικής με Μπουλιανά αντικείμενα, που αποτελούν ωστόσο πολύ σπάνιες και ειδικές περιπτώσεις.
- Αυτοαναφορική: η μουσική, αντίθετα με τη γλώσσα, μπορεί να αναφερθεί μόνο σε μουσικά ερεθίσματα, συνήθως εντός του ίδιου κομματιού (και πάλι με εξάιρεση τις παραπάνω ασυνήθιστες περιπτώσεις).
- Πολιτισμική: η μουσική, όπως και η γλώσσα, αποτελεί πολιτισμικό αντικείμενο και προϋποθέτει προσπολιτισμό (πολιτισμική εκμάθηση) προκειμένου να γίνει αντιληπτή, και έτσι εξαρτάται σε μεγάλο βαθμό από την μάθηση.
- Μαγευτική: η μουσική έχει ισχυρή συνειδητή και μη συνειδητή επίδραση στην ανθρώπινη νόηση και προκαλεί κινητικές και συναισθηματικές αντιδράσεις, σε τέτοιο βαθμό που οι αντιδράσεις αυτές να περιγράφονται ως “ακούσιες”, και να ωθούν άτομα να σπαταλούν μεγάλο μέρος των χρημάτων τους σε μουσική.

Η φυσική γλώσσα μοιράζεται κάποια από τα παραπάνω χαρακτηριστικά, αλλά φέρει το βάρος της αναφοράς στον πραγματικό κόσμο και του κοινωνικού νοήματος, που κάνουν τη μελέτη της ιδιαίτερα δύσκολη στην αναγωγική επιστήμη. Η μουσική, αντίθετα, αποτελεί ένα “κλειστό σύστημα”, το οποίο μπορεί να μελετηθεί πλήρως με βάση την αντίληψη, αναπαράσταση, νόηση και μνήμη που σχετίζονται με αυτή, και έτσι προσφέρεται για άμεση εμπειρική μελέτη και σύνδεση με τη γλώσσα, μέσω των μηχανισμών και πόρων που πιθανώς μοιράζονται.

1.2. Ιδιαιτερότητες της μοντελοποίησης της μουσικής

Ο τομέας της υπολογιστικής μουσικολογίας, που ξεκίνησε ήδη από τη δεκαετία του '50, με τις πρώτες μουσικές συνθέσεις από ηλεκτρονικό υπολογιστή (Illiac suite, 1957 – Hiller & Issacson), δραστηριοποιείται σήμερα στον τομέα του Music Information Retrieval (MIR – Ανάκτηση Μουσικής Πληροφορίας), στη μουσική θεωρία και ανάλυση, την εθνομουσικολογία, την έρευνα της μουσικής εκτέλεσης, αλλά και στον τομέα της γνωσιακής επιστήμης. Στην υπολογιστική προσέγγιση της νόησης, έχουν προταθεί τις τελευταίες τρεις δεκαετίες πολλά μοντέλα, τα οποία αφορούν συγκεκριμένες μουσικές διεργασίες (εύρεση τονικότητας, μέτρου), τη μουσική αντίληψη (αναγνώριση τόνου, ομαδοποίηση και τεμαχιοποίηση, αρμονική ανάλυση,

ανακάλυψη μοτίβων, τονική αφαίρεση) αλλά και πιο γενικά, ενοποιητικά μοντέλα της μουσικής εμπειρίας, εκτέλεσης και σύνθεσης (Temperley, 2013).

Η πληθώρα, ωστόσο, μοντέλων έχει φέρει στο προσκήνιο και το ζήτημα της σύγκρισης μεταξύ τους, και της επιλογής αυτών που αναπαριστούν καλύτερα τη γνωστική διεργασία ή διεργασίες που μοντελοποιούν, σε σχέση με άλλα μοντέλα. Ο Honing σημειώνει ότι ο συνηθισμένος έλεγχος καλής προσαρμογής (goodness-of-fit test) συχνά δεν αρκεί για την σύγκριση μεταξύ μοντέλων, και χρειάζεται να ληφθούν υπ' όψη και άλλοι παράγοντες, όπως η απλότητα του κάθε μοντέλου (σύμφωνα με το ξυράφι του Occam) και το κατά πόσο παράγει αναπάντεχα αποτελέσματα σε σχέση με προηγούμενες δοκιμές, τα οποία ωστόσο εντάσσονται εντός των προβλεπόμενων αποτελεσμάτων (Honing, 2006). Την πρόταση αυτή έχουν ενστερνιστεί και άλλοι υπολογιστικοί γνωσιακοί επιστήμονες, μεταξύ των οποίων και οι Pearce και Wiggins, όπως θα δούμε παρακάτω.

Είδαμε ήδη στο εισαγωγικό κομμάτι πως η μουσική ως τομέας μελέτης είναι απλούστερος από τη γλώσσα, καθώς δεν περιέχει το σημασιολογικό της βάρος. Από την άλλη μεριά όμως, η μουσική είναι εγγενώς και αναγκαστικά πιο περίπλοκη από γλώσσα με την εξής έννοια: Η γλωσσική επικοινωνία περιλαμβάνει πολλές αντιληπτικές διαστάσεις (φώνημα, τονικό ύψος, ακουστότητα, χρονισμός κτλ.), ωστόσο η διάσταση του φωνήματος, και οι ιεραρχικές δομές που χτίζονται με βάση αυτή είναι τόσο κυρίαρχες, που καθιστούν εφικτή τη μελέτη της γλώσσας μόνο μέσω αυτών, γεγονός που καθρεφτίζεται και στον ίδιο το γραπτό λόγο. Το γεγονός αυτό έχει οδηγήσει στη χρήση περιορισμένων, μονοδιάστατων Μαρκοβιανών μοντέλων στην υπολογιστική γνωσιακή μελέτη της γλωσσολογίας. Στη μουσική, αντίθετα, αυτού του είδους η μονοδιάστατη προσέγγιση δεν μπορεί να λειτουργήσει, λόγω της θεμελιώδους φύσης της, δηλαδή της ταυτόχρονης ανάπτυξης και αλληλεπίδρασης πολλών αντιληπτικών διαστάσεων, που καθιστά απαραίτητη τη δημιουργία ειδικών τεχνικών αξιολόγησης ώστε να αποφευχθεί η ασάφεια κατά τη μελέτη της.

Η πολυδιάστατη φύση της μουσικής έχει ωθήσει την έρευνα στην αναζήτηση βελτιωμένων τεχνικών μοντελοποίησης, με ένα χαρακτηριστικό παράδειγμα την σημαντική εργασία των Conklin & Witten. Η εργασία αυτή συνίσταται στην επέκταση των μονοδιάστατων μοντέλων μεταβλητής τάξης που είχε χρησιμοποιήσει ο Witten σε προηγούμενη γλωσσολογική έρευνα, σε πολυδιάστατα μέσα πρόβλεψης (1995). Μία δεκαετία αργότερα, ο Marcus Pearce έδειξε ότι η μέθοδος των Conklin & Witten, ενταγμένη στο γνωσιακό πλαίσιο του συστήματος IDyOM (Information Dynamics of Music), μοντελοποιεί την ανθρώπινη αντίληψη μελωδίας με εξαιρετική ακρίβεια, ενώ έχει χρησιμοποιηθεί και στη γλωσσολογία, με ενδιαφέροντα αποτελέσματα (2005). Τα Μαρκοβιανά μοντέλα, στις διάφορες παραλλαγές τους που έχουν χρησιμοποιηθεί στην έρευνα, έχουν βρεθεί αποτελεσματικά σε ποικίλες μουσικές εργασίες τις τελευταίες δεκαετίες. Βρίσκονται στην απέναντι όχθη από τις προσεγγίσεις στη γνωσιακή μουσικολογία που εμπνέονται από την συντακτική γλωσσολογία του Chomsky, οι οποίες

θέτουν κεντρική σημασία στον κατωφερή έλεγχο των συντακτικών δομών, σε αντίθεση με τις ανωφερείς στατιστικές κανονικότητες που αποτυπώνουν τα Μαρκοβιανά μοντέλα.

1.3. Μουσική ομοιότητα και μνήμη

Η εφήμερη φύση της μουσικής ενισχύει το σειριακό της ποιόν, γεγονός που έχει επιπτώσεις και για τη μνήμη της μουσικής. Πιο συγκεκριμένα, η σύγκριση μεταξύ μουσικών αντικειμένων, δηλαδή η αντίληψη της μουσικής ομοιότητας, είναι αναγκαστικά διαδοχική, καθώς η ταυτόχρονη ακρόαση των μουσικών ερεθισμάτων είναι κακόφωνη και ακατανόητη. Επομένως, ο μόνος τρόπος ένας άνθρωπος να συγκρίνει δύο κομμάτια μουσικής είναι να συγκρίνει το καθένα με τη μνήμη του άλλου. Αυτό τίθεται σε άμεση αντιδιαστολή με τα οπτικά ερεθίσματα, στα οποία είναι εφικτό να συγκριθεί άμεσα κάθε χαρακτηριστικό, με την γρήγορη εναλλαγή προσοχής μεταξύ των ερεθισμάτων.

Ως άμεση συνέπεια της εξ' ορισμού επιβεβλημένης σειριακής φύσης της μουσικής, έχουμε το φαινόμενο της προέγερσης (priming effect), κατά το οποίο ένα πρώτο ερέθισμα το οποίο γίνεται αντικείμενο αντιληπτικής και νοητικής επεξεργασίας επηρεάζει την αντιληπτική και νοητική επεξεργασία ενός δεύτερου ερεθίσματος, το οποίο παρουσιάζεται αργότερα, οδηγώντας έτσι αναπόφευκτα σε ασύμμετρες κρίσεις ομοιότητας μεταξύ τους (Tversky, 1977). Μία άλλη συνέπεια της διαδοχικότητας της μουσικής είναι πως η μελέτη της μουσικής ομοιότητας συνιστά μελέτη της μουσικής μνήμης. Αν λάβουμε υπ' όψη και την ανεπιστημονικότητα και αυτοαναφορικότητα της μουσικής, τότε μπορούμε λογικά να υποθέσουμε πως η μουσική ομοιότητα ταυτίζεται με την δομική ομοιότητα στη μνήμη, εφόσον δεν υπάρχει αναφορά σε κάτι εξωτερικό (μεμονωμένες εξωμουσικές συνδέσεις είναι στατιστικά ασήμαντες). Επομένως, είναι πιθανό η μουσική να προσφέρεται για τη μελέτη της μνήμης μέσω της μουσικής ομοιότητας, περισσότερο απ' ότι άλλοι τομείς.

Η αναφορά στη μουσική, η ομοιότητα δηλαδή σε αντιληπτικό και γνωστικό επίπεδο, εμφανίζεται τόσο στο επίπεδο μίας μουσικής φράσης, ενός τμήματος ή και ενός ολόκληρου κομματιού. Στη σχετική έρευνα έχει μελετηθεί η μουσική ομοιότητα σε πολλαπλά επίπεδα, συχνά για εμπορικούς σκοπούς. Ωστόσο, η μηχανική προσέγγιση στο ζήτημα οδηγεί σε πολλές μελέτες στην παράλειψη του γνωσιακού στοιχείου του προβλήματος, την αναγωγή του σε πρόβλημα επεξεργασίας ακουστικού σήματος, και την παραμέληση του προβλεπτικού χαρακτήρα της μουσικής αντίληψης και νόησης. Η προσέγγιση αυτή οδηγεί συχνά στον προβληματισμό σχετικά με το λεγόμενο "σημειολογικό κενό" ("semantic gap") μεταξύ υπολογιστικής ανάλυσης και νόησης, που ωστόσο πιθανώς αντιμετωπίζεται με την μελέτη της μουσικής ομοιότητας ως συνάρτηση της μνήμης (Wiggins, 2007).

Η θεωρία "Cue Abstraction" ("Αφαίρεσης Σημάτων") της Deliège (Deliège & Mélen, 2004), που συνδέεται στενά με την μουσικολογική σημειωτική ανάλυση του Ruwet (Ruwet & Everist, 1987), πραγματεύεται τη σχέση μεταξύ νόησης και μουσικής δομής με άμεσο τρόπο,

υποστηρίζοντας πως μία θεμελιώδης διαδικασία ακρόασης είναι η αναγνώριση σημάτων (cues), τα οποία μπορεί να αποτελούνται από μουσικά θέματα, φράσεις ή ακόμη και σύντομα μοτίβα. Ο ακροατής δημιουργεί νοητικές αναπαραστάσεις των σημάτων αυτών κατά την ακρόαση στις διάφορες επανεμφανίσεις και παραλλαγές τους. Οι αναπαραστάσεις εξελίσσονται παράλληλα με τα σήματα, ορίζοντας τι αναγνωρίζεται ως μέλος μίας κατηγορίας και τι όχι, πράγμα που γίνεται εφικτό μέσω της επίδρασης του φαινομένου της προέγερσης και της δυναμικής φύσης των αναπαραστάσεων.

Στην έρευνα της μουσικής νόησης, η μοντελοποίηση των μουσικών αντιλημμάτων και δομών χρησιμοποιώντας γεωμετρικούς χώρους χαμηλών διαστάσεων έχει συνεισφέρει στην κατανόηση της αντιληπτικής ομοιότητας. Υπάρχει πληθώρα τέτοιων μοντέλων, κυρίως για το αντίλημμα του τόνου, και κατ' επέκταση της τονικής αρμονίας, αλλά και για το ρυθμό. Οι μαθηματικοί χώροι αυτοί έχουν τυποποιηθεί από τον Gärdenfors, και περιλαμβάνουν διακριτές περιοχές, στις οποίες ανατίθενται συμβολικά ονόματα, όπως για παράδειγμα στους γνωστούς χώρους των χρωμάτων και φωνηέντων (Gärdenfors, 2000). Σύμφωνα με τη θεωρία, αν υπάρχει τουλάχιστον ένα ορισμένο σημείο στον χώρο, είναι δυνατή η σύγκριση άλλων σημείων με αυτό. Με βάση ένα σύνολο βασικών γεωμετρικών ορισμών, μπορούμε να διακρίνουμε μεταξύ δύο αναπαραστάσεων στον χώρο: η μία είναι χονδρική και διακριτή ενώ η άλλη λεπτομερής και συνεχής. Στη μουσική, ο τόνος προσφέρεται ιδιαιτέρως για διευκρίνηση των παραπάνω εννοιών. Πιο συγκεκριμένα, τα δυτικά πληκτροφόρα όργανα είναι συνήθως κουρδισμένα σε ισοσυγκερασμό, δηλαδή κάθε νότα απέχει ίση απόσταση από τις γειτονικές της στην λογαριθμική κλίμακα των μουσικών διαστημάτων, και συγκεκριμένα απόσταση ημιτονίου (ορίζεται ως συντελεστής συχνοτήτων και ισούται με τη δωδέκατη ρίζα του δύο). Αυτό το κούρδισμα δεν χρησιμοποιείται από ικανούς μουσικούς οργάνων με δυνατότητα μικροεπέμβασης στον τόνο, καθώς προτιμώνται πιο σύμφωνα διαστήματα σε σχέση με την τονική του κάθε κομματιού που έχει ως αποτέλεσμα η απόσταση μεταξύ γειτονικών νοτών να μην είναι πάντα ίδια. Με αυτή την έννοια, ένα πιάνο είναι φάλτσο, αλλά για τους περισσότερους ακροατές το γεγονός αυτό περνάει απαρατήρητο, παρόλο που η διαφορά είναι εύκολα αντιληπτή όταν τα κούρδισματα συγκρίνονται άμεσα. Η εξήγηση για το φαινόμενο αυτό αποδίδεται στην κατηγορική αντίληψη, που προσδίδει στον ακροατή που την διαθέτει τα εξής πλεονεκτήματα: επιτρέπει την αναγνώριση ακόμα και προσεγγίσεων γνωστών ερεθισμάτων και οι διακριτές αναπαραστάσεις της είναι ευκολότερο να αποθηκευτούν και ανακληθούν από συνεχείς αναπαραστάσεις.

1.4. Η νοητική αναπαράσταση του τόνου

Ένα ενδιαφέρον κομμάτι της έρευνας στη μουσική μνήμη αποτελεί η διαχρονική ανάπτυξη της τονικής αντίληψης σε μεμονωμένα άτομα. Συγκεκριμένα, στοιχεία δείχνουν ότι βρέφη στον δυτικό κόσμο κατέχουν τους πρώτους μήνες ζωής τους απόλυτη ακοή (perfect pitch), στην οποία χάνουν μετέπειτα συνειδητή πρόσβαση, και η τονική τους αντίληψη γίνεται σχετική,

δηλαδή ορίζεται μεταξύ διαδοχικών τόνων. Μία πιθανή εξελικτική εξήγηση του φαινομένου είναι η ανάγκη κατανόησης φωνητικών σχημάτων που χρησιμοποιούνται και από τους δύο γονείς και τα αδέρφια ενός βρέφους, τα οποία εκφράζονται σε διαφορετικά τονικά εύρη. Ωστόσο, μερικοί άνθρωποι διατηρούν συνειδητή πρόσβαση στην απόλυτη ακοή όλη τους τη ζωή, ενώ υπάρχουν στοιχεία που δείχνουν ότι ενήλικες χωρίς απόλυτο αυτί διατηρούν απόλυτη μνήμη άδηλα.

Το αντίλημμα του μουσικού τόνου μπορεί να αναπαρασταθεί ως μία ευθεία γραμμή, το αριστερό άκρο της οποίας ανταποκρίνεται στα 44Hz και το δεξί στα 22kHz, όσο περίπου και το ακουστικό εύρος των ανθρώπων. Η κλίμακα αυτή είναι λογαριθμική, και μας δίνει την ιδέα του “τονικού ύψους”, με βάση τη θέση ενός τόνου στην γραμμή. Συνολικά, η κλίμακα περιλαμβάνει περίπου 110 δυτικά ισοσυγκερασμένα ημιτόνια, τα οποία απαιτούν τουλάχιστον επτά bits για να αποθηκευτούν. Υπάρχουν, ωστόσο, και άλλοι τρόποι να συγκρίνουμε τόνους. Ένας ιδιαίτερα θεμελιώδης σχετίζεται με τη φυσιολογία του αυτιού και την αρμονική στήλη. Πιο συγκεκριμένα, το αντίλημμα της οκτάβας εμφανίζεται όταν ο λόγος μεταξύ των συχνοτήτων δύο τόνων είναι ακέραια δύναμη του 2, δηλαδή η θεμελιώδης συχνότητα του ενός τόνου ταυτίζεται με τον δεύτερο αρμονικό του άλλου, ή αν η απόσταση μεταξύ τους είναι μεγαλύτερη, η ίδια σχέση εφαρμόζεται αναδρομικά. Επομένως, υπάρχει μία κυκλική δομή εντός της αρμονικής στήλης, την οποία τα αυτιά των ανθρώπων έχουν εξελιχθεί να ακούνε. Η δομή αυτή μπορεί να αναπαρασταθεί ως έλικας σε τρισδιάστατο χώρο, όπου ο κατακόρυφος άξονας εκφράζει την απόσταση οκτάβων μεταξύ τόνων (Shepard, 1982). Η αναπαράσταση αυτή δηλώνει πως η ομοιότητα μεταξύ νοτών που απέχουν οκτάβα είναι μεγαλύτερη από την ομοιότητα γειτονικών νοτών (12 ημιτονίων στη δυτική μουσική), γεγονός που ανταποκρίνεται στην αντιληπτική εμπειρία των δυτικών μουσικών, αν και δεν έχει αποκρυσταλλωθεί ακόμη σε τι βαθμό το αντίλημμα της οκτάβας είναι μαθημένο και σε τι βαθμό εγγενές. Μάλιστα, πρόσφατες έρευνες σε γεωγραφικά απομονωμένο πολιτισμό, χωρίς επαφή με τη δύση, έχουν δείξει ότι η ισοδυναμία της οκτάβας δεν είναι εξέχων χαρακτηριστικό της μουσικής αντίληψης των κατοίκων του πολιτισμού αυτού (Jacoby et al., 2019). Το ζήτημα του κατά πόσο νοητικές δομές όπως γραμμές, έλικες και δένδρα είναι εγγενείς, και καθεμία χρησιμοποιείται για την αναπαράσταση κατάλληλων εννοιών έχει μελετηθεί από τους Kemp & Tenenbaum, και η αντίληψη της οκτάβας ταιριάζει με τα ευρήματά τους (Kemp & Tenenbaum, 2008).

Εξαιτίας της λογαριθμικής κατασκευής της, η γραμμή των τόνων, όταν χρησιμοποιείται για τη δημιουργία κλιμάκων, εμφανίζει ασυμμετρίες. Έχοντας ως δεδομένες τις 12 νότες του ισοσυγκερασμένου συστήματος μπορούμε να ορίσουμε κλίμακες με βάση το υποσύνολο της χρωματικής κλίμακας που τις συνιστά. Η ιστορικά δημοφιλέστερη κλίμακα στη δυτική μουσική είναι η μείζονα κλίμακα, πιθανώς λόγω της καλής αντιστοιχίας της με την αρμονική στήλη. Η μείζονα κλίμακα μπορεί κατασκευαστεί ξεκινώντας από οποιαδήποτε νότα, λαμβάνοντας το εξής υποσύνολο της χρωματικής κλίμακας: {0, 2, 4, 5, 7, 9, 11}.

Τόσο η δομή των διαφόρων κλιμάκων, ορισμένες κατ' αυτό τον τρόπο, όσο και ο τονικός έλικας και η σχετική ακοή προσφέρουν αποδοτικότητα στις αναπαραστάσεις του τόνου. Πιο αναλυτικά, η σχετική ακοή μας επιτρέπει να ακούσουμε κάθε νότα ενός κομματιού σε σχέση με την τονική νότα του, ενώ η ασυμμετρία των κλιμάκων μας επιτρέπει να αναγνωρίσουμε την κάθε νότα με βάση την μετατόπισή της από την τονική. Επιπλέον, έχει υποστηριχθεί ότι για τον εντοπισμό δομών σε μία αντιληπτική διάσταση είναι απαραίτητος ένας μικρός αριθμός διακριτών αντιληπτικών κατηγοριών και ένας μικρός αριθμός στοιχείων προς απομνημόνευση (McAdams, 1987). Η χρήση ισοδυναμίας μεταξύ οκτάβων (octave equivalence) οδηγεί σε αναπαραστάσεις δύο διαστάσεων των επτά κατηγοριών η καθεμία, σε αντίθεση με τη χρήση μόνο της τονικής γραμμής, που οδηγεί σε αναπαραστάσεις μίας διάστασης των 110 κατηγοριών. Ίσως έτσι εξηγείται το γεγονός ότι η συντριπτική πλειονότητα των τονικών συστημάτων ανά τον κόσμο που έχουν μελετήσει οι μουσικολόγοι χρησιμοποιούν ισοδυναμία μεταξύ οκτάβων.

Στη πράξη, τα παραπάνω φαινόμενα εμφανίζονται κατά την ανάπτυξη των ικανοτήτων ενός μουσικού. Για παράδειγμα, ένας αρχάριος μουσικός που παίζει ένα όργανο που απαιτεί κούρδισμα της κάθε νότας μαθαίνει γρήγορα να ακούει την θέση της κάθε νότας σε σχέση με την τονική ως ποιότητα του ήχου (δηλαδή qualia της κάθε βαθμίδας). Σε μετέπειτα στάδιο της μουσικής εκπαίδευσης, είναι σε θέση να φανταστεί μία τονική και έτσι να ακούσει επόμενες νότες με βάση την τονική αυτή, ικανότητα που μειώνει σε μεγάλο βαθμό το μέγεθος των τονικών αναπαραστάσεων. Τέλος, ένας προχωρημένος μουσικός είναι σε θέση να συνδυάσει το διάστημα και τη βαθμίδα και έτσι μπορεί να διακρίνει μεταξύ εμφανίσεων ενός συγκεκριμένου διαστήματος ανάλογα με τη βαθμίδα από την οποία ξεκινά, παρότι σε ισοσυγκερασμένο κούρδισμα δεν υπάρχει φυσική διαφορά μεταξύ τους, αν η τονική δεν είναι ακουστή.

Οι αρχές που περιγράψαμε για την αναπαράσταση μεμονωμένων τονικών υψών εφαρμόζονται και για την αναπαράσταση της μουσικής αρμονίας, στην οποία μελετάμε ομάδες ταυτόχρονων τονικών υψών. Παρόλο που η σχέση μεταξύ συγχορδιών και των νοτών που τις συνιστούν είναι ξεκάθαρη (και είναι δυνατόν κανείς να εκπαιδευτεί στο να ξεχωρίζει μεμονωμένες νότες μίας συγχορδίας), και οι ίδιες οι συγχορδίες γίνονται αντιληπτές ως “αντιληπτικά άτομα”. Έτσι, είναι δυνατό να χρησιμοποιήσουμε αντίστοιχες αναπαραστάσεις με αυτές που χρησιμοποιήσαμε για την αναπαράσταση του τόνου, όπως έλικες σε τρισδιάστατο χώρο (Chew, 2014) ή την επιφάνεια ενός τόρου (Lerdahl, 1988), διατηρώντας τις ιδιότητες της επαυξημένης αποδοτικότητας της αναπαράστασης και της αντιληπτικής ισοδυναμίας, που στην περίπτωση του τόνου εντοπίσαμε στο αντίλημμα της οκτάβας και στην περίπτωση της αρμονίας εντοπίζουμε στις μετατροπίες. Αυτός ο χώρος πρέπει επίσης να μαθευτεί από έναν μουσικό, επομένως οι σταθερές εκφάνσεις ενός ερεθίσματος παίζουν καθοριστικό ρόλο στη δημιουργία των νοητικών αναπαραστάσεων: η γνωσιακή διαδικασία που ακολουθείται φαίνεται να συνίσταται σε αναζήτηση για τον βέλτιστο τρόπο να αναπαρασταθεί η πληροφορία, βάσει της δομής της.

1.5. Μουσική και αντίληψη αλληλουχίας

Η εφήμερη φύση της μουσικής, εκτός από την επίδρασή της στη μνήμη, την οποία είδαμε ήδη, σχετίζεται επίσης στενά με την ικανότητα της τεμαχιοποίησης (chunking) στο χρόνο. Τα ιεραρχικά μοντέλα που προκύπτουν από τη διαδικασία της τεμαχιοποίησης εμφανίζονται ακόμη και σε απλή μουσική, ενώ έχουν χρησιμοποιηθεί σε τεχνικές μουσικής ανάλυσης του τελευταίου αιώνα. Ακόμη, είναι εφικτό μία μουσική ανάλυση να αναπαρασταθεί ως δένδρο, σε αναλογία με τα συντακτικά δένδρα της γλωσσικής ανάλυσης, με διαφορά ωστόσο την έλλειψη συντακτικών κατηγοριών (Lerdahl & Jackendoff, 1983). Έχει παρατηρηθεί σε έρευνες των τελευταίων δεκαετιών, πως οι ίδιες ανωφερείς στατιστικές μέθοδοι προβλέπουν επιτυχώς την ανθρώπινη κρίση τόσο για τεμαχιοποίηση της μουσικής όσο και για τεμαχιοποίηση της γλώσσας (Goldwater et al., 2006; M. T. Pearce, Müllensiefen, et al., 2010). Μπορούμε έτσι να συμπεράνουμε πως οι δύο τομείς πιθανώς συνδέονται μέσω ενός κοινού μηχανισμού, ενώ για την τεμαχιοποίηση μπορούμε να υποθέσουμε ότι πρόκειται για γενικότερη ιδιότητα, αφού εμφανίζεται και στη μουσική και τη γλώσσα.

Τα σήματα της θεωρίας αφαίρεσης σημάτων της Deliège αποτελούν συνήθως αλληλουχίες νοτών (1987) και επομένως είναι λογικό να υποθέσουμε ότι αυτές οι αλληλουχίες πρέπει να εμπεριέχονται σε αντιληπτικά τεμάχια, προκειμένου να διευκολύνουν την ανάκληση. Το γεγονός αυτό υποδηλώνει πως η αναπαράσταση των τεμαχίων δουλεύει με τέτοιο τρόπο, ώστε να διευκολύνει τις κρίσεις τοπικής ομοιότητας. Ο τόνος και χρόνος για μεμονωμένες νότες δύναται να αναπαρασταθεί ευθέως, χρησιμοποιώντας τις γεωμετρικές προσεγγίσεις που παρατέθηκαν παραπάνω αλλά οι αλληλουχίες είναι πιο προβληματικές. Μπορούν για παράδειγμα να αναπαρασταθούν ως μαθηματικές ακολουθίες ή ως τροχιές σε έναν χώρο τόνου και χρόνου. Ωστόσο, με οποιονδήποτε τρόπο και αν αναπαρασταθούν, οι αλληλουχίες πρέπει να είναι προσβάσιμες ιεραρχικά και τα μεμονωμένα στοιχεία της ιεραρχίας προσβάσιμα ως μονάδες, καθώς αυτού του είδους οι χειρισμοί εκτελούνται συνεχώς από μουσικούς, οι οποίοι με ευκολία εστιάζουν σε όλα τα μουσικά επίπεδα, όταν συζητούν και δοκιμάζουν μουσική. Ακόμη, είναι σημαντικό να μπορεί να οριστεί ο βαθμός ομοιότητας μεταξύ αλληλουχιών, και αν ο βαθμός αυτός είναι μεγάλος να μπορεί να οδηγήσει σε ενοποίηση τους.

1.6. Μουσική και προσδοκία

Έχει υποστηριχτεί από τη δεκαετία του '50 στη μελέτη της μουσικής, πως η προσδοκία παίζει σημαντικό ρόλο στη μουσική εμπειρία (Meyer, 1956). Ο Huron υποστηρίζει την ίδια άποψη και την επεκτείνει στην ψυχολογική του θεωρία για τη μουσική (2006), ενώ η εργασία του Pearce προσφέρει εμπειρική επιβεβαίωση στην υπόθεση. Στην εργασία του ο Huron, μάλιστα, επιχειρηματολογεί πειστικά πως η αίσθηση αβεβαιότητας ως προς το τι θα ακολουθήσει, η οποία μεταβάλλεται συνεχώς κατά την ακρόαση, ευθύνεται για μεγάλο μέρος της λεπτομερούς αισθητικής εμπειρίας της μουσικής. Η αύξηση και μείωση της μουσικής τάσης, που οφείλεται

στις αυξομειώσεις της αβεβαιότητας των προσδοκιών, και οδηγεί τελικά στην επιβεβαίωση ή διάψευση τους, συμβάλλει στην αντίληψη της ροής της μουσικής.

Ωστόσο, εκτός από το περιεχόμενο των προσδοκιών και την έκβασή τους, μεγάλη σημασία για την αίσθηση της μουσικής έχει και ο συγχρονισμός των προσδοκιών, που επιτρέπει σε ομάδες ανθρώπων να εκτελούν μουσική ταυτόχρονα, με εντυπωσιακή ακρίβεια, φαινόμενο που έχει παρατηρηθεί μόνο σε ανθρώπους. Η ικανότητα αυτή αντίληψης και προσδοκίας ενός παλμού είναι γνωστός ως *συγχρονισμός (entrainment)*¹, και για μερικούς ανθρώπους αποτελεί τόσο ισχυρή λειτουργία που συχνά το βρίσκουν δύσκολο να μην συγχρονιστούν, όταν ακούν μουσική που βρίσκουν ελκυστική (Fitch, 2013). Το φαινόμενο αυτό έχει μελετηθεί εκτενώς στον τομέα της γνωσιακής μουσικολογίας (Grahn, 2012), και έχει διαπιστωθεί πως, παρόλο που άλλα ζωικά είδη με ικανότητες φωνητικής μίμησης παρουσιάζουν προσωρινό συγχρονισμό σε μουσική όταν ενθαρρύνονται να το κάνουν είτε συγχρονίζονται αντανακλαστικά, ο αυθόρμητος και παρατεταμένος συγχρονισμός από προσωπική βούληση είναι αποκλειστικά ανθρώπινο χαρακτηριστικό.

Το σύστημα IDyOM μοντελοποιεί το περιεχόμενο των ανθρώπινων προσδοκιών με μεγάλη ακρίβεια, και λαμβάνει υπ' όψη τον χρονισμό των ακολουθιών που δέχεται για τη δημιουργία των προσδοκιών του, ωστόσο δεν δύναται να επεξεργαστεί μουσική σε πραγματικό χρόνο. Ο ρυθμικός συγχρονισμός είναι, όμως, μία από τις βασικότερες ικανότητες που απαιτούνται στη μουσική εκτέλεση και σχετίζεται άμεσα με την κινητική προετοιμασία, ώστε το αποτέλεσμα να συντονισμένο στο χρόνο. Ο συγχρονισμός μπορεί, μάλιστα, να γίνει εξαιρετικά περίπλοκος, περιλαμβάνοντας μη κανονικές ρυθμικές δομές με παρατεταμένους κύκλους (ινδική κλασική ή ελληνική παραδοσιακή μουσική), ταυτόχρονα διακριτά επίπεδα ρυθμικής αγωγής (πολυρυθμίες) και λεπτές παρεκκλίσεις από τον παλμό, που παίζουν ωστόσο σημαντικό μουσικό ρόλο (funk, jazz). Συμπεραίνουμε, έτσι, πως με εκπαίδευση, η ικανότητα συγχρονισμού των ανθρώπων μπορεί να φτάσει υψηλά επίπεδα πολυπλοκότητας.

Όπως αναφέρθηκε παραπάνω, η ικανότητα συγχρονισμού φαίνεται να συσχετίζεται με την ικανότητα φωνητικής μίμησης στα ζώα (Schachner et al., 2009). Είναι λοιπόν πιθανό, ο συγχρονισμός να σχετίζεται με τη φωνητική μίμηση. Ο τομέας έρευνας της γνωσιακής επιστήμης που ασχολείται με την ανάπτυξη προσφέρει πολλά στοιχεία για την εμπλοκή της φωνητικής μίμησης στη διαδικασία εκμάθησης της ομιλίας, και απαιτεί κάποιου είδους αντίληψη ομιλίας (ακόμη και χωρίς σημασιολογικές συνδέσεις). Ενδέχεται η ικανότητα του συγχρονισμού να συνδέεται με όλες αυτές τις διαδικασίες, για λόγους γνωσιακής αποδοτικότητας, ως εξής: Η εστιασμένη προσοχή στον λόγο, όπως σε οτιδήποτε, είναι ακριβή από άποψη ενέργειας που σπαταλιέται. Ο κατάλληλος χρονισμός, έτσι, των περιόδων προσοχής, μέσω της πρόβλεψης του χρόνου εμφάνισης της επόμενης μονάδας πληροφορίας

1 Ο αγγλικός όρος *entrainment* έχει διαφορετικό περιεχόμενο από τον όρο *synchronization* στην καθημερινή του χρήση, καθώς αναφέρεται στη βιολογική αντίδραση των ανθρώπων σε περιοδικά φαινόμενα και σχετίζεται με την τάση για περιοδική ρυθμική κίνηση σε παλμό (beat).

ενός συνομιλητή, βελτιστοποιεί την αποδοτικότητα της προσοχής (Large & Jones, 1999). Υπάρχει διαφωνία σχετικά με το αν αυτός ο έλεγχος της προσοχής επιτυγχάνεται με τη χρήση νοητικών ταλαντωτών ή χρονομετρητών διαστημάτων (oscillators, interval timers). Επιπλέον, με βάση την οπτική αυτή, ο κοινός συγχρονισμός μεταξύ συνομιλητών είναι καθοριστικό κριτήριο για αποτελεσματική και παρατεταμένη συζήτηση. Ακόμη, μπορούμε να υποθέσουμε πως μία θετική ηδονική αντίδραση στον κατάλληλο χρονισμό θα ήταν εξελικτικά χρήσιμη, καθώς θα προωθούσε, μέσω της συνομιλίας και ενδεχομένως ομαδικής δημιουργίας μουσικής, τους κοινωνικούς δεσμούς μεταξύ ομάδων, που ιστορικά ήταν απαραίτητοι για την επιβίωση των πρώιμων ανθρώπων. Ίσως με αυτό τον τρόπο να εξηγείται η ανθρώπινη απόλαυση του μουσικού συγχρονισμού.

1.7. Μουσική και συναίσθημα

Οι εσωτερικές συναισθηματικές εμπειρίες που προκαλούνται από τη μουσική, παρόλο που για πολλούς αποτελούν την κύρια λειτουργία της, δεν μελετήθηκαν συστηματικά στον κλάδο της μουσικής ψυχολογίας παρά μόνο τις τελευταίες δεκαετίες. Πρωτοπόροι ερευνητές όπως ο Meyer, ο Berlyne και ο Clynes, εξέδωσαν ωστόσο σημαντικά έργα για τη μουσική και το συναίσθημα κατά τις δεκαετίες του '50, '60 και '70 αντίστοιχα, τα οποία έγιναν αντικείμενο ευρείας μελέτης από τη δεκαετία του '90 και ύστερα. Θέματα τα οποία απασχολούν την σημερινή έρευνα στον κλάδο περιλαμβάνουν τη μέτρηση μουσικών συναισθημάτων, μελέτη του κοινωνικού πλαισίου στο οποίο αυτά προκύπτουν, τον ρόλο τους στην υγεία, τη φαινομενολογία τους, συγκρίσεις διαφορετικών πολιτισμών, τους υποκείμενους μηχανισμούς, το στοιχείο του χρόνου και την νευρωνική απεικόνισή τους (Juslin & Sloboda, 2010). Όσον αφορά την εξελικτική κατανόηση των συναισθηματικών αποτελεσμάτων της νοητικής επεξεργασίας της μουσικής, υπάρχουν διαφορετικές απόψεις στη βιβλιογραφία. Χαρακτηριστικά παραδείγματα αποτελούν η θεωρία του Pinker, που υποστηρίζει πως δεν υπάρχει τέτοια εξελικτική εξήγηση, ενώ ο Huron προτάσσει πως οι συναισθηματικές αντιδράσεις στη μουσική προκαλούνται από τον χειρισμό των προσδοκιών των ακροατών, ενώ υπάρχει πληθώρα άλλων θεωριών για το θέμα (Huron, 2006; Pinker, 2009).

2. Υπολογιστική Δημιουργικότητα

2.1. Κλάδος και αντικείμενο

Ο τομέας της υπολογιστικής δημιουργικότητας, στον οποίο η έρευνα ξεκίνησε πραγματικά κατά τη δεκαετία του '90, αποτελεί έναν υποκλάδο της τεχνητής νοημοσύνης και ορίζεται από τους Colton & Wiggins ως εξής: "Η φιλοσοφία, επιστήμη και μηχανική υπολογιστικών συστημάτων τα οποία, αναλαμβάνοντας συγκεκριμένες αρμοδιότητες, επιδεικνύουν συμπεριφορές που μη βεβιασμένοι παρατηρητές θα χαρακτήριζαν δημιουργικές" (2012). Οι συμπεριφορές αυτές

περιλαμβάνουν τόσο καλλιτεχνικές δραστηριότητες, που παραδοσιακά θεωρούνταν δημιουργικές, αλλά και επιστημονικές, όπως τα μαθηματικά. Ο τομέας αποτελείται από επιστήμονες με ποικίλες ειδικεύσεις, κάποιοι εκ των οποίων επικεντρώνονται σε συγκεκριμένες δημιουργικές εργασίες (όπως παραγωγή μουσικής, ζωγραφικής, ποίησης, αστείων ή αφήγησης), κάποιοι σε γενικότερους δημιουργικούς μηχανισμούς (όπως δημιουργική χρήση γλώσσας ή νέες συλλογιστικές μεθόδους) ενώ άλλοι στο γενικό πλαίσιο εντός του οποίου εμφανίζεται η δημιουργικότητα και πως αξιολογείται. Όπως και σε άλλους υποκλάδους της τεχνητής νοημοσύνης, κάποιοι εστιάζουν στην ανθρώπινη δημιουργικότητα και άλλοι στην κατασκευή συστημάτων που είναι δημιουργικά υπό τους δικούς τους όρους. Μία διαφορά του τομέα από τη ευρύτερη έρευνα της τεχνητής νοημοσύνης είναι πως συχνά δεν υπάρχει καλώς ορισμένη λύση σε κάποιο πρόβλημα, ούτε και καλώς ορισμένο πρόβλημα. Το ίδιο φαινόμενο παρατηρείται και στη μελέτη της Τεχνητής Γενικής Νοημοσύνης (Artificial General Intelligence, AGI), ενώ η γνωσιακή αρχιτεκτονική αποτελεί μία προσέγγιση που δύναται να ενώσει τους δύο υποκλάδους της τεχνητής νοημοσύνης.

2.2. Ορισμοί της δημιουργικότητας

Ήδη από τις αρχές του 20ου αιώνα, δημοσιεύτηκαν πολλές θεωρίες για τη δημιουργικότητα, μεταξύ άλλων από τους Guilford και Csikszentmihalyi. Η θεωρία του Wallas είναι ίσως η πιο συγκεκριμένη από αυτές τις πρώιμες θεωρίες, και έτσι προσφέρεται για υπολογιστική μελέτη. Πιο συγκεκριμένα, ο Wallas εντόπισε τέσσερα δημιουργικά στάδια (1926):

1. Προετοιμασία: ο δημιουργικός σκοπός ταυτοποιείται και παρατηρείται,
2. Επώαση: Δεν υπάρχει συνειδητή προσπάθεια για δημιουργικότητα, αλλά μη συνειδητή,
3. Διαφώτιση: Η στιγμή που μία ιδέα εμφανίζεται στη συνείδηση (“Aha” moment),
4. Επαλήθευση: Η νέα ιδέα εφαρμόζεται.

Η θεωρία αυτή κάνει τη χρήσιμη διάκριση μεταξύ συνειδητής και μη συνειδητής αυθόρμητης δημιουργικότητας, που στον πραγματικό κόσμο εντοπίζεται στις διαφορετικές διαδικασίες που ακολουθούνται από έναν επαγγελματία δημιουργό, ο οποίος δεν μπορεί να περιμένει την έμπνευση για να δημιουργήσει, και έναν καλλιτέχνη χωρίς χρονικούς περιορισμούς, ο οποίος βασίζεται στην αυθόρμητη δημιουργικότητα.

Οι πρώτες συστηματικές μελέτες σχετικά με την ανθρώπινη δημιουργικότητα στη σύγχρονη γνωστική ψυχολογία πραγματοποιήθηκαν μεταξύ άλλων τον Rhodes (1961), ο οποίος μελέτησε τους παράγοντες που υποκινούν την δημιουργικότητα στους ανθρώπους και την Amabile (1983), η οποία μελέτησε τους εξωγενείς παράγοντες που την επηρεάζουν. Πρόσφατα, η Jordanous (2016) επανάφερε τη θεωρία του Rhodes στην τρέχουσα επιστημονική συζήτηση, και συγκεκριμένα την υποδιαίρεση της δημιουργικότητας σε τέσσερις προοπτικές ονόματι “The four P’s of creativity: Person, Process, Product, Press”. Η προοπτική του ανθρώπου ασχολείται

με τον δημιουργό (που μπορεί να είναι και μη ανθρώπινος πράκτορας) και τι τον κάνει δημιουργικό. Η προοπτική της διαδικασίας ασχολείται με τις δράσεις που λαμβάνονται ώστε να παραχθεί ένα δημιουργικό κατασκεύασμα. Η προοπτική του προϊόντος, το οποίο μπορεί να είναι ένα αντικείμενο τέχνης ή μια επιστημονική ιδέα, ασχολείται με τα χαρακτηριστικά του αυτά, που το κάνουν άξιο να αποκληθεί δημιουργικό. Η προοπτική του τύπου μελετά την περιβάλλουσα κουλτούρα που ορίζει τι εκλαμβάνεται ως δημιουργικό και τι όχι σε κοινωνικό επίπεδο.

Οι Jordanous & Keller δίνουν μία ενδελεχή προσέγγιση στον ορισμό της δημιουργικότητας χρησιμοποιώντας στατιστικές μεθόδους ώστε να αποσπάσουν δεκατέσσερα βασικά χαρακτηριστικά της δημιουργικότητας, όπως αυτά εντοπίζονται σε σχετικές επιστημονικές έρευνες, τα οποία δίνονται σε μη ιεραρχημένη λίστα: ενεργή συμμετοχή και επιμονή, αντιμετώπιση της αβεβαιότητας, βαθιά γνώση τομέων, γενική νοητική ικανότητα, παραγωγή αποτελεσμάτων, ανεξαρτησία και ελευθερία, πρόθεση και συναισθηματική ανάμιξη, πρωτοτυπία, εξέλιξη και ανάπτυξη, κοινωνική διάδραση και επικοινωνία, αυθόρμητη/υποσυνείδητη επεξεργασία, σκέψη και εκτίμηση, αξιολόγηση, ποικιλία-απόκλιση-πειραματισμός (2016).

Μία ακόμη ενδιαφέρουσα άποψη για την υπολογιστική δημιουργικότητα, σε σχέση με την ανθρώπινη δημιουργικότητα έρχεται από τον Marvin Minsky, τον αποκαλούμενο «πατέρα της τεχνητής νοημοσύνης», ο οποίος με τα γραπτά του και την ενεργή παρουσία του για 60 χρόνια στον τομέα, υποστήριξε την συμβολική προσέγγιση της έρευνας στο AI, εξέλιξε μαθηματικές μεθόδους (Perceptrons, Turing Machines), πρότεινε υπολογιστικά μοντέλα (όπως το γνωστικό πολυπρακτορικό μοντέλο "Society of Mind", 1986) ασκώντας μεγάλη επιρροή. Η βαθιά κατανόησή του για το πεδίο γίνεται φανερό και στο ότι παρότι έχουν περάσει σχεδόν 40 χρόνια από τότε που έγραψε το σχετικό άρθρο του, είναι ακόμα επίκαιρο (1982). Ο Minsky, λοιπόν, αρχικά αναφέρει την συνηθισμένη άποψη που παρουσιάστηκε στην εισαγωγή, ότι δηλαδή οι υπολογιστές είναι αποκλειστικά μηχανές για «βαρετούς» υπολογισμούς και ότι δεν θα μπορέσουν ποτέ να είναι δημιουργικοί, αυθόρμητοι ή συναισθηματικοί ούτε θα μπορέσουν πραγματικά να νομίσουν, πιστέψουν ή καταλάβουν οτιδήποτε όπως κάνουν οι άνθρωποι. Ωστόσο, τίθεται ενάντια σε αυτή, υποστηρίζοντας πως πρώτα πρέπει να αναπαραγάγουμε την απλή ανθρώπινη συμπεριφορά με υπολογιστικά μέσα πριν προσπαθήσουμε να το κάνουμε για σπουδαία έργα της ανθρώπινης νόησης. Σχετικά με την εξέλιξη των υπολογιστών, τονίζει ότι με την πάροδο του χρόνου οδηγούμαστε σε όλο και πιο ευέλικτα προγράμματα που έχουν όλο και λιγότερη ανάγκη προγραμματιστών και πως το τελικό στάδιο θα είναι προγράμματα τα οποία για οποιοδήποτε πρόβλημα εφαρμόζουν τις καλύτερες γνωστές μεθόδους από μνήμης, συμπεριλαμβανομένου του προβλήματος να γράψουν ένα καλύτερο πρόγραμμα για να αντικαταστήσουν τα ίδια. Αναφερόμενος στην ανθρώπινη δημιουργικότητα, υποστηρίζει την άποψη πως αυτή δεν υπάρχει πραγματικά με την έννοια της μαγείας υπό την οποία συχνά την αντιλαμβανόμαστε, αλλά ότι αποτελεί αποτέλεσμα της αδυναμίας του εκάστοτε ανθρώπινου κριτή να καταλάβει την διανοητική διαδικασία του καλλιτέχνη. Στρέφει λοιπόν την προσοχή του

στην καθημερινή δημιουργικότητα, που όλοι οι άνθρωποι εφαρμόζουμε κατά τη διάδραση με το περιβάλλον μας εν αγνοία μας, αλλά που αποτελεί ένα από τα μεγαλύτερα άλυτα μυστήρια της επιστήμης: ποιος είναι δηλαδή ο τρόπος με τον οποίο ο εγκέφαλος μας υπολογίζει όλες τις πολυπληθείς παραμέτρους τόσο γρήγορα και αποτελεσματικά που να μας κάνουν λειτουργικούς στα περιβάλλοντά μας.

Σχετικά με τη διαφορά των καθημερινών μυαλών με τα εξαιρετικά πρωτοπόρα/δημιουργικά, ο Minsky την εντοπίζει σε έναν ιδιαίτερο συνδυασμό δεξιοτήτων, που κάποιος έτυχε να αποκτήσει κατά τη διάρκεια της ζωής του, μαζί με έντονο ενδιαφέρον για κάποιον τομέα, μεγάλη ευχέρεια σε αυτόν και αυτοπεποίθηση ώστε να μπορέσει να ξεφύγει από τα πρότυπα. Συμπεραίνει έτσι πως η κοινή καθημερινή λογική, που είναι τόσο απρόσιτη στην επιστήμη, όταν ισορροπείται με σωστό τρόπο και με αποκτά έντονο κίνητρο, παράγει δημιουργική ιδιοφυΐα. Ο τρόπος που συμβαίνει αυτό είναι συνδεδεμένος με τα μετα-επίπεδα της μάθησης. Οι δημιουργικοί άνθρωποι μαθαίνουν καλύτερους τρόπους για την ίδια τη μάθηση και ως αποτέλεσμα αποκτούν περισσότερες και βαθύτερες δεξιότητες, καθώς και αποτελεσματικούς τρόπους διαχείρισης και εξισορρόπησης τους (παρατίθεται το παράδειγμα του συνθέτη, ο οποίος συνδυάζει δεξιότητες παραγωγής μουσικών φράσεων και θεμάτων -ανάλογων της καθημερινής ομιλίας- με δεξιότητες της φόρμας -ανάλογων της απλής αφήγησης ιστοριών- με τέτοιο τρόπο ώστε τα στοιχεία αυτά να βρίσκονται σε ισορροπία). Οι περισσότεροι άνθρωποι δεν ενδιαφέρονται ποτέ στη ζωή τους να αποκτήσουν ικανότητες που θα τους βοηθήσουν να μάθουν γρηγορότερα, καθώς ο κόπος που θα έπρεπε να καταβάλλουν δεν αποδίδει άμεσα. Μάλιστα, ο τρόπος που ένα παιδί μαθαίνει να μαθαίνει αποτελεσματικότερα από ένα άλλο, μπορεί κάλλιστα να αποτελεί θέμα τύχης, που παραγάγει εν καιρό μεγάλη ποσοτική διαφορά στις ικανότητες των παιδιών που εκλαμβάνεται ως ποιοτική διαφορά και «βαφτίζεται» δημιουργικότητα. Γιατί όμως, αφού είναι τόσο θεμιτό και προσβάσιμο, δεν είναι περισσότεροι άνθρωποι δημιουργικοί; Ο Minsky απαντά πως πιθανώς η αιτία οφείλεται σε εξελικτικές κοινωνικές δυνάμεις, που για να διατηρήσουν τον κοινωνικό ιστό σταθερό, προωθούν την ταύτιση τρόπων σκέψης μεταξύ ατόμων, σε βάρος της ελεύθερης ανάπτυξης της δημιουργικότητας του κάθε ατόμου.

2.3. Εννοιολογικοί χώροι και δημιουργικότητα

Η γνωσιακή επιστήμονας Margaret Boden ασχολήθηκε εκτενώς με την αποσαφήνιση της διαδικασίας της δημιουργικότητας και εισήγαγε τόσο έννοιες από άλλες θεωρίες της δημιουργικότητας όσο και νέες έννοιες, οι οποίες έκτοτε χρησιμοποιήθηκαν ευρέως από τον ερευνητικό τομέα της υπολογιστικής δημιουργικότητας (1998, 2009) Μία σπουδαία συμβολή της Boden στη σκέψη για τη δημιουργικότητα ήταν η εισαγωγή του όρου «εννοιολογικός χώρος» (conceptual space), δηλαδή ο χώρος όπου βρίσκονται οι διαθέσιμες ιδέες, κάποιες από τις οποίες έχουν εξερευνηθεί και κάποιες από τις οποίες δεν έχουν ακόμα ανακαλυφθεί.

Στη θεωρία της η Boden κάνει ξεκάθαρο πως θεωρεί τη δημιουργικότητα βασικό χαρακτηριστικό της ανθρώπινης νοημοσύνης και όχι το προνόμιο λίγων, όπως συχνά προβάλλεται στην κοινωνία. Ορίζει λοιπόν τη δημιουργικότητα ως την ικανότητα παραγωγής νέων και πολύτιμων ιδεών, αποσαφηνίζοντας ότι πολύτιμες εννοεί τις αυτές που είναι χρήσιμες για κάποιο σκοπό, αισθητικά απολαυστικές, πλούσιες σε περιπλοκότητα ή ενδιαφέρουσες με άλλους τρόπους και ότι οι ιδέες αυτές δεν χρειάζεται να είναι μονάχα θεωρίες, αλλά μπορεί να είναι αφηγήσεις, ερμηνείες, έννοιες αλλά και να επιδέχονται φυσική υλοποίηση όπως πίνακες, γλυπτά, σπίτια ή ακόμα και κινητήρες αεροπλάνων. Όσον αφορά τον χαρακτηρισμό νέες για τις ιδέες αυτές, κάνει τον διαχωρισμό μεταξύ P-Creativity (Psychological-Ψυχολογική) και H-Creativity (Historical-Ιστορική), η πρώτη εκ των οποίων δηλώνει την δημιουργία μίας ιδέας νέας για το ον που την παρήγαγε ενώ η δεύτερη δηλώνει μια ιδέα που δεν έχει παραχθεί ξανά στην ιστορία της ανθρωπότητας. Τονίζει μάλιστα ότι είναι το πρώτο είδος στο οποίο οφείλει να επικεντρωθεί η έρευνα της τεχνητής νοημοσύνης, καθώς αν επιτευχθεί H-Creativity τότε το P-Creativity θα είναι εντός της έκτασης της τεχνολογίας, και καλεί προς αποφυγή του superhuman-human fallacy, κατά το οποίο η μη ικανότητα της τωρινής τεχνολογίας να παραγάγει P-Creativity σημαίνει πως δεν θα μπορέσει ποτέ να την επιτύχει.

Η Boden κάνει και έναν δεύτερο διαχωρισμό της δημιουργικότητας, ανάλογα με το είδος εξερεύνησης του εννοιολογικού χώρου στο οποίο βασίζεται. Συγκεκριμένα, διαχωρίζει την συνδυαστική (combinational), την διερευνητική (exploratory) και την μεταμορφωτική (transformational) δημιουργικότητα. Στην συνδυαστική δημιουργικότητα δημιουργούνται συνδέσεις μεταξύ οικείων, αλλά μέχρι τότε ασύνδετων ιδεών, όπως για παράδειγμα ένα μεγάλο μέρος της ποίησης, τα collage και άλλες οπτικές συνδυαστικές τέχνες, η μίμηση πουλιών στην ορχήστρα αλλά ακόμα και η χρήση αναλογιών στην επιστήμη. Στη διερευνητική δημιουργικότητα εξερευνώνται άγνωστες πτυχές του χώρου και δοκιμάζονται τα όρια και οι δυνατότητες του χώρου αυτού, και είναι αυτή η κατηγορία στην οποία εμπίπτει το μεγαλύτερο μέρος τέχνης και επιστήμης. Στη μεταμορφωτική δημιουργικότητα αλλάζουν οι ίδιοι οι κανόνες που όριζαν έναν εννοιολογικό χώρο, παράγοντας έτσι έναν καινούργιο χώρο που περιλαμβάνει ιδέες οι οποίες δεν ήταν μέχρι τότε προσιτές, και που είναι μη ενστικτώδεις για τους περισσότερους ανθρώπους, με συχνά τραγικά αποτελέσματα για τους δημιουργούς των ιδεών αυτών (η Boden παραθέτει τα παραδείγματα των Ignaz Semmelweis και του Vincent Van Gogh).

Πραγματεύεται ακόμα το ζήτημα της αποτίμησης των παραγόμενων ιδεών από ένα σύστημα AI το οποίο επιχειρεί να πετύχει δημιουργική συμπεριφορά, θέτοντας το πρόβλημα της παραγωγής μεγάλου αριθμού άχρηστων ιδεών σε σχέση με τον αριθμό των χρήσιμων στα συστήματα αυτά. Με αυτό εννοείται ότι σε πολλά υπάρχοντα δημιουργικά συστήματα AI, το μεγαλύτερο μέρος των παραγόμενων προϊόντων τους, παρά την καινοτομία τους, δεν εκτιμώνται ως αξιόλογα από τους ανθρώπους. Μία πιθανή λύση στο πρόβλημα αυτό θα ήταν ίσως η χρήση εσωτερικών μηχανισμών αξιολόγησης, όπως για παράδειγμα ένας γενετικός αλγόριθμος που εξελίσσεται και ο ίδιος μαζί με ολόκληρο το σύστημα. Για να γίνουν τέτοιοι μηχανισμοί δυνατοί, θα πρέπει οι προγραμματιστές τους να μπορούν να εκφράσουν με

ακρίβεια τις σχετικές ανθρώπινες αξίες ώστε να αναγνωριστούν από το σύστημα, πράγμα ευκολότερο για πολιτισμικά αποδεκτούς εννοιολογικούς χώρους, παρά για διευρυμένους χώρους, οι οποίοι θα χρειάζονταν εν μέρει διαφορετικούς κανόνες αξιολόγησης από αυτούς του χώρου από τον οποίο δημιουργήθηκαν. Η δυσκολία στη διαδικασία αυτή συνδέεται με την δυσκολία των ανθρώπων να εκφράσουν λεκτικά ή με υπολογιστικά μέσα, που ακριβώς έγκειται η εκτίμησή ή απέχθειά τους για ένα καλλιτεχνικό προϊόν, που έχει να κάνει ίσως με την στήριξη τέτοιων ειδών αξιολόγησης σε κινητήριους και συναισθηματικούς παράγοντες και που αλλάζει μεταξύ κουλτουρών και εποχών. Τα επιστημονικά κριτήρια της θεωρητικής κομψότητας και συνοχής είναι ίσως ευκολότερο να οριστούν και να εφαρμοστούν, ωστόσο και στον επιστημονικό τομέα εμφανίζεται το φαινόμενο της «μόδας», που έχει να κάνει με πολιτικά-κοινωνικά ζητήματα που σχετίζονται με τις επιστημονικές ανακαλύψεις, χωρίς ωστόσο να γίνονται αυτά συνειδητά από τους ίδιους τους επιστήμονες.

Ο Geraint Wiggins δημιούργησε ένα γενικό μαθηματικό πλαίσιο για τη θεωρία της Boden, το Creative Systems Framework (CSF – Πλαίσιο Δημιουργικών Συστημάτων), ώστε να γίνει εφικτή η σύγκριση και έτσι η καλύτερη κατανόηση των δημιουργικών συστημάτων (2006). Η χρήση του όρου του “εννοιολογικού χώρου” από την Boden είναι λιγότερη ακριβής και αυστηρά ορισμένη από την τυπική διατύπωση του Wiggins αλλά και τους γεωμετρικούς εννοιολογικούς χώρους του Gärdenfors.

3. Θεωρία πληροφορίας

3.1. Ιστορικό Πλαίσιο

Η θεμελιώδης σημασία εργασία του Shannon στη θεωρία της πληροφορίας (1948) επηρέασε εκτός από τεχνολογικούς τομείς, την ψυχολογία, τη γλωσσολογία αλλά και τη μουσική, στην οποία βρήκε εφαρμογές κατά τις δεκαετίες του '50 και '60. Ωστόσο, η εφαρμογή της θεωρίας αντιμετώπισε δυσκολίες λόγω ανεπαρκειών των Μαρκοβιανών μοντέλων για τους σκοπούς της μοντελοποίησης (ειδικά της γλώσσας όπως ανέφερε ο Chomsky) καθώς και λόγω περιορισμών στο μέγεθος των διαθέσιμων μουσικών συλλογών και της επεξεργαστικής ισχύος των υπολογιστών. Έτσι, η θεωρία του Shannon έπεσε στην αφάνεια κατά τις επόμενες δεκαετίες, κατά τη διάρκεια δηλαδή της γνωσιακής επανάστασης, όπου χρησιμοποιούνταν κατά κύριο λόγο συμβολικά μοντέλα. Αυτό άλλαξε με την άνοδο του συνδεδετισμού στα μέσα της δεκαετίας του '80 που στη μουσική μουσική έφτασε στα μέσα επόμενης δεκαετίας, όπου οι Conklin και Witten αντιμετώπισαν πολλές από τις αρχικές δυσκολίες εφαρμογής των μεγεθών της θεωρίας πληροφορίας (1995). Στη γνωσιακή επιστήμη είναι σήμερα κοινά αποδεκτή η οπτική ότι το μυαλό λειτουργεί σαν μία μηχανή επεξεργασίας πληροφοριών που προσπαθεί να είναι όσο πιο αποδοτική όσο γίνεται στο περιβάλλον του, επομένως η θεωρία πληροφορίας μπορεί να χρησιμοποιηθεί άμεσα στη σχετική θεωρητική και εμπειρική έρευνα.

3.2. Εφαρμογή σε διαδοχικές δομές και κατηγοριοποίηση

Σύμφωνα με τον Shannon, η μεταφορά πληροφορίας από μία πηγή σε έναν δέκτη δύναται να κωδικοποιηθεί σε bits (binary digits = δυαδικά ψηφία), που αποτελούν τα μικρότερα δυνατά άτομα πληροφορίας. Στην περίπτωση που υπάρχει ήδη αποκτημένη κοινή πληροφορία μεταξύ πηγής και δέκτη, τότε το μέγεθος ενός μηνύματος που περιλαμβάνει μέρος αυτής της κοινής πληροφορίας, μετρημένο σε bits, μειώνεται. Μάλιστα, αν πάρουμε ένα στατιστικό μοντέλο ενός αλφάβητου, τότε είναι δυνατό να υπολογίσουμε προσεγγιστικά την πιθανότητα εμφάνισης κάποιου συμβόλου του αλφαβήτου, σύμφωνα με τον τύπο εντροπίας του Shannon:

$$H = - \sum_{i \in A} p_i \log_2 p_i$$

όπου H =εντροπία (αποτελεί μέτρο της αταξίας της κατανομής, το αντίθετο της δομής), A =αλφάβητο, i =σύμβολο στο αλφάβητο και p_i =πιθανότητα εμφάνισης του συμβόλου i . Αν σε κάποια στιγμή, ένα σύμβολο έχει πιθανότητα εμφάνισης 1, τότε η εντροπία μηδενίζεται και η κατανομή δεν περιλαμβάνει καθόλου πληροφορία. Όταν, αντίθετα, είναι εξίσου πιθανό να εμφανιστεί οποιοδήποτε σύμβολο, τότε η εντροπία μεγιστοποιείται και παίρνει τιμή ανάλογη με το μέγεθος του αλφάβητου.

Η ερμηνεία αυτή εφαρμόζεται και σε δυναμικές κατανομές, n -οστής τάξης, και έτσι καθίσταται δυνατός ο υπολογισμός της πιθανότητας εμφάνισης κάποιου συμβόλου και της εντροπίας της εκάστοτε κατανομής. Αυτοί οι προσεγγιστικοί υπολογισμοί μεταβάλλονται κατά τη διάρκεια μιας ακολουθίας συμβόλων καθιστώντας ένα τέτοιο μοντέλο δυναμικό. Υπάρχουν, επιπλέον, πειραματικές ενδείξεις πως η εντροπία συνδέεται με την αίσθηση αβεβαιότητας των ανθρώπων, υποδεικνύοντας πως ίσως και οι ανθρώπινες προσδοκίες αναπαρίστανται ως στατιστικές κατανομές (Hansen & Pearce, 2014).

Σύμφωνα με τον MacKay, μπορούμε να χρησιμοποιήσουμε μία μαθηματικά απλουστευμένη μορφή της εντροπίας του Shannon, ώστε να υπολογίσουμε προσεγγιστικά τον αριθμό απαραίτητων bits για ένα συγκεκριμένο, γνωστό σύμβολο (2003). Η τιμή αυτή λέγεται ποσότητα πληροφορίας (information content) και ορίζεται ως εξής:

$$h_i = -\log_2 p_i$$

όπου h =ποσότητα πληροφορίας, και τα υπόλοιπα σύμβολα είναι ίδια με αυτά της εντροπίας.

Η τιμή αυτή χρησιμοποιείται όταν μελετάται ο ρυθμός μετάδοσης πληροφορίας στη φυσική γλώσσα, αλλά και στη μουσική, που εντός του πλαισίου του συστήματος IDyOM, προβλέπει την ανθρώπινη εμπειρία της απροσδοκίας (δηλαδή της αίσθησης εμφάνισης ενός απρόοπτου γεγονότος) με μεγάλη επιτυχία. Όσον αφορά την βελτιστοποίηση των αναπαραστάσεων του μοντέλου, αυτή επιτυγχάνεται μέσω ελαχιστοποίησης της διασταυρωμένης εντροπίας (cross-entropy).

Η εντροπία και η ποσότητα πληροφορίας έχουν χρησιμοποιηθεί και για εργασίες τεμαχιοποίησης (chunking) στη γλώσσα και τη μουσική, όπου παρουσιάζουν συντακτικά παρόμοια μοτίβα. Ο τρόπος εφαρμογής της τεμαχιοποίησης είναι ο εξής: Όταν ξεκινά η επεξεργασία ενός τεμαχίου, του οποίου η αλληλουχία συμβόλων έχει απομνημονευτεί και περάσει στην στατιστική κατανομή, τότε κατά τη διάρκεια της επεξεργασίας του, η τιμή της εντροπίας είναι χαμηλή επειδή το σύστημα “γνωρίζει” πιο σύμβολο είναι πιθανότερο να ακολουθήσει σε κάθε στιγμή. Ωστόσο, μετά το τέλος του τεμαχίου, η εντροπία της κατανομής και η ποσότητα πληροφορίας του επόμενου συμβόλου αυξάνονται και έτσι μπορούμε να εντοπίσουμε τεμάχια με βάση τις διακυμάνσεις στις τιμές αυτές.

Τα μεγέθη που προκύπτουν από τη θεωρία πληροφορίας μπορούν να εφαρμοστούν και στην κατηγοριοποίηση εντός της μνήμης, εργασία η οποία επιδέχεται διαφορετικές προσεγγίσεις. Μία από αυτές είναι η μαθηματική/υπολογιστική προσέγγιση, όπως ο αλγόριθμος ID3 (Iterative Dichotomiser), ο οποίος ελαχιστοποιεί την εντροπία προκειμένου να μεγιστοποιήσει την ποσότητα πληροφορίας για ένα δοσμένο σύστημα κατηγοριοποίησης. Από την άλλη, υπάρχει η γνωσιακή προσέγγιση, η οποία δεν χρησιμοποιεί κάποιο ορισμένο εκ των προτέρων σύστημα κατηγοριοποίησης, αλλά δημιουργεί καινούργια συστήματα, με βάση τα δεδομένα που έχει, και μάλιστα μπορεί να τροποποιήσει το σύστημα κατηγοριοποίησης που

χρησιμοποιεί, αν οι τιμές της εντροπίας και ποσότητας πληροφορίας νέων δεδομένων το απαιτούν, προκειμένου να διατηρηθεί η αποδοτικότητα του συστήματος.

4. Εννοιολογικοί χώροι

Όπως αναφέρθηκε, ο Gärdenfors ανέπτυξε και επέκτεινε μία γεωμετρική θεωρία εννοιολογικών χώρων, που στην ουσία επιχειρεί να γεφυρώσει το επιστημονικό κενό μεταξύ συμβολικών – διακριτών (όπως η μαθηματική λογική) και των μη-συμβολικών – συνεχών (όπως τα νευρωνικά δίκτυα) αναπαραστάσεων (2014). Τα δύο άκρα αυτά συνδέονται μέσω ενός ενδιάμεσου χώρου, ο οποίος είναι συνεχής, λίγων διαστάσεων και ο οποίος επιδέχεται γεωμετρικό λογισμό. Ο χώρος αυτός είναι το μέσο αντιστοίχισης συμβολικών εννοιών με την πολύπλοκη δομή των αντιληπτικών δεδομένων που διαχειρίζεται για παράδειγμα ένα νευρωνικό δίκτυο. Η αντιστοίχιση αυτή, έχει δειχθεί, ότι είναι ικανή να εξηγήσει την αντίληψη του χρώματος αλλά και κινήσεων στο φυσικό κόσμο. Στο χώρο της μουσικής, έχουν προταθεί γεωμετρικοί χώροι για τον τόνο και το ρυθμό (βλ. 1.4., Forth et al., 2010). Χωρίς επιπλέον ορισμούς και προδιαγραφές για τους εννοιολογικούς χώρους, αυτοί είναι όμοιοι με τους χώρους της Boden, όπως τυποποιήθηκαν από τον Wiggins, με την προσθήκη γεωμετρικών ιδιοτήτων και συνέχειας.

Στη μελέτη των αναπαραστάσεων του νοήματος έχουν προταθεί πολλών ειδών λύσεις, που προωθούν τη δημιουργικότητα, ωστόσο οι αναπαραστάσεις που χρησιμοποιούν δίκτυα συχνά είναι περιορισμένες ή δύσκολες στην εφαρμογή τους (Xiao et al., 2019). Αντίθετα, οι αναπαραστάσεις σε εννοιολογικούς χώρους προσφέρουν σαφείς απαντήσεις σε ερωτήσεις αναφορικά με τη σχέση μεταξύ εννοιών του χώρου αυτού. Πιο συγκεκριμένα, είναι εφικτό να θέσει κανείς τις ερωτήσεις: πόσο κοντά βρίσκεται αυτή η έννοια σε μία άλλη; ποια έννοια βρίσκεται ανάμεσά τους; είναι αυτή η έννοια ίδια με μία άλλη; που τοποθετείται μία νέα έννοια στον εννοιολογικό χώρο; Οι απαντήσεις δίνονται με απλή εφαρμογή γεωμετρικού λογισμού στον χώρο λίγων διαστάσεων στον οποίον βρίσκονται οι έννοιες.

Ωστόσο, η θεωρία του Gärdenfors προδιαγράφει αναλυτικά ακόμη περισσότερες ιδιότητες των εννοιολογικών χώρων από αυτές. Οι χώροι στη θεωρία του έχουν διαφορετικούς τύπους. Κάποιοι, όπως ο χώρος των χρωμάτων, ορίζονται με βάση ακέραια σύνολα διαστάσεων, οι οποίες δεν έχουν νόημα, εκτός αν όλες οι διαστάσεις είναι παρούσες (π.χ. αντίθεση, φωτεινότητα, ευκρίνεια). Από την άλλη, υπάρχουν χώροι που αποτελούνται από διαχωριστές διαστάσεις, οι οποίες μπορούν να χρησιμοποιηθούν ανεξάρτητα από την παρουσία άλλων διαστάσεων. Επιπλέον, κάποιοι εννοιολογικοί χώροι ορίζουν την έννοια της απόστασης χρησιμοποιώντας Ευκλείδεια μετρική, ενώ άλλοι με μετρική Manhattan (νορμική ταξί). Ακόμη, είναι δυνατό ορισμένοι χώροι να αναφέρονται σε άλλους χώρους, όπως όταν, για παράδειγμα, ένας χώρος αναπαριστά τις σχετικές θέσεις δύο αντικειμένων, τα οποία αναπαριστώνται σε

άλλους χώρους. Είναι, επιπροσθέτως, εφικτό να υπερτεθούν συλλογές χώρων ώστε να δημιουργήσουν πολύπτυχα (manifold) πολλών διαστάσεων τα οποία είναι ικανά να περιγράψουν, θεωρητικά, κάθε δομή.

Το νόημα, αντιληπτικό ή μη, αναπαρίσταται ως συγκεκριμένες περιοχές στους εννοιολογικούς χώρους αυτούς, ενώ σύμφωνα με τον Gärdenfors, οι κυρτές περιοχές συνιστούν “φυσικές έννοιες” (natural concepts), καθώς σε αυτούς ανήκει κάθε σημείο το οποίο βρίσκεται μεταξύ δύο άλλων σημείων της περιοχής. Κατά συνέπεια, ορίζονται η ισοδυναμία και έγκλειση μεταξύ περιοχών, καθώς και η δημιουργία νέων περιοχών, με βάση τις ιδιότητες του πολύπτυχου. Επίσης, είναι εφικτό να μοντελοποιηθεί η θεωρία πρωτοτύπων, κάνοντας χρήση των κέντρων των περιοχών ενός χώρου. Με βάση τα παραπάνω, μπορούμε να σκεφτούμε τις αναπαραστάσεις του τόνου, τις οποίες αναλύσαμε παραπάνω, ως εννοιολογικούς χώρους, με άξονες που δηλώνουν το απόλυτο τονικό ύψος ή κάποια συγκεκριμένη αναλογία συχνοτήτων και περιοχές σε αυτό τον χώρο που αναπαριστούν η καθεμία διαφορετική νότα ή τονική κλάση.

Όσον αφορά τις δυναμικές, διαδοχικές δομές, αυτές δύνανται να αναπαρασταθούν ως τροχιές σε έναν εννοιολογικό χώρο. Δυστυχώς, η προσέγγιση αυτή είναι περιοριστική, καθώς είναι δύσκολη η σύγκριση μεταξύ τέτοιων τροχιών, πράγμα που υπονομεύει την απλότητα της θεωρίας του Gärdenfors. Μία εναλλακτική πρόταση προσφέρεται από τους Chella et al., και συμπεριλαμβάνει τη χρήση φασματικών αναπαραστάσεων των τροχιών κάποιων αλληλουχιών σε έναν εννοιολογικό χώρο, δηλαδή χρήση του μετασχηματισμού Φουριέ (2015; 2004). Με αυτό τον τρόπο, καθίσταται εφικτή η αναπαράσταση των τροχιών ως σημεία σε εννοιολογικούς χώρους υψηλότερου επιπέδου, εφόσον ορίζονται και αυτοί ενδελεχώς. Χρησιμοποιείται δηλαδή η ίδια η θεωρία εννοιολογικών χώρων για την επίλυση του προβλήματος. Η επέκταση αυτή της θεωρίας του Gärdenfors δεν αλλάζει την βασική ουσία της, αλλά την εμπλουτίζει με νέες δυνατότητες και τρόπους κατανόησης και εξήγησης των φαινομένων που περιγράφει.

5. Γνωσιακή Αρχιτεκτονική

5.1. Στόχοι και μεθοδολογία της έρευνας στον τομέα της γνωσιακής αρχιτεκτονικής
Η έρευνα στις γνωσιακές αρχιτεκτονικές, που γίνεται στα πλαίσια της γνωσιακής επιστήμης, διαφέρει από το μεγαλύτερο μέρος της έρευνας στον κλάδο. Πιο αναλυτικά, στο μεγαλύτερο κομμάτι της γνωσιακής έρευνας μελετώνται μηχανισμοί με σκοπό να εξηγηθούν συγκεκριμένα συμπεριφοριστικά φαινόμενα, αλλά στην έρευνα της γνωσιακής αρχιτεκτονικής μελετώνται γενικοί μηχανισμοί και σκοπός είναι η εύρεση μίας συνολικής λειτουργικής θεωρίας της δομής της νόησης, ενίοτε λαμβάνοντας υπ’ όψη την βιολογική δομή του εγκεφάλου (Langley et al., 2009). Επομένως, ο γενικός στόχος είναι η κατανόηση υψηλού επιπέδου για τη νόηση, το πως οι μηχανισμοί της υπάρχουσας έρευνας συνυπάρχουν και πως είναι δυνατό, σε ορισμένες

περιπτώσεις, να εξηγηθούν φαινόμενα χρησιμοποιώντας λιγότερους, πιο γενικούς μηχανισμούς.

Ο ερευνητής της επιστήμης υπολογιστών και γνωστικής ψυχολογίας Allen Newell, που μαζί με τους Laird και Rosenbloom δημιούργησε τη γνωσιακή αρχιτεκτονική Soar, δίνει μία λίστα ζητούμενων για ενοποιητικές θεωρίες της νόησης (1990):

1. Προσανατολισμός σε στόχους: αναφέρεται στην παραδοσιακή στον τομέα της τεχνητής νοημοσύνης ιδέα, πως οι άνθρωποι είναι θεμελιωδώς εστιασμένοι σε συνειδητούς στόχους. Ωστόσο, σύγχρονες έρευνες έχουν δείξει ότι πιθανώς πολλοί από τους στόχους μας στην πραγματικότητα σχηματίζονται από ένα υποσυνείδητο κύκλωμα, και στη συνέχεια εισέρχονται στη συνειδητή προσοχή μας (Bode et al., 2011).

2. Πλούσιο περιβάλλον: η νόηση είναι θεμελιωδώς ενσωματωμένη σε ένα περιβάλλον, και οι εξελικτικές δυνάμεις που την επιλέγουν εξαρτώνται από την περιπλοκότητα του περιβάλλοντος.

3. Πλούσια γνώση: για την επίλυση ακόμη και απλών προβλημάτων είναι απαραίτητο ένα μεγάλο σώμα γνώσης σε ένα περίπλοκο περιβάλλον.

4. Σύμβολα και αφαιρέσεις: η ικανότητα αναπαράστασης γνώσης σε πολλαπλά επίπεδα αφαίρεσης είναι βασική για έναν συνειδητό πράκτορα, έτσι ώστε να είναι σε θέση να διακρίνει περιπτώσεις από κλάσεις και ιδιότητες από αντικείμενα. Η ικανότητα αφαίρεσης είναι θεμελιώδης για την γενίκευση και τη συλλογιστική.

5. Ευελιξία σε περιβαλλοντικές αλλαγές: όσο πιο νοητικά εξελιγμένο είναι ένα σύστημα, τόσο καλύτερα θα έπρεπε να αντιμετωπίζει αλλαγές στο περιβάλλον.

6. Μάθηση: όσο πιο νοητικά εξελιγμένο είναι ένα σύστημα, τόσο πιο αποτελεσματικά πρέπει να είναι σε θέση να μάθει από την εμπειρία του και να εφαρμόσει τις γνώσεις του.

Όλα τα παραπάνω ζητούμενα εντοπίζονται σε κάποια μορφή εντός των αρχιτεκτονικών που αναφέρονται παρακάτω.

Ένα σημαντικό ζήτημα στην έρευνα των γνωσιακών αρχιτεκτονικών είναι η στενή σύνδεσή τους με την συμβολική προσέγγιση στην τεχνητή νοημοσύνη και πιο συγκεκριμένα η τάση κάποιων ερευνητών να υποβιβάζουν τη νόηση σε επίλυση προβλημάτων. Ωστόσο, η επίλυση προβλημάτων αποτελεί μία μόνο από τις δραστηριότητες του μυαλού, και η υιοθέτηση αυτής της προσέγγισης οδηγεί σε άδηλες υποθέσεις, οι οποίες δεν είναι πάντα επιθυμητές.

Η έρευνα των γνωσιακών αρχιτεκτονικών μπορεί να ομαδοποιηθεί σε δύο βασικές προσεγγίσεις: τη γνωσιακή (cognitivist), όπου υπάρχουν προκαθορισμένα σύμβολα που δίνουν νόημα στο σύστημα και την αναδυόμενη (emergent), όπου το ίδιο το σύστημα σχηματίζει το νόημα και την εσωτερική γλώσσα αναπαράστασης, που κατά κανόνα δεν είναι αναγνώσιμη από ανθρώπους. Μία τρίτη προσέγγιση είναι υβρίδιο των δύο παραπάνω, και επιχειρεί να

εξαγάγει σύμβολα και κανόνες βάσει αντιληπτικού περιεχομένου, που δίνει σε ένα τέτοιο σύστημα το πλεονέκτημα πως η συλλογιστική εντός αυτού μπορεί να εξηγηθεί σύμφωνα με τα σύμβολα που το ίδιο δημιουργεί, βασισμένα στην αντιληπτική εμπειρία, η οποία μπορεί να αλλάξει επηρεάζοντας συγχρόνως και τα σύμβολα. Οι γνωσιακές αρχιτεκτονικές IDyOT και Micro-PSI που αναλύονται στη συνέχεια ανήκουν και οι δύο σε αυτή την υβριδική κατηγορία μεταξύ συμβολικής και αναδυόμενης αρχιτεκτονικής.

Μία από τις πιο αναλυτικές και πρόσφατες επιθεωρήσεις της γνωσιακής αρχιτεκτονικής είναι αυτή του Vernon, όπου τα ζητούμενα που έθεσε ο Newell επεκτείνονται ώστε να συμπεριλάβουν τα εξής: ενσωμάτωση, αντίληψη, δράση, προσμονή, προσαρμογή, κινητοποίηση και αυτονομία (2014). Έχουν προταθεί, επίσης, κριτήρια αξιολόγησης γνωσιακών αρχιτεκτονικών, τα βασικότερα εκ των οποίων είναι η εμπειρική κάλυψη και η οικονομία (Varma, 2011). Επιπλέον μεθοδολογικά κριτήρια για μία γνωσιακή αρχιτεκτονική αποτελούν η ύπαρξη υποκειμενικής και διαυποκειμενικής σημασίας στα πλαίσια της ερευνητικής κοινότητας, η προσφορά ιδιωμάτων τα οποία χρησιμοποιούνται σε μετέπειτα έρευνες και η παραδοξότητα, που συνδυάζει πρωτοτυπία, απροσδοκία και εγκυρότητα. Τα κριτήρια αυτά δεν έχουν να κάνουν άμεσα με την ίδια την αρχιτεκτονική, αλλά συνδέονται με θέματα στη φιλοσοφία της επιστήμης (Lakatos, 1970).

5.2. Χαρακτηριστικά γνωσιακών αρχιτεκτονικών

Ο Anderson, ψυχολόγος και δημιουργός μίας εκ των πρώτων γνωσιακών αρχιτεκτονικών, της ACT (και μετέπειτα ACT-R), παρουσιάζει τον κύριο ισχυρισμό του ως εξής: “Το μόνο που είναι η νοημοσύνη είναι η απλή συσσώρευση και συντονισμός πολλών μικρών μονάδων γνώσης, τα οποία στο σύνολό τους παράγουν σύνθετη νόηση. Το όλο δεν είναι κάτι παραπάνω από το άθροισμα των μερών του, αλλά έχει πολλά μέρη” (All that there is to intelligence is the simple accrual and tuning of many small units of knowledge that in total produce complex cognition. The whole is no more than the sum of its parts, but it has a lot of parts. 1996).

Μία ανάγνωση του παραπάνω ισχυρισμού αντιτίθεται στον δυισμό, όμως μία άλλη διατείνεται πως η διαδικασία συλλογής και οργάνωσης των μερών της νόησης, καθώς και της δυναμικής συνεχούς αναδιοργάνωσης τους δεν είναι κεντρικής σημασίας. Η αρχή του Anderson εντοπίζεται και στις γνωσιακές αρχιτεκτονικές Soar, LIDA, EPIC, οι οποίες είναι βασισμένες σε προγραμματισμένη γνώση και κανόνες, τα οποία αντιμετωπίζονται ως πράκτορες που αντιδρούν σε και μαθαίνουν από δεδομένα (Franklin et al., 2014; Kieras & Meyer, 1995; Laird et al., 1987). Ουσιαστικά, οι αρχιτεκτονικές αυτές προσφέρουν ένα κοινό υπόστρωμα για την δοκιμή πολλών ψυχολογικών θεωριών ταυτόχρονα. Ο Newell γράφει σχετικά πως ΣΥΜΠΕΡΙΦΟΡΑ = ΑΡΧΙΤΕΚΤΟΝΙΚΗ + ΠΕΡΙΕΧΟΜΕΝΟ (όπου το περιεχόμενο περιλαμβάνει κανόνες), πράγμα που δεν αποκλείει την προσθήκη νέων ιδιοτήτων μέσω προγραμματισμού ή μάθησης.

Μία εναλλακτική προσέγγιση λαμβάνεται στις γνωσιακές αρχιτεκτονικές IDyOM και CLARION, η θεωρία των οποίων ορίζει την ίδια την αρχιτεκτονική και είναι βασισμένη σε ένα κεντρικό σύστημα μάθησης. Συγκεκριμένα, η αρχιτεκτονική CLARION αποτελεί συνδυαστικό σύστημα βασισμένο σε τεχνητά νευρωνικά δίκτυα και διαχωρίζει μεταξύ άδηλης και έκδηλης γνώσης, η πρώτη εκ των οποίων αποκτάται ασυνείδητα, ενώ η δεύτερη συνειδητά. Και τα δύο είδη γνώσης επηρεάζουν τη μάθηση του συστήματος, ωστόσο οι τρόποι με τους οποίους επιδρούν στη γενίκευση επί συμβόλων και τεμαχίων είναι δύσκολο να γίνουν αντιληπτοί λεπτομερώς, λόγω της δυσκολίας κατανόησης των νευρωνικών δικτύων (Hélie & Sun, 2010; Wiggins, 2020).

Ένα ζήτημα το οποίο τίθεται στην έρευνα των γνωσιακών επιστημών είναι ο τρόπος που συνδέεται ένα σύστημα με το περιβάλλον του και αποκτά νέα γνώση. Όλες οι αρχιτεκτονικές που αναφέρθηκαν παραπάνω είναι αισθητικές στη θεωρία τους, δηλαδή προτάσσουν ότι νέα γνώση προέρχεται από μεταγωγή από το περιβάλλον. Αν δεν συμπεριλαμβανόταν καμία μορφή μεταγωγής τότε θα ήταν αδύνατο να μελετηθεί η σχέση μεταξύ αναπαράσταση και νοήματος, ενώ υπάρχουν στοιχεία παράλληλης εξέλιξης μεταξύ αναπαραστάσεων και αντιληπτικών δεδομένων (Roberson et al., 2006). Ωστόσο, ο τρόπος που αυτή η μεταγωγή εφαρμόζεται, ποια χαρακτηριστικά των αντιληπμάτων του περιβάλλοντος λαμβάνονται περισσότερο υπ' όψη, πως επηρεάζεται αυτή η διαδικασία από πρότερη γνώση και προβλέψεις για το μέλλον είναι χαρακτηριστικά που μοντελοποιούνται σε διαφορετικούς βαθμούς και με διαφορετικούς τρόπους σε κάθε αρχιτεκτονική.

Όσον αφορά τις αναπαραστάσεις στη μνήμη του συστήματος, η προσέγγιση των περισσότερων γνωσιακών αρχιτεκτονικών έγκειται στην σχεδίαση της μνήμης από ανθρώπους προγραμματιστές, με βάση ψυχολογικές θεωρίες μνήμης που έχουν εμπειρική υποστήριξη (με κυριότερη τη θεωρία των Baddeley & Hitch, 1974). Έτσι, η μνήμη ενός συστήματος συχνά διαχωρίζεται σε βραχύχρονη και μακρόχρονη, εργαζόμενη, επεισοδιακή, σημασιολογική κτλ.

Κάθε σύστημα ή πράκτορας χρειάζεται κάποια εγγενή κινητοποίηση για να πράξει οτιδήποτε και η λειτουργία αυτή ονομάζεται *γνωσιακός κύκλος (cognitive cycle)*. Οι Russell & Norvig δίνουν μία ταξινόμια πρακτόρων και των γνωστικών τους κύκλων, ανάλογα με την ικανότητα επιλογής πράξεων που διαθέτουν (2021):

1. Πράκτορες χωρίς κατάσταση: δεν έχουν μνήμη και αντιδρούν αντανακλαστικά (πχ. θερμοστάτης)
2. Πράκτορες με μνήμη: έχουν εσωτερική κατάσταση και μπορούν να συλλογιστούν για πληροφορία προηγούμενων καταστάσεων, αλλά παραμένουν ουσιαστικά αντανακλαστικοί, απλώς καλύτερα ενημερωμένοι (πχ. σύστημα αυτόματης μαγνητοσκόπησης τηλεόρασης)
3. Πράκτορες με σκοπούς: διαθέτουν στόχους πέρα απ' την αντίδραση σε μεμονωμένα ερεθίσματα και οι αντιδράσεις τους σε αυτά προσαρμόζονται βάσει των μακροπρόθεσμων στόχων τους (πχ. Αυτοκινούμενο όχημα χωρίς χάρτη, αλλά με δυνατότητα οδήγησης σε δρόμους και εφοδιασμένο με πυξίδα, ώστε να μπορεί να επιλέγει κατευθύνσεις στροφών)

4. Πράκτορες βασισμένοι στη χρησιμότητα: χρησιμοποιούν συλλογισμό ώστε να πετύχουν τους στόχους τους με την μεγαλύτερη πιθανή αποδοτικότητα (πχ. Αυτοκινούμενο όχημα το οποίο χρησιμοποιεί χάρτες και ζωντανή πληροφορία για την κίνηση ώστε να εντοπίσει τη συντομότερη διαδρομή)

Ένα είδος πρακτόρων που λείπει από αυτή τη λίστα είναι οι προβλεπτικοί πράκτορες, και αυτό ακριβώς είναι ένα από τα ζητήματα που αναλύονται παρακάτω, υπό την οπτική γωνία της γνωσιακής αρχιτεκτονικής IDyOT.

Β' ΜΕΡΟΣ: Η ΘΕΩΡΙΑ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ IDYOT

6. Εισαγωγικά στοιχεία και επιστημονική προέλευση της θεωρίας IDyOT

6.1. Στόχοι του IDyOT και υπολογιστική μοντελοποίηση

Η γνωσιακή αρχιτεκτονική IDyOT (Information Dynamics of Thinking, δηλαδή Πληροφοριακή Δυναμική της Σκέψης) αποτελεί μία προσπάθεια εξήγησης ανθρώπινων συμπεριφορών που σχετίζονται με την επεξεργασία διαδοχικών δεδομένων, τη δημιουργικότητα και άλλων φαινομένων, εντός ενός κοινού, απλού λειτουργικού πλαισίου (Wiggins & Forth, 2015). Η αρχική πρόταση που κινητοποίησε την εργασία είναι πως ο μηχανισμός της μη-συνειδητής δημιουργικότητας βασίζεται σε στατιστική πρόβλεψη που ρυθμίζεται από πληροφοριακά μεγέθη. Η δομή που περιγράφεται στη θεωρία του IDyOT είναι αφαιρετική και βασίζεται στη λειτουργία και αλληλεπίδραση των μερών που περιγράφονται, ωστόσο η θεωρία σχετίζεται τόσο με νευροφυσιολογικές όσο και εξελικτικές προτάσεις. Ακόμη, η θεωρία του IDyOT συμμαρτίζεται την αρχή της οικονομίας, γνωστή και ως “ξυράφι του Ockham”, δηλαδή προτιμά απλά μοντέλα και εξηγήσεις για τα φαινόμενα που μελετά σε σχέση με σύνθετα μοντέλα με περιττά χαρακτηριστικά.

Το ότι το IDyOT αποτελεί υπολογιστικό μοντέλο πρακτικά σημαίνει ότι έχει τη μορφή υπολογιστικού προγράμματος, το οποίο ενσαρκώνει μία θεωρία και τη δοκιμάζει με διάφορους τρόπους. Αν κάποια από τα αποτελέσματα των δοκιμών είναι μη αναμενόμενα, αυτό αποτελεί ένδειξη ότι ίσως έχει αποφευχθεί η υπερμοντελοποίηση (overfitting), κατά την οποία ένα μοντέλο μπορεί να εξηγήσει σε ικανοποιητικό βαθμό μόνο τα δεδομένα που του δόθηκαν και δεν αποτελεί πλήρη εξήγηση του φαινομένου το οποίο μοντελοποιεί (Honing, 2006). Επίσης, το IDyOT επιχειρεί να είναι επεξηγηματική και όχι περιγραφική θεωρία, που σημαίνει ότι σκοπός είναι η πλήρης κατανόηση των φαινομένων για τα οποία ενδιαφέρεται με βάση το υπόστρωμα απ' όπου αναδύονται και τις δυναμικές που τα τροποποιούν. Πιο συγκεκριμένα, το επίπεδο αφαίρεσης του IDyOT είναι πολύ απομακρυσμένο από το επίπεδο των βιολογικών νευρώνων του εγκεφάλου, και μελετά συνδυασμούς αφαιρετικών μαθηματικών διαδικασιών ώστε να χτίσει ένα επεξηγηματικό μοντέλο σε αυτό το επίπεδο.

6.2. Προσδοκία και αναπαράσταση στο IDyOM

Το σύστημα IDyOM αποτελεί πρόγονο της γνωσιακής αρχιτεκτονικής IDyOT και ενσαρκώνει πολλές από τις αρχές που αναλύθηκαν στο πρώτο μέρος. Ένα χαρακτηριστικό παράδειγμα είναι πως αναδεικνύει τη σημασία της προσδοκίας στη μουσική (μέσω θεωρίας που είναι βασισμένη

στη δουλειά των Meyer, Huron, Pearce & Wiggins, μεταξύ άλλων). Το IDyOM αποτελεί μοντέλο της επεξεργασίας μονοφωνικής μελωδίας, βασίζεται σε Μαρκοβιανή πρόβλεψη και πολυδιάστατες αναπαραστάσεις οι οποίες καταφέρνουν να προσεγγίσουν τις ανθρώπινες μελωδικές προσδοκίες. Η μελωδία είναι μία διακριτή μουσική δομή που εμφανίζεται με διαφορετικό τρόπο σε κάθε πολιτισμό, ενώ μπορεί να διαχωριστεί στη χρονική της συνιστώσα (οικουμενική) και το μουσική της περιεχόμενο (εξαρτώμενο από τον εκάστοτε πολιτισμό). Το σύστημα αποδίδει καλά και σε εργασίες τεμαχιοποίησης, τόσο σε μουσική όσο και σε γλώσσα ενώ παρουσιάζει συμβατότητα με τα ευρήματα της νευροφυσιολογίας της προσδοκίας στους ανθρώπους (M. T. Pearce, Ruiz, et al., 2010).

Ακόμη, υπάρχουν στοιχεία που υποδηλώνουν ότι η μουσική προσδοκία συνεισφέρει στην μουσική συναισθηματική εμπειρία, και συγκεκριμένα οι προβλέψεις του IDyOM σχετίζονται με τις συναισθηματικές αντιδράσεις ανθρώπων ακροατών (Egermann et al., 2013). Έτσι, παρέχεται εμπειρική υποστήριξη στην υπόθεση πως η προσδοκία είναι μία θεμελιώδης λειτουργία για τους ανθρώπους η οποία προσαρμόστηκε εξελικτικά (exapted) ώστε να δημιουργήσει τη μουσική εκτίμηση. Ο μηχανισμός του IDyOM είναι εντελώς γενικός και ανεξάρτητος από τον τομέα που μοντελοποιείται καθώς χρησιμοποιεί μόνο την χρονική ακολουθία των δεδομένων για να κάνει τις προβλέψεις του. Επομένως, η μάθηση γλώσσας και μουσικής είναι στην ουσία η ίδια διαδικασία, εφαρμοσμένη σε διαφορετικές αναπαραστάσεις δεδομένων. Οι προσδοκίες αναπαριστώνται ως στατιστικές κατανομές με βάση τα γνωστά σύμβολα.

Το σύστημα IDyOM έχει επίσης δυναμικές αναπαραστάσεις για τα δεδομένα των ακολουθιών του. Δοθέντων ορισμένων χαρακτηριστικών, χρησιμοποιείται hill-climbing ώστε να επιλεγεί η πιο συμπίεσμένη αναπαράσταση (μέσω της ελαχιστοποίησης της διασταυρωμένης εντροπίας ή cross-entropy). Η υπόθεση πίσω από την υπολογιστική διεργασία αυτή είναι ότι ο εγκέφαλος ψάχνει αναπαραστάσεις που αποθηκεύουν τα μαθημένα δεδομένα όσο πιο αποδοτικά γίνεται, ενδεχομένως ως μέρος της μνημονικής παγίωσης (memory consolidation). Το μοντέλο αυτό είναι περιγραφικό και απαιτεί πολλούς υπολογιστικούς πόρους, αλλά ανακαλύπτει γρήγορα την σχετική αναπαράσταση του τόνου (ξεκινώντας από την απόλυτη αναπαράσταση του) καθώς και αναπαραστάσεις βασισμένες σε μουσικές κλίμακες και τρόπους. Τα παραπάνω στοιχεία του συστήματος IDyOM εντοπίζονται σε κάποια μορφή και στη θεωρία του IDyOT που αναλύεται παρακάτω.

6.3. Global Workspace Theory

Η θεωρία του Οικουμενικού Χώρου Εργασίας (Global Workspace Theory) αποτελεί ένα θεωρητικό και πρακτικό πλαίσιο εντός του οποίου μπορεί να μοντελοποιηθεί η συνείδηση, και είναι βασισμένο σε μία πολυπρακτορική αρχιτεκτονική που χρησιμοποιεί κάτι σαν AI blackboard για την επικοινωνία μεταξύ πρακτόρων, με κάποιους ιδιαίτερους περιορισμούς (Baars, 1988). Παρόλο που ουσιαστικά πρόκειται για μία γνωσιακή αρχιτεκτονική, δεν έχει συγκεκριμένη υλοποίηση, αλλά διάφορες άλλες αρχιτεκτονικές αποτελούν εφαρμογή της

θεωρίας, όπως η αρχιτεκτονική LIDA. Η θεωρία αποφεύγει την ‘δύσκολη’ ερώτηση της συνείδησης (‘hard’ problem of consciousness) και επικεντρώνεται στην ερώτηση “ποιο είναι το περιεχόμενο της συνείδησης και ποιοι μηχανισμοί την υποστηρίζουν”.

Σύμφωνα με τη θεωρία, ο μη συνειδητός νους συνίσταται σε μία μεγάλη συλλογή ειδικών γεννητριών, οι οποίες εκτελούν εργασίες εφαρμόζοντας αλγορίθμους σε δεδομένα και ανταγωνίζονται μαζικά και παράλληλα για πρόσβαση στον οικουμενικό χώρο εργασίας (global workspace), αποκλειστικά μέσω του οποίου δύναται να ανταλλαχθεί πληροφορία μεταξύ των γεννητριών. Η πληροφορία που φέρει κάθε γεννήτρια πρέπει να περάσει ένα κατώφλι ‘σημασίας’ προκειμένου να περάσει στον οικουμενικό χώρο εργασίας. Ο χώρος αυτός είναι πάντοτε ορατός σε όλες τις γεννήτριες και περιλαμβάνει τις πληροφορίες για τις οποίες ο οργανισμός είναι συνειδητός ανά πάσα στιγμή. Ο οικουμενικός χώρος εργασίας είναι ικανός να περιλαμβάνει μόνο ένα ‘πράγμα’ σε δεδομένη χρονική στιγμή, αλλά μπορεί να είναι πληροφορία πολλών διαφορετικών επιπέδων. Το Global Workspace (GW) εξαρτάται ακόμη σε μεγάλο βαθμό από τα συμφραζόμενα και το νόημα που δημιουργείται εντός του, δηλαδή είναι ευαίσθητο και δομημένο με βάση το γενικότερο πλαίσιο, το οποίο μπορεί να συμπεριλαμβάνει σκοπούς, επιθυμίες κ.α.

Ένα πρόβλημα με την αρχική διατύπωση της θεωρίας από τον Baars, είναι το παράδοξο του κατωφλιού (threshold paradox). Το παράδοξο έγκειται στο ότι μόνο ομάδες γεννητριών μπορούν να φτάσουν το κατώφλι για είσοδο στο GW, αλλά η επικοινωνία μεταξύ τους και η δημιουργία ομάδων γεννητριών είναι εφικτή μόνο μέσω του GW. Η θεωρία του IDyOT παρουσιάζει μία λύση στο παράδοξο αυτό, χρησιμοποιώντας μεγέθη της θεωρίας πληροφορίας (Wiggins, 2012). Η δημιουργικότητα στη θεωρία οικουμενικού χώρου εργασίας ταυτίζεται με την είσοδο του περιεχομένου μίας γεννήτριας στο GW, που αποτελεί ίσως το ίδιο φαινόμενο με τη διαφώτιση (‘Aha’ moment) του Wallas. Η ίδια υπόθεση χρησιμοποιείται και στη θεωρία του IDyOT (Wiggins, 2020).

7. Οπερατική χωρωδία του νου και οι γεννήτριες του IDyOT

7.1. Η αναλογία της οπερατικής χωρωδίας

Μέρος των ιδεών σχετικά με το σύστημα IDyOT προέρχονται από την πρωτοπόρα εργασία του Taine, ο οποίος το 1872 εξέδωσε ένα βιβλίο σχετικά με τη νοημοσύνη, όπου παρουσιάζεται η αναλογία του μυαλού με θέατρο (Taine & Hays, 1872). Σύμφωνα με αυτή, στο θέατρο του μυαλού υπάρχουν φώτα τα οποία μπορούν να φωτίζουν μόνο έναν ηθοποιό κάθε στιγμή, ενώ οι υπόλοιποι ηθοποιοί συνεχίζουν να δρουν στο σκοτάδι. Η αναλογία αυτή θυμίζει πολυπρακτορικές θεωρίες του νου, όπως το Society of Mind του Minsky και το Global Workspace Theory του Baars (Baars, 1988; Minsky & Lee, 1988). Εφόσον έχουμε να κάνουμε με θέατρο, τίθεται το ερώτημα ποιο είναι το κοινό. Η απάντηση που δίνει ο Wiggins στο

συγκεκριμένο ερώτημα είναι πως δεν υπάρχει κοινό (το οποίο μπορεί να οδηγήσει στο φιλοσοφικό πρόβλημα του homonculus), αλλά το θέατρο είναι ο τόπος όπου εδράζεται η συνείδηση. Η θέση αυτή βασίζεται στην λειτουργική άποψη του Shanahan και αποφεύγει το 'δύσκολο' πρόβλημα της συνείδησης, προσπαθώντας αρχικά να καταλάβει την θέση της στον χάρτη της νόησης (Shanahan, 2010). Βασικότερα από τη συνείδηση ωστόσο, ο σκοπός της θεωρίας της οπερατικής χορωδίας του νου, που δημιούργησε ο Wiggins στηριζόμενος στην εργασία του Taine και των άλλων επιστημόνων που αναφέρθηκαν, είναι η εξήγηση ενός πιθανού μηχανισμού της δημιουργικότητας, μέσω της σύνδεσης της συνειδητής προσοχής και την ασυνειδητής, χαμηλού επιπέδου νοητικής επεξεργασίας εντός ενός εξελικτικού πλαισίου (Wiggins, 2012).

Σύμφωνα, λοιπόν, με την επέκταση της αναλογίας του θεάτρου από τον Wiggins που την μετέτρεψε σε οπερατική σκηνή με χορωδούς, η συνειδητή προσοχή αναπαρίσταται από το φωτισμένο μέρος της σκηνής του θεάτρου. Αυτό έχει παρόμοιους περιορισμούς με την προσοχή των ανθρώπων: έχει δηλαδή πεπερασμένη χωρητικότητα (περιορισμένη δυνατότητα φωτισμού ηθοποιών στη σκηνή), ωστόσο είναι ευέλικτο και μπορεί να αλλάξει το περιεχόμενό του με ευκολία (τα φώτα μπορούν να κινηθούν εύκολα από κάθε μέρος της σκηνής σε κάθε άλλο). Ανά πάσα στιγμή υπάρχουν πολλαπλά μέλη της χορωδίας στη σκηνή, παρόλο που φωτίζονται λίγα ή μόνο ένα από αυτά. Τα μέλη που δεν φωτίζονται, ωστόσο, αυτοσχεδιάζουν συνεχώς και δίνουν προσοχή στο φωτισμένο μέρος της σκηνής, απ' όπου γίνεται ανταλλαγή πληροφορίας μεταξύ των χορωδών. Οι αυτοσχεδιασμοί των μη φωτισμένων μελών της χορωδίας μπορεί να μην είναι ακουστοί, όταν όμως αρκετοί χορωδοί συγχρονιστούν ή τραγουδήσουν σε ταυτοφωνία, τότε μπορεί να γίνουν το κέντρο της προσοχής και να φωτιστούν, αν είναι αρκετά ηχηροί.

7.2. Οι γεννήτριες και η εξελικτική τους υποστήριξη

Πως όμως μεταφράζεται αυτή η αναλογία σε υπολογιστικό πρόγραμμα και ποια είναι η εξελικτική υποστήριξη της; Η λύση δίνεται από πολλαπλές γεννήτριες, όπως περιγράφεται στη θεωρία Οικουμενικού Χώρου Εργασίας του Baars, καθεμία από τις οποίες αναλαμβάνει το ρόλο των χορωδών του οπερατικού θεάτρου. Οι γεννήτριες αυτές παράγουν προβλέψεις με βάση τη γνώση που έχει ήδη ένας οργανισμός, ενώ ο ίδιος προβλεπτικός μηχανισμός είναι δυνατό να χρησιμοποιηθεί για πολλούς σκοπούς, δηλαδή σε πολλαπλές διαφορετικές αναπαραστάσεις, πράγμα που ενισχύει την εξελικτική του ευλογοφάνεια. Παράλληλα, οι γεννήτριες αυτές επιτρέπουν τη μάθηση και προσαρμογή της συμπεριφοράς του οργανισμού σε νέα, άγνωστα περιβάλλοντα, ενώ μπορούν να συνδυάσουν πολλών ειδών αναπαραστάσεις.

Ωστόσο, τα δεδομένα που αντιλαμβάνεται ένας οργανισμός σε ένα περιβάλλον είναι συχνά σύνθετα και η μάθηση απλών στατικών συσχετισμών, βασισμένων σε συνύπαρξη στοιχείων, δεν αρκούν για την κατανόηση και επιβίωσή του. Έτσι, χρειάζεται η εκμάθηση ακολουθιών δεδομένων και η παραγωγή προβλέψεων με βάση αυτές, ώστε να γίνει η αβεβαιότητα του

περιβάλλοντος διαχειρίσιμη. Επιπλέον, εκτός από ένα μηχανισμό επεξεργασίας διαδοχικών δεδομένων, χρειάζεται και μία μετα-στρατηγική η οποία έγκειται στο εξής: αν ο οργανισμός βρίσκεται σε καταστάσεις στις οποίες η πρόβλεψη με βάση παρελθοντικές εμπειρίες δεν είναι έμπιστη, τότε είναι περισσότερο άγρυπνος και προσεκτικός στο περιβάλλον του. Αυτού του είδους η αλλαγή στην νοητική κατάσταση και επομένως και στη συναισθηματική κατάσταση του οργανισμού, ανάλογα με την αβεβαιότητα των προβλέψεων του είναι ακριβώς το είδος μηχανισμού που ο Huxon υποστηρίζει ότι γέννησε την αισθητική της μουσικής (Huxon, 2006). Τέτοιοι μηχανισμοί εντοπίζονται τόσο σε ζώα όσο και σε ανθρώπους και παρόλο που δεν βασίζονται σε συνειδητό έκδηλο συλλογισμό, τον ενεργοποιούν όταν οι συνθήκες είναι αβέβαιες. Με βάση αυτούς τους μηχανισμούς ο οργανισμός καταφέρνει να μάθει πιο αποδοτικά, και σύμφωνα με τον Gärdenfors η αντιληπτική μάθηση κινητοποιείται από την ανάγκη να γίνουν αντιληπτές οι ομοιότητες και διαφορές μεταξύ οντοτήτων στον κόσμο και οι παρατηρήσεις να τοποθετηθούν στα κατάλληλα σημεία με βάση τις ήδη υπάρχουσες εμπειρίες (Gärdenfors, 2000).

7.3. Πρόβλεψη, στατιστική συχνότητα και απροσδοκία

Η σημαντικότερη ιδιότητα ενός αυτόνομου συστήματος που μαθαίνει είναι η ικανότητά του να φαντάζεται και να αντιδρά εκ των προτέρων. Δοθέντος ενός κατηγοριοποιημένου μοντέλου του κόσμου, ένα σύνολο γεννητριών μπορεί να χρησιμοποιήσει το μοντέλο αυτό για να κάνει προβλέψεις βασισμένες σε διαδοχικούς συνδυασμούς. Οι προβλέψεις στο σύστημα IDyOT γίνονται παράλληλα από έναν μεγάλο αριθμό γεννητριών και έτσι προβλέψεις που γίνονται από πολλές γεννήτριες συγχρόνως γίνονται αντιληπτές ως ηχηρές, ενώ προβλέψεις που γίνονται από λίγες γεννήτριες συγχρόνως δεν γίνονται αντιληπτές. Η πρόταση αυτή βασίζεται σε στατιστικές έννοιες της μάθησης, σύμφωνα με τις οποίες πολλαπλά ανεξάρτητα υποσυστήματα κάνουν προβλέψεις μέσω δειγματοληψίας από ένα προβλεπτικό αλλά ατελές (για λόγους υπολογιστικού κόστους) στατιστικού μοντέλου. Έτσι, τα πιο αναμενόμενα γεγονότα είναι τα πιο πιθανά να προβλεφθούν.

Παρ' όλα αυτά, αν ένας οργανισμός στηριζόταν μόνο σε αυτή την αρχή της στατιστικής συχνότητας θα αποτύγχανε, καθώς δεν θα διέθετε την ικανότητα να φανταστεί απίθανες καταστάσεις και να προετοιμαστεί για αυτές. Για την επίλυση αυτού του θέματος είναι απαραίτητα τα μεγέθη της θεωρίας πληροφορίας που μελετήθηκαν παραπάνω και συγκεκριμένα η ποσότητα πληροφορίας h (information content) και η εντροπία H (entropy). Επιπλέον, ορίζουμε την ένταση p μίας γεννήτριας ως την πιθανότητα επιβεβαίωσης του προβλεπόμενου αποτελέσματος και την απροσδοκία h_t ενός μερικού μοντέλου της πραγματικής τρέχουσας εμπειρίας στην κατάσταση t . Όταν η τρέχουσα εμπειρία είναι πιθανή είναι και αναμενόμενη, επομένως η απροσδοκία h_t χαμηλή. Αντίθετα, όταν η τρέχουσα εμπειρία είναι μη πιθανή είναι και μη αναμενόμενη, επομένως η απροσδοκία h_t υψηλή. Κάθε γεννήτρια είναι ευαίσθητη στη δική της τιμή απροσδοκίας και μεταβάλλει την έντασή της

ανάλογα με αυτή. Η περίπτωση αυτή ονομάζεται recognition-h case. Υπάρχει και μία δεύτερη τιμή απροσδοκίας h_{t+1} , που αναφέρεται στην επόμενη στιγμή από τη στιγμή της αναγνώρισης, εκφράζει την απροσδοκία της προβλεπόμενης κατάστασης, και συνιστά την περίπτωση prediction-h case. Η τελευταία τιμή αποτελεί μία πιθανή εξήγηση γιατί αναπάντεχες προβλέψεις είναι πιο πιθανό να τραβήξουν την προσοχή σε σχέση με πιο κοινότερες προβλέψεις, ενώ υπερβολική χρήση ή ελαττώματα σε αυτό τον μηχανισμό μπορεί να ευθύνονται για φαινόμενα άγχους και νοητικής ασθένειας σε ανθρώπους.

Η συνολική πρόβλεψη ενός γεγονότος βασίζεται σε δύο αντικρουόμενες δυνάμεις, των οποίων οι επιδράσεις πολλαπλασιάζονται ως εξής:

Πιθανότητα του γεγονότος (αριθμός γεννητριών που το προβλέπουν) × Prediction-h (ένταση των γεννητριών με βάση την απροσδοκία του γεγονότος)

Ο υπολογισμός αυτός της πρόβλεψης εμφανίζει αρνητική προκατάληψη έναντι προβλέψεων που είναι πολύ πιθανές ή πολύ απίθανες.

7.4. Γεννήτριες και Οικουμενικός χώρος εργασίας

Ο μηχανισμός των γεννητριών λοιπόν, δημιουργεί τόσο περιεχόμενο αναγνώρισης (recognition-h) όσο και περιεχόμενο πρόβλεψης (prediction-h), το οποίο σχετίζεται με την πιθανότητα του περιεχομένου στο στατιστικό μοντέλο. Στη συνέχεια, οι γεννήτριες και των δύο ειδών περιεχομένου ανταγωνίζονται η μία την άλλη, όπως οι χορωδοί στην οπερατική σκηνή, για πρόσβαση στον Οικουμενικό Χώρο Εργασίας. Έτσι, για παράδειγμα, όταν ο οργανισμός βρίσκεται σε καταστάσεις ξεκάθαρου και παρόντα κινδύνου ή ωφέλειας, το περιεχόμενο αναγνώρισης υπερτερεί του περιεχομένου πρόβλεψης καθώς οι πιθανές προβλέψεις είναι πολύ περισσότερες από το πραγματικό αντιληπτικό περιεχόμενο και έτσι η μάζα της πιθανότητας τους είναι πολύ πιο απλωμένη, μην επιτρέποντάς τους να γίνουν ακουστές. Συγχρόνως, μπορούμε να υποθέσουμε ότι υπάρχει κάποιος τρόπος να γίνει γίνει διάκριση μεταξύ των δύο ειδών περιεχομένων που περνούν στην συνειδητή προσοχή (αλλιώς ένας οργανισμός δε θα μπορούσε να διακρίνει μεταξύ του κόσμου όπως τον αντιλαμβάνεται και όπως τον προβλέπει.

7.5. Η αίσθηση της αβεβαιότητας

Παραπάνω είδαμε τον τρόπο με τον οποίο χρησιμοποιείται στο IDyOT η ποσότητα πληροφορίας, δηλαδή ως μέτρο της απροσδοκίας αντιληπτικού περιεχομένου ή προβλεπτικού περιεχομένου. Το άλλο μέγεθος της θεωρίας πληροφορίας του Shannon, η εντροπία H , έχει νόημα μόνο στο προβλεπτικό πλαίσιο. Πιο συγκεκριμένα, μετρά τη διαφορά μεταξύ ορισμένων βέβαιων προβλεπόμενων εκβάσεων σε καταστάσεις έναντι αβέβαιων εκβάσεων, μοντελοποιώντας έτσι την υποκειμενική αίσθηση της αβεβαιότητας. Όσο υψηλότερη είναι η

τιμή της εντροπίας H σε προβλεπτικές γεννήτριες, τόσο χαμηλότερη είναι η συνολική ένταση τους, επομένως οι λιγότερο βέβαιες παραγόμενες προβλέψεις υποβαθμίζονται. Ο μηχανισμός αυτός αποτρέπει την υπερφόρτωση του GW, επιτρέποντας είσοδο μόνο σε στατιστικά ασφαλείς προβλέψεις. Όσο μεγαλύτερο είναι το εύρος των πιθανοτήτων, τόσο μεγαλύτερη είναι και η εντροπία και επομένως τόσο αυξημένη είναι και η αβεβαιότητα. Ένα παράδειγμα της λειτουργίας αυτού του μηχανισμού είναι ένας οργανισμός με λιγότερη εμπειρία, όπως ένα παιδί, το οποίο είναι πιο πιθανό να εισάγει απίθανες προβλέψεις στη συνείδησή του σε σχέση με κάποιον πιο έμπειρο οργανισμό, όπως έναν ενήλικα, που διαθέτει ένα καλύτερο στατιστικό μοντέλο για να κάνει τις προβλέψεις του.

7.6. Το κατώφλι εισαγωγής στον Οικουμενικό χώρο εργασίας
Η συνολική ένταση για κάθε γεννήτρια δίνεται από τον εξής τύπο:

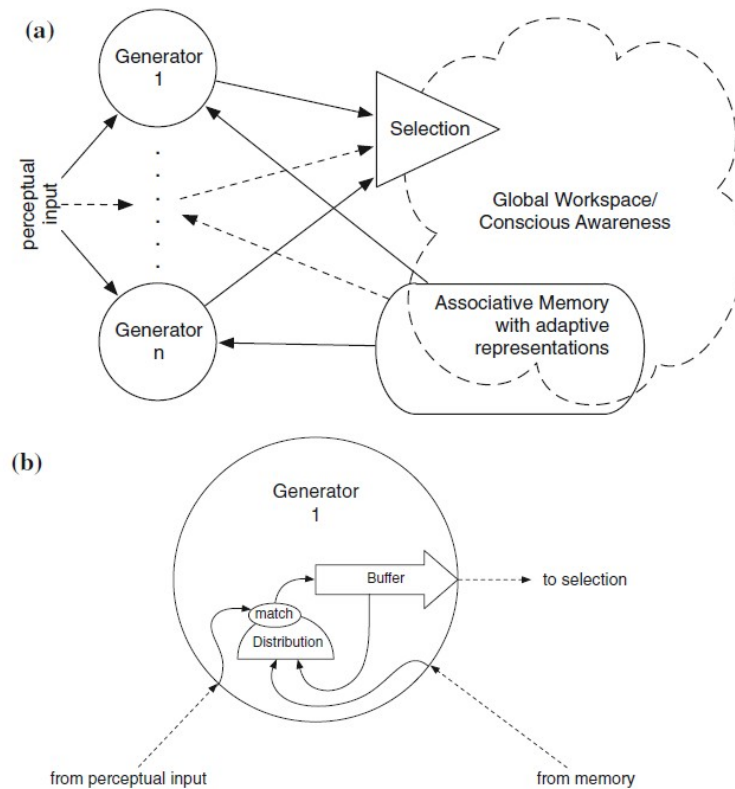
$T = \frac{p * h}{H}$, όπου T είναι συνολική ένταση, p είναι η πιθανότητα του ενδεχομένου, h είναι η απροσδοκία αναγνώρισης ή πρόβλεψης και H είναι η εντροπία της πρόβλεψης.

Η τιμή αυτή χρησιμοποιείται για να αποφασιστεί ποιο από τα πιθανά περιεχόμενα των γεννητριών, που προκύπτουν από αντιστοίχιση του αντιληπτικού περιεχομένου που λαμβάνει ένας οργανισμός με το στατιστικό μοντέλο που έχει στη μνήμη του, εισάγονται στον Οικουμενικό Χώρο Εργασίας. Το περιεχόμενο των γεννητριών αποθηκεύεται σε έναν buffer (ενδιάμεση μνήμη για προσωρινή αποθήκευση πληροφορίας), ο οποίος αδειάζει όταν η ένταση της γεννήτριας περνάει το κατώφλι για εισαγωγή στη συνειδητή προσοχή. Είναι αξιοσημείωτο ότι το περιεχόμενο των γεννητριών ενδέχεται να είναι και λανθασμένη αντίληψη η οποία είναι πιθανό να περάσει στον οικουμενικό χώρο εργασίας. Το ίδιο το κατώφλι για εισαγωγή στο GW δεν είναι αριθμητικά συγκεκριμένο, αλλά είναι αποτέλεσμα του συναγωνισμού μεταξύ των γεννητριών. Με αυτό τον μαθηματικό μηχανισμό, λύνεται το παράδοξο του κατωφλιού της θεωρίας του Baars στη θεωρία του IDyOT.

7.7. Σύνοψη

Συνολικά, είδαμε ότι νοητικές γεννήτριες αναλαμβάνουν το ρόλο των χρωδών του οπερατικού θεάτρου της δημιουργικής συνείδησης, και ότι οι γεννήτριες αυτές μαθαίνουν από την αντιληπτή εμπειρία. Τα μαθημένα σώματα γνώσης χρησιμοποιούνται για συνεχή παράλληλη πρόβλεψη πιθανών μερικών εκβάσεων της εκάστοτε τρέχουσας αντιληπτής κατάστασης του κόσμου. Επίσης χρησιμοποιούνται μεγέθη της θεωρίας πληροφορίας, και συγκεκριμένα η ποσότητα πληροφορίας (information content) και η εντροπία (entropy). Παρουσιάζεται αντιστάθμιση (trade-off) μεταξύ της ποσότητας πληροφορίας και της πιθανότητας αναγνώρισης ή πρόβλεψης μίας κατάστασης, που ρυθμίζει την συνολική νοητική προσπάθεια ενός οργανισμού, αφαιρώντας υπερβολικά πιθανά ή απίθανα περιεχόμενα. Η αίσθηση της

αβεβαιότητας μοντελοποιείται από την εντροπία, η οποία καταστέλλει αβέβαιες προβλέψεις και υποστηρίζει τις περισσότερο βέβαιες. Είναι ευλογοφανές να υποθέσουμε ότι μέσω της εξέλιξης αναπτύξαμε περίπλοκους στατιστικούς μηχανισμούς και μέτρα θεωρίας πληροφορίας για την διαρρύθμιση της προσοχής μας. Η δημιουργικότητα, όπως θα δούμε και παρακάτω, αποτελεί ένα παράπλευρο αποτέλεσμα αυτού του μηχανισμού.



Διάγραμμα 1 - (α) Η λειτουργία των γεννητριών σε σχέση με τον οικουμενικό χώρο εργασίας (Global Workspace) και τη μνήμη. (β) Η εσωτερική δομή μίας γεννήτριας (Wiggins & Forth, 2015)

8. Υψηλότερα χαρακτηριστικά της αρχιτεκτονικής

8.1. Αναπαράσταση και μνήμη

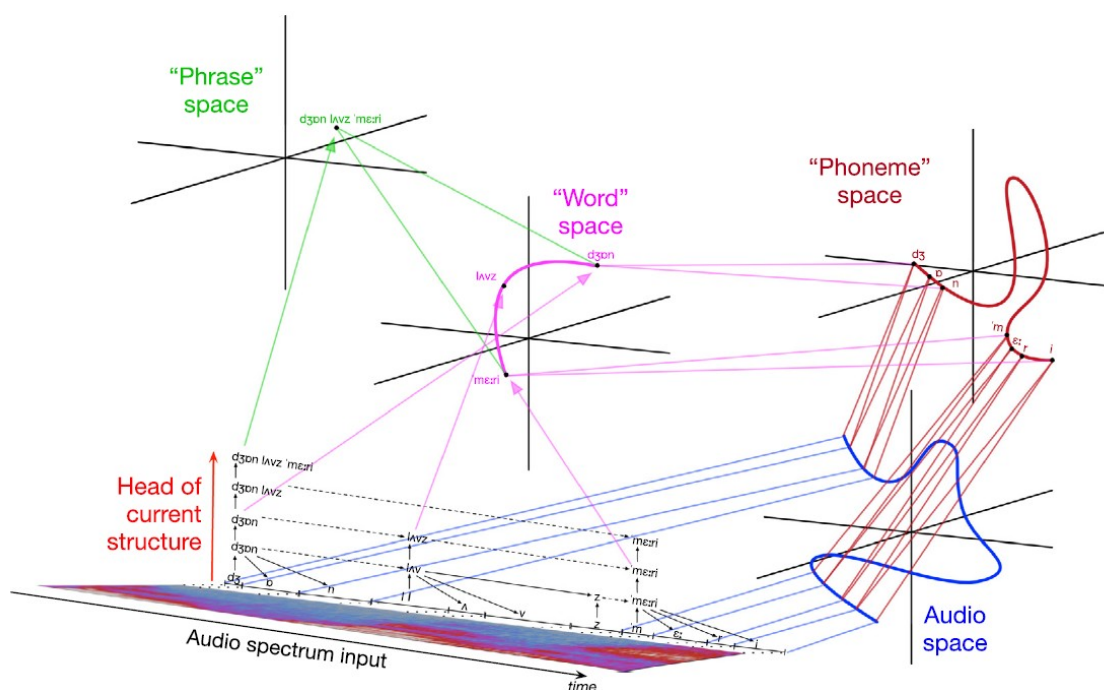
Όπως σε κάθε γνωσιακή αρχιτεκτονική, οι αναπαραστάσεις και η δομή της μνήμης της παίζουν καθοριστικό ρόλο στη λειτουργία της. Το μνημονικό μοντέλο του IDyOT βασίζεται σε πολυδιάστατες χρονικές ακολουθίες αντιληπτικού περιεχομένου. Καθεμία από αυτές τις διαστάσεις έχει τον δικό της εννοιολογικό χώρο, όπως για παράδειγμα το ακουστικό περιεχόμενο από το όργανο του Corti, το οποίο αναπαρίσταται ως χρονικό φάσμα. Χρησιμοποιώντας την εντροπία για να την τεμαχιοποίηση (ανακάλυψη ορίων) στις ακολουθίες υπό επεξεργασία, δημιουργούνται σύμβολα σε υψηλότερα επίπεδα στην ιεραρχία εννοιολογικών χώρων της μνήμης. Λέγεται ότι ένα υψηλότερο σύμβολο “υποτείνει” (subtends) την χαμηλότερη ακολουθία που αναπαριστά. Το χαμηλότερο από όλα τα επίπεδα, που στο παράδειγμα της ακοής θα ήταν η ακουστική αισθητηριακή μνήμη, πρέπει να απορριφθεί σε υπολογιστικές υλοποιήσεις (και μπορούμε να υποθέσουμε και στους ανθρώπους) καθώς περιέχει υπερβολική ποσότητα πληροφορίας για αποθήκευση. Έτσι η αφαιρετική δομή της συμβολικής ιεραρχίας εννοιολογικών χώρων οδηγεί σε πληροφοριακή αποδοτικότητα.

Η ανθρώπινη μνήμη μπορεί να διακριθεί σε δηλωτική (ή έκδηλη) και άδηλη μνήμη (Foerde & Poldrack, 2009), και η ίδια διάκριση γίνεται και στο IDyOT. Η δηλωτική μνήμη αποκτιέται μέσω συνειδητής προσοχής, ενώ η άδηλη χωρίς αυτή. Πρακτικά αυτό σημαίνει ότι οι πληροφορίες που περνούν από το GW αποτελούν τη δηλωτική μνήμη, ενώ αυτές που δεν περνούν από το GW αποτελούν την άδηλη μνήμη (όπως για παράδειγμα η στατιστική μάθηση της γλώσσας που παρουσιάζεται από παιδιά μικρής ηλικίας). Ωστόσο, η τελική αναπαράσταση της μνήμης είναι ίδια, ανεξάρτητα από τον τρόπο με τον οποίο αποκτήθηκε.

Μία ακόμη διάκριση της δηλωτικής μνήμης είναι αυτή που τη χωρίζει σε επεισοδιακή και σημασιολογική μνήμη (Tulving, 1972). Στο IDyOT η επεισοδιακή μνήμη συνίσταται από ένα πολυδιάστατο, ιεραρχικό στατιστικό μοντέλο, το οποίο καταγράφει αλληλουχίες χαμηλού επιπέδου από αντιληπτικά δεδομένα και χτίζει περαιτέρω επίπεδα από την τεμαχιοποίηση τους. Έτσι, η επεισοδιακή μνήμη αποτελεί μία κυριολεκτική αναπαράσταση της αντιληπτικής πληροφορίας και των συναγωγών που προέκυψαν από αυτή, παρόλο που τα χαμηλότερα επίπεδα της αναπαράστασης τελικά απορρίπτονται για λόγους χωρητικότητας. Η σημασιολογική μνήμη, αντιθέτως, αποτελεί μία ιεραρχία εννοιολογικών χώρων που σχετίζονται με τα σύμβολα που χρησιμοποιούνται για τον χαρακτηρισμό των δεδομένων αλληλουχιών και τα τεμάχιά τους. Καθένα από τα στοιχεία στους εννοιολογικούς χώρους αυτούς είναι ανεξάρτητο από τις αλληλουχίες στις οποίες εμφανίζεται και αποτελεί αναπαράσταση ενός συσχετισμού μεταξύ νοημάτων, δηλαδή λαμβάνει την έννοια του με βάση τα αντιλήμματα από τα οποία προέκυψε και τη σχέση του με άλλα στοιχεία. Το Διάγραμμα 1 παρουσιάζει ένα

απλουστευμένο παράδειγμα της λειτουργίας της δηλωτικής μνήμης κατά την επεξεργασία της φράσης σε φυσική γλώσσα.

Μία τελευταία διάκριση της μνήμης γίνεται μεταξύ βραχύχρονης και μακρόχρονης μνήμης (Shiffrin & Atkinson, 1968). Σε αντίθεση με μεγάλο μέρος των γνωσιακών επιστημόνων, η θεωρία του IDyOT υποστηρίζει την *in situ* προσέγγιση στην επεξεργασία δεδομένων, σύμφωνα με την οποία η διάκριση μεταξύ βραχύχρονης και μακρόχρονης μνήμης δεν είναι όσο ξεκάθαρη όσο σε έναν σύγχρονο ηλεκτρονικό υπολογιστή. Πιο συγκεκριμένα, σύμφωνα με την προσέγγιση αυτή, η διαφορά μεταξύ των δύο ειδών μνήμης είναι ότι η βραχύχρονη μνήμη περιλαμβάνει την μνημονική δομή που υποτείνεται από το υψηλότερο σύμβολο στη μνήμη, ενώ η μακρόχρονη μνήμη δεν περιλαμβάνει τη δομή αυτή μέχρι να ολοκληρωθεί, καθώς δεν έχει προσδοκίες για άμεση συνέχιση.

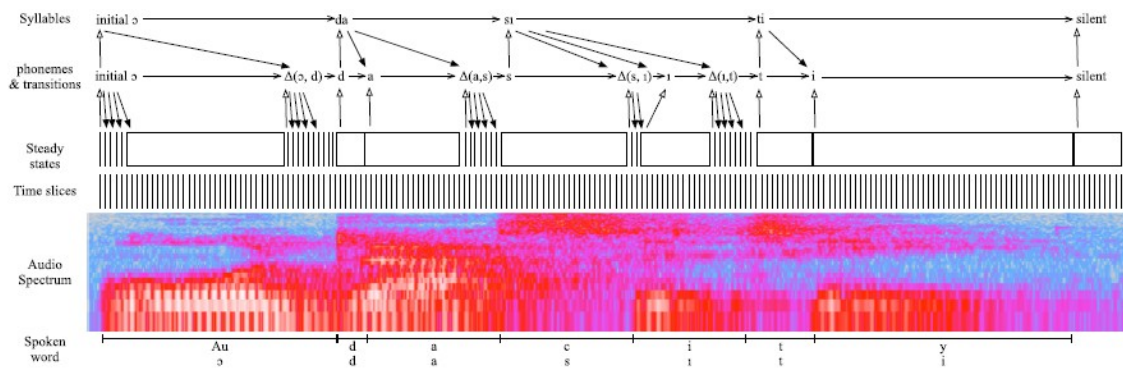


Διάγραμμα 2 - Η επεισοδιακή (μαύρη) και σημασιολογική (πολλαπλών χρωμάτων) μνήμη του IDyOT κατά την επεξεργασία της φράσης "John loves Mary" (Wiggins, 2020)

8.2. Φασματική φύση των εννοιολογικών χώρων

Το IDyOT χρησιμοποιεί τη θεωρία των εννοιολογικών χώρων του Gärdenfors για τη σημασιολογική του μνήμη, ωστόσο επεκτείνει μαθηματικά τη θεωρία από τη χαμηλών διαστάσεων φύση της, στους ισχυρούς διανυσματικούς χώρους Hilbert. Η αλλαγή αυτή φαίνεται να μην τηρεί την αρχή του ξυραφιού του Ockham, επομένως χρειάζεται δικαιολόγηση.

Αρχικά, η ταλαντευόμενη φύση των ομάδων νευρώνων του εγκεφάλου που σχετίζονται με την ακοή (Cariani, 1999) απαιτεί μαθηματικά εργαλεία ικανά να την αποδώσουν, όπως τους χώρους Hilbert, οι οποίοι μπορούν να αναπαραστήσουν αλληλουχίες ανεξάρτητα του χρόνου, σε στατική μορφή. Επιπλέον, είναι μέρος της λειτουργίας της μνήμης του IDyOT πως χρονικές τροχιές διαφορετικού μήκους, τόσο στην επεισοδιακή όσο και στην σημασιολογική μνήμη, θα χρειαστεί να συγκριθούν. Για παράδειγμα, για τη δημιουργία δομών ανάλογων με συντακτικές κατηγορίες ή την αναπαράσταση κλάσεων κίνησης και ήχου οι οποίες μπορεί να λάβουν χώρα σε διαφορετικές ταχύτητες. Με βάση τους Chella et al. (2015; 2004), το IDyOT χρησιμοποιεί μετασχηματισμό Fourier για την φασματική αναπαραστασιακή αντιστοίχιση αντιλημμάτων χαμηλού επιπέδου σε σύμβολα υψηλού επιπέδου. Ακόμη, ένα στοιχείο που υποστηρίζει τη χρήση των φασματικών αναπαραστάσεων είναι η αναπαράσταση Mel-Frequency Cepstral Coefficient (MFCC), που χρησιμοποιείται στην ανάλυση ομιλίας και λειτουργεί με παρόμοιο τρόπο με το IDyOT. Τέλος, η ίδια η θεωρία του Gärdenfors αποτελεί μία ειδική περίπτωση των ισχυρότερων μαθηματικών που περιγράφηκαν, και συγκεκριμένα χρησιμοποιεί μόνο πραγματικούς αριθμούς και απλά μετρικά απόστασης, επομένως η χρήση χώρων Hilbert αποτελεί συντηρητική επέκταση της θεωρίας του Gärdenfors.



Διάγραμμα 3 - Παράδειγμα λειτουργίας των φασματικών αναπαραστάσεων στο IDyOT με τεμαχιοποίηση της λέξης "audacity" (Wiggins, 2020)

8.3. Μνημονική παγίωση

Είδαμε λοιπόν πως η λειτουργεί η σταδιακή μάθηση του IDyOT, από το επίπεδο των γεννητριών στο επίπεδο των υψηλότερων εννοιολογικών χώρων της σημασιολογικής μνήμης, όπου βρίσκονται τα ευρύτερα σύμβολα του συστήματος. Αποτέλεσμα αυτού του τρόπου λειτουργίας είναι ότι κατά τη λειτουργία του συστήματος θα δημιουργηθούν τεμάχια τα οποία θα αποδειχτεί από επόμενα δεδομένα ότι είναι λανθασμένα, σε αντιστοιχία με τον τρόπο που οι

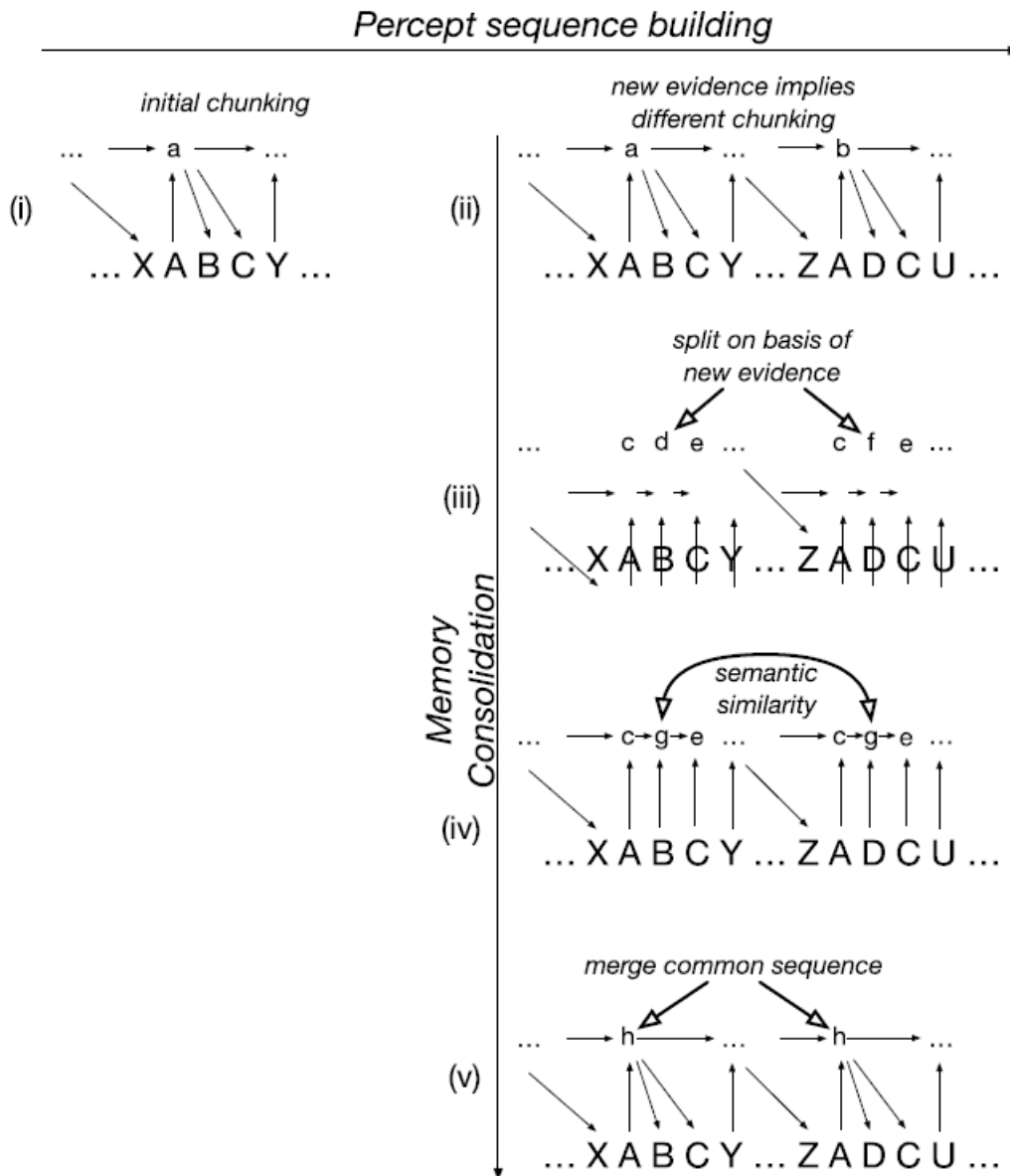
άνθρωποι αντιλαμβανόμαστε ότι κάποιες έννοιες που έχουμε στη μνήμη μας αποδεικνύονται λανθασμένες από μετέπειτα εμπειρίες. Ως αποτέλεσμα, είναι απαραίτητη μία περιοδική αναθεώρηση των συμβόλων και δομών στη μνήμη με σκοπό την μείωση του αναπαραστασιακού φορτίου, διαδικασία που στο IDyOT παίζει η μνημονική παγίωση (με έμπνευση από την αντίστοιχη ανθρώπινη διαδικασία, Wixted, 2004). Πιο αναλυτικά, η μνημονική παγίωση ταυτίζεται με την επαν-αναπαράσταση (re-representation) της μαθημένης πληροφορίας ώστε να βελτιωθεί τόσο η εκφραστική ισχύς της αναπαράστασης όσο και η προβλεπτική ικανότητα του μοντέλου που χρησιμοποιείται για να την εκφράσει. Για να είναι αποδοτική, η μνημονική παγίωση λαμβάνει χώρα όταν η μνήμη δε λαμβάνει νέα πληροφορία, σε αντιστοιχία με τον ύπνο στους ανθρώπους.

Στόχος της μνημονικής παγίωσης είναι η εύρεση της βέλτιστης κατηγοριοποίησης, μέσω της ελαχιστοποίησης της διασταυρωμένης εντροπίας (cross-entropy) τόσο στο αντιληπτικό όσο και στο σειριακό μοντέλο. Συγκεκριμένα εξερευνώνται εναλλακτικές αναπαραστάσεις με βάση την συνεμφάνιση και την σειριακή στατιστική. Παρουσιάζεται τάση για λιγότερα σύμβολα, καθώς περισσότερα σύμβολα αυξάνουν την εντροπία, ωστόσο αυτή αντισταθμίζεται από την προϋπόθεση της προβλεπτικής ακρίβειας, η οποία χάνεται όταν συγχωνεύονται υπερβολικά πολλά σύμβολα. Η παγίωση συνίσταται σε τρία διαφορετικά αλλά διασυνδεδεμένα φαινόμενα:

- 1) Ορισμοί συμβόλων: Περιοχές σε συμβολικούς εννοιολογικούς χώρους μπορεί να προσαρμοστούν, να συγχωνευτούν ή να χωριστούν, όπως οι δομές στη διαδικασία μάθησης της γλώσσας σύμφωνα με τον Kirby (2000).
- 2) Εννοιολογικοί χώροι: Ενδέχεται να χρειάζεται να αλλάξει η γεωμετρία των ίδιων των εννοιολογικών χώρων ώστε να επαν-αναπαραστήσουν νέα δεδομένα αποδοτικά. Αυτό επιτυγχάνεται με αναζήτηση μεταξύ μαθηματικών τιμών διαφόρων συναρτήσεων στο χώρο, κάποιες από τις οποίες είναι και γνωσιακά εύλογες, δηλαδή θα μπορούσαν να αποτελούν μέρος της λειτουργίας των νευρώνων του εγκεφάλου (Kemp et al., 2004).
- 3) Αλλαγές ιεραρχίας: Ως αποτέλεσμα των παραπάνω αλλαγών, ενδέχεται να μετατραπεί και η ιεραρχία της μνήμης του IDyOT με την προσθήκη νέων αναπαραστάσεων (για παράδειγμα στη μουσική, με προσθήκη της αναπαράστασης του τονικού διαστήματος από την αναπαράσταση του τόνου), επηρεάζοντας και τις σχετικές προβλέψεις.

Το διάγραμμα 4 παρουσιάζει τη διαδικασία της παγίωσης, η οποία σχετίζεται με τη διαδικασία επιλογής viewpoint η οποία εφαρμόστηκε σε προγόνους του συστήματος IDyOT, αλλά έχει επεκταθεί ώστε να δημιουργεί νέες, μη προγραμματισμένες από ανθρώπους αναπαραστάσεις για την βελτίωση των προβλέψεων και συμπίεση των αποθηκευμένων δεδομένων. Ουσιαστικά, στη μνημονική παγίωση το IDyOT παίρνει κάθε μαθημένο εννοιολογικό χώρο, ξεκινώντας από τον χαμηλότερο και προχωρώντας προς τον υψηλότερο και υπολογίζει τα αποτελέσματα από πιθανές αλλαγές των σχετικών τεμαχίων. Σκοπός είναι η ελαχιστοποίηση του cross-entropy μεταξύ του μοντέλου και γνωστών δεδομένων. Ο χώρος αναζήτησης (search space)

εναλλακτικών δομών για επαν-αναπαράσταση συμβόλων μεγαλώνει εκθετικά με τον αριθμό τους, επομένως απαιτείται η χρήση ευρετικών μεθόδων και όριο στο ποσό αλλαγής που επιτρέπεται. Στο πρακτικό επίπεδο των γεωμετρικών εννοιολογικών χώρων, οι σημασιολογικές περιοχές που αντιστοιχούν σε ένα επεισοδιακό σύμβολο πρέπει να προσαρμόζονται έτσι ώστε να συμπεριλαμβάνουν τα σημεία που βρίσκονται πιο κοντά στο κέντρο τους απ' ότι στο κέντρο άλλων περιοχών. Αυτή η αρχή οδηγεί γενικά προς την κυρτότητα των περιοχών, ακολουθώντας τη θεωρία του Gärdenfors, αλλά μόνο αν μειώνεται η συνολική διασταυρωμένη εντροπία του μοντέλου (cross-entropy). Η διαδικασία λειτουργεί σαν διόρθωση λαθών κατηγοριοποίησης, και μπορεί να διαγράψει σύμβολα και περιοχές που δεν είναι πλέον χρήσιμες. Η επίδραση της μνημονικής παγίωσης μπορεί να είναι μεγάλη και να εξαπλωθεί σε όλη τη μνήμη του μοντέλου, επομένως είναι χρήσιμο να υπάρχει όριο στην εφαρμογή της. Το όριο αυτό τίθεται από την εξισορρόπηση μεταξύ μεγέθους του αλφαβήτου του μοντέλου και της προβλεπτικής του ικανότητας, που επηρεάζουν άμεσα την εντροπία.



Διάγραμμα 4 - Παράδειγμα μνημονικής παγίωσης κατά την οποία το αρχικό τεμάχιο συγχωνεύεται με ένα άλλο που εμφανίζεται αργότερα, σχηματίζοντας έτσι ένα σύμβολο υψηλότερου επιπέδου (Wiggins, 2020)

8.4. Πρόβλεψη

Όπως ήδη είδαμε, οι γεννήτριες του IDyOT λαμβάνουν δείγματα από το στατιστικό μοντέλο αλληλουχιών της μνήμης του IDyOT και εκτελούν Μαρκοβιανή πρόβλεψη, λαμβάνοντας υπ' όψη τα συμφραζόμενα. Η πρόβλεψη γίνεται με τη χρήση ενός Μαρκοβιανού πίνακα μετάβασης

πρώτης τάξης για κάθε επίπεδο της ιεραρχίας της μνήμης και σε τρεις κατευθύνσεις: εμπρός, κατά μήκος της αλληλουχίας του εκάστοτε επιπέδου, προς τα πάνω, στο πιο αφηρημένο ανώτερο επίπεδο και προς τα κάτω, στο λιγότερο αφηρημένο κατώτερο επίπεδο. Η τρέχουσα υπόθεση στη θεωρία του IDyOT είναι πως τα προβλεπτικά μοντέλα αυτά είναι πρώτης τάξης, ενώ οι αλληλεπιδράσεις μεταξύ των επιπέδων προσφέρουν πιο αναλυτικά στατιστικά δεδομένα που υπολογίζονται από τον αλγόριθμο PPM*, ο οποίος χρησιμοποιείται και στο σύστημα IDyOM (Cleary & Witten, 1984). Επιπροσθέτως, όπως και στο σύστημα πολλαπλών οπτικών γωνιών (multiple viewpoints) και στο IDyOM, προβλέψεις μοντέλων διαφορετικών χαρακτηριστικών συνδυάζονται, ανάλογα με το βαθμό στατιστικής συσχέτισης μεταξύ των χαρακτηριστικών αυτών. Ομοίως, προβλέψεις ενός επιπέδου είναι πιθανό να επηρεάσουν προβλέψεις άλλων επιπέδων (Wiggins & Sanjekdar, 2019). Το αντιληπτικό περιεχόμενο που λαμβάνει ο οργανισμός από το περιβάλλον του αντιστοιχίζεται σε ένα χώρο ο οποίος τροποποιείται σύμφωνα με την τρέχουσα κατανομή των πιθανών εκβάσεων, έτσι ώστε πιο αναμενόμενα σύμβολα να είναι πιο ανεκτικά σε αποκλίσεις.

8.5. Γνωσιακός κύκλος

Όπως αναφέρθηκε στην ενότητα για τις γνωσιακές αρχιτεκτονικές, αυτές διαθέτουν μία σειρά διαδικασιών που κινητοποιούν και ορίζουν τη συνολική τους λειτουργία. Η σειρά αυτή λέγεται γνωσιακός κύκλος και στη θεωρία του IDyOT συνδέεται στενά με τον οικουμενικό χώρο εργασίας (Global Workspace) του Baars. Πιο αναλυτικά, οι γεννήτριες του IDyOT λαμβάνουν δείγματα από τη στατιστική μνήμη, συγχρονιζόμενες από το αντιληπτικό περιεχόμενο (όταν αυτό δεν παρέχεται από το περιβάλλον, τότε οι γεννήτριες δειγματοληπτούν ελεύθερα σε μία διαδικασία που, όπως θα δούμε παρακάτω, ίσως να αποτελεί εξήγηση της μη-συνειδητής δημιουργικότητας). Στη συνέχεια, το αντιληπτικό περιεχόμενο που λαμβάνει το σύστημα από το περιβάλλον αντιστοιχίζεται με τις προβλέψεις των γεννητριών και όταν είτε αντιστοιχίες (recognition-h), είτε προβλέψεις από αντιστοιχίες (prediction-h) έχουν υψηλό περιεχόμενο πληροφορίας, τότε οι γεννήτριες που τις περιλαμβάνουν αδειάζουν το περιεχόμενό τους στον οικουμενικό χώρο εργασίας. Όταν καινούργια πληροφορία εισέρχεται στο GW η προηγούμενη πληροφορία μετατοπίζεται στη μνήμη, όπου ενημερώνει το στατιστικό μνημονικό μοντέλο του συστήματος. Γεννήτριες των οποίων οι buffers δεν καταφέρνουν να περάσουν το κατώφλι εισαγωγής στη συνειδητή προσοχή ξεχνιούνται (λεπτομέρειες σχετικά με τη διάρκεια και άλλες πιθανές αιτίες του φαινομένου είναι υπό μελέτη). Μπορούμε να υποθέσουμε την ύπαρξη σταθερών γεννητριών για κάθε αισθητική τροπικότητα (sensory modality, δηλαδή όραση, ακοή, κτλ). Με βάση αυτή την υπόθεση, μπορούμε να θεωρήσουμε πως ξαφνική αύξηση στην ποσότητα πληροφορίας σε κάποια αίσθηση συνεπάγεται την ενεργοποίηση περισσότερων γεννητριών, προσφέροντας εξήγηση στο γνωστικό φαινόμενο του “τραντάγματος” (cognitive jolt) όταν η προσοχή ανακατανέμεται απότομα.

9. Γνωσιακή υποστήριξη και εφαρμογές του IDyOT

9.1. Γνωσιακά στοιχεία για την εγκυρότητα της θεωρίας

Παραπάνω αναφερθήκαμε στη θεωρία του Meyer σχετικά με τη σημασία των πιθανοτήτων για το μουσικό νόημα και συναίσθημα. Πιο συγκεκριμένα, η θεωρία αυτή προτάσσει πως οι προσδοκίες που παράγονται από μουσικές δομές συνιστούν ένα σύνθετο σύστημα πιθανοτήτων. Η υπόθεση αυτή υιοθετείται τόσο από την εργασία των Conklin & Witten, από το σύστημα IDyOM του Pearce αλλά και από το IDyOM. Μάλιστα, όλες οι θεωρίες αυτές χρησιμοποιούν τιμές της θεωρίας πληροφορίας για να τη δημιουργία και επεξεργασία του συστήματος μουσικών πιθανοτήτων που υλοποιούν. Παρακάτω, παρατίθενται πειραματικά στοιχεία που επιβεβαιώνουν την εγκυρότητα της θεωρίας σε γνωσιακό επίπεδο.

Καταρχάς, οι άνθρωποι είναι ικανοί να αναφερθούν στην συνειδητή επίγνωση μίας προσδοκίας, καθώς οι νότες υψηλής ή χαμηλής πιθανότητας γίνονται αντιληπτές ως τέτοιες. Επιπλέον, πειράματα με χρόνους αντίδρασης υποδεικνύουν ότι νότες υψηλής πιθανότητας γίνονται προϊόν ταχύτερης επεξεργασίας απ' ότι νότες χαμηλής πιθανότητας. Ακόμη, έχει εντοπιστεί σταθερή αύξηση στον συγχρονισμό ζώνης βήτα (beta-band synchronization) όταν το αντίλημμα που παρουσιάζεται είναι μη αναμενόμενο. Συνδέοντας τα πειραματικά ευρήματα αυτά με την αναλογία του θεάτρου, θα μπορούσαμε να εξηγήσουμε τα παραπάνω φαινόμενα χρησιμοποιώντας την πιθανότητα κάθε ενδεχομένου και την τιμή αναγνώρισης h . Αναλυτικότερα, σύμφωνα με το μηχανισμό των γεννητριών του IDyOT, η εισαγωγή χορωδών που σχετίζονται με πολύ πιθανές προβλέψεις στο φωτισμένο μέρος της σκηνής (συνείδηση) είναι λιγότερο πιθανή. Ο συγχρονισμός αποτελεί πιθανό μηχανισμό ή αποτέλεσμα (δεν είναι ακόμη ξεκάθαρο ποιο απ' τα δύο) της αυξημένης έντασης των χορωδών με υψηλότερο recognition-h.

Όσον αφορά την εντροπία H , έχει επιδειχθεί ότι έχει μία μικρή αλλά σταθερή επίδραση στην μουσική εκτέλεση. Συγκεκριμένα, έχει διαπιστωθεί πειραματικά πως υπάρχει μεγαλύτερη παρέκκλιση μεταξύ του εκφραστικού χρονισμού διαφορετικών εκτελεστών, σε σημεία υψηλής εντροπίας δοσμένων κομματιών (Gingras et al., 2016). Επομένως, γίνεται φανερό η τουλάχιστον υποσυνείδητη επίγνωση της εντροπίας από τους μουσικούς.

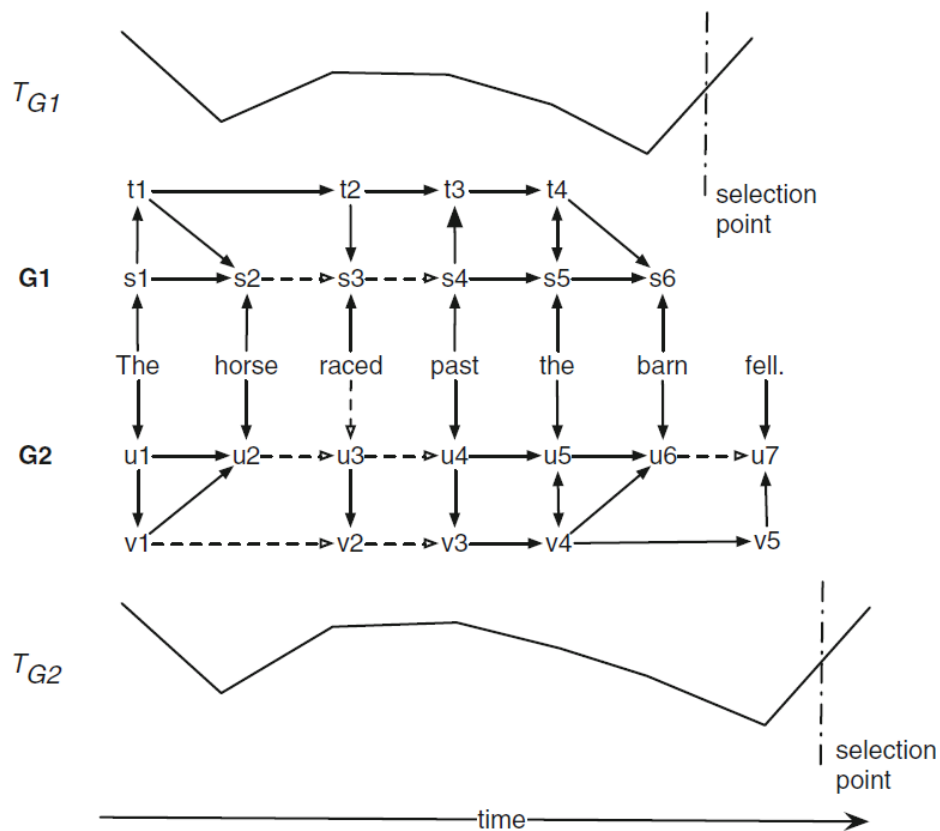
9.2. Τεμαχιοποίηση γλώσσας και μουσικής

Είναι γνωστό στην κοινότητα της γνωσιακής επιστήμης ότι αλληλουχίες από μουσικά και γλωσσικά σύμβολα μπορούν να τμηματοποιηθούν (segmentation), έτσι ώστε να χωριστούν σε αντιληπτικά τεμάχια (perceptual chunks), με χρήση μεγεθών της θεωρίας πληροφορίας όπως η ποσότητα πληροφορίας και η εντροπία (Gobet et al., 2001). Η εξελικτική συλλογιστική πίσω

από τη λειτουργία αυτή είναι ξεκάθαρη: η συμπύεση της πληροφορίας σε τεμάχια είναι πιθανότερο να προσδώσει αποδοτικότερες αναπαραστάσεις και επομένως καλύτερη προβλεπτική ικανότητα σε έναν οργανισμό. Ωστόσο, οι υπάρχουσες μέθοδοι τμηματοποίησης δεν εντάσσονται σε μία γενικότερη γνωσιακή αρχιτεκτονική και χρησιμοποιούν κυρίως περιγραφικούς κανόνες για να εντοπίσουν όρια τεμαχίων. Αντιθέτως, στο IDyOT η τεμαχιοποίηση αποτελεί αναδυόμενη ιδιότητα της λειτουργίας του συστήματος. Πιο συγκεκριμένα, οι ακολουθίες που περνούν στους buffers των γεννητριών μπορούν να εντοπίσουν τα τέλη τεμαχίων σύμφωνα με την τιμή της εντροπίας των προβλέψεών τους σε σχέση με τις υπόλοιπες γεννήτριες.

Ο μηχανισμός αυτός του IDyOT μοντελοποιεί και φαινόμενα σχετικά με την ασάφεια των ορίων στη γλώσσα, για παράδειγμα το “garden path” effect, κατά το οποίο προτάσεις όπως η “The horse raced past the barn fell”, αρχικά αναλύεται ως ολοκληρωμένη μετά τη λέξη “barn”, και η εμφάνιση της τελευταίας λέξης “fell” προκαλεί έκπληξη. Το σύστημα είναι ικανό να προβλέψει συγχρόνως και συνεκτικά σε πολλαπλά επίπεδα αφαίρεσης, διαδικασία που θυμίζει την πολλαπλή επεξεργασία προτάσεων της γλωσσολογίας, αλλά με την προσθήκη πιθανολογικής βελτιστοποίησης ώστε να αφαιρούνται σχετικά απίθανες ερμηνείες. Επομένως, και οι δύο ερμηνείες της παραπάνω garden path πρότασης αναλύονται συγχρόνως, αλλά η ορθή είναι σχετικά απίθανη και έτσι η αποπλανητική λανθασμένη ερμηνεία λαμβάνει προβάδισμα στη συνειδητή εμπειρία λόγω του εσφαλμένα προβλεπόμενου τέλους της πρότασης μετά τη λέξη “barn” το οποίο αποδίδει μία πρόβλεψη υψηλής εντροπίας. Όταν η εναλλακτική ανάγνωση περνά στο GW, ως αποτέλεσμα της εισόδου της τελευταίας λέξης “fell”, είναι απαραίτητη η αναθεώρηση μεταξύ της προηγούμενης και τωρινής ερμηνείας, και είναι αυτή η απότομη αλλαγή που οδηγεί στο γνωστικό τράνταγμα (cognitive jolt) του garden path effect.

Αυτή η διαδικασία ανάλυσης μέσω πρόβλεψης απεικονίζεται στο Διάγραμμα 5, όπου μία πρόταση αναλύεται ταυτόχρονα από δύο γεννήτριες (παρόλο που στην πραγματική λειτουργία του συστήματος θα υπήρχαν πολύ περισσότερες). Δημιουργούνται προβλέψεις, στις οποίες δίνονται αυθαίρετα ονόματα, αλλά οι οποίες συντίθενται ανωφερώς και δεν επιβάλλονται ανωφερώς. Η συνάρτηση T_{Gi} αποτελεί μέτρο της θεωρίας πληροφορίας το οποίο χρησιμοποιείται για να συγκρίνει και να επιλέξει το περιεχόμενο για εισαγωγή στο GW. Η υπόθεση, λοιπόν, του IDyOT είναι ότι το σύστημα θα πρέπει να μπορεί να τμηματοποιεί εξίσου πετυχημένα με κατωφερή μοντέλα τόσο σε μουσική όσο και σε γλώσσα, καθώς και να προσομοιάζει το garden path effect χωρίς επιπλέον περιγραφικούς κανόνες.



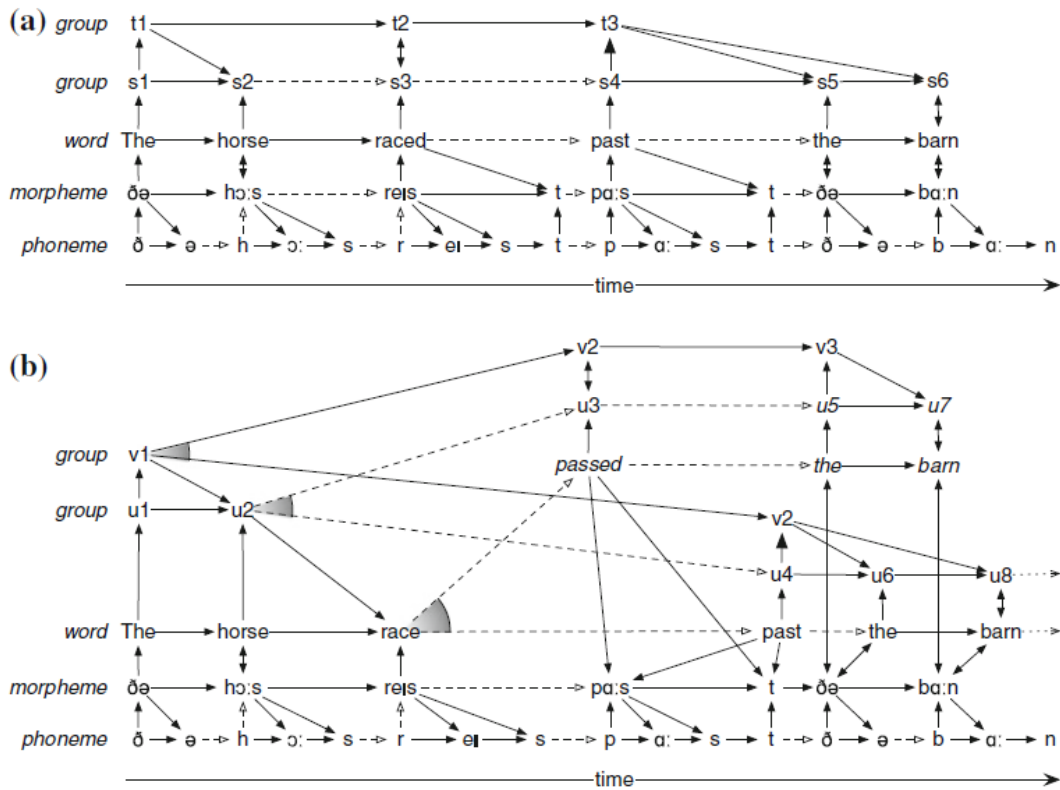
Διάγραμμα 5 - Δύο ταυτόχρονες προβλέψεις αλληλουχιών του συστήματος IDyOT για την ίδια πρόταση, με την οποία επιδεικνύεται το *garden path effect*. (Wiggins & Forth, 2015)

9.3. Λεξιλογική αμφισημία

Εφόσον το IDyOT μπορεί να εκτελέσει εργασίες τμηματοποίησης, θα έπρεπε να είναι και σε θέση να αναλύει (parse). Ωστόσο, η έννοια της ανάλυσης στη θεωρία του IDyOT διαφέρει από την έννοια που λαμβάνει στη σχολή γλωσσολογίας του Chomsky. Στην τελευταία υπάρχει η έννοια της σωστής ανάλυσης, σε αντίθεση με το IDyOT όπου υπάρχει η έννοια του βαθμού αντιστοίχισης με την παρατηρούμενη πιθανότητα. Επομένως, το σύστημα δεν εξαρτάται από ολόκληρες προτάσεις, ή άλλες προκαθορισμένες συντακτικές μονάδες, αλλά η επαγωγική του

μέθοδος λειτουργεί ανωφερώς, συνδυάζει και προβλέπει από τα δεδομένα που του δίνονται και τις στατιστικές δομές που έχει μάθει. Όταν το αποτέλεσμα της διαδικασίας αυτής έχει αρκετή ποσότητα πληροφορίας περνά στον οικουμενικό χώρο εργασίας. Όλη αυτή η διεργασία, ξεκινώντας από το επίπεδο των μορφημάτων (αν και το σύστημα κανονικά ξεκινάει από raw audio) απεικονίζεται στο Διάγραμμα 6. Η δεύτερη εικόνα μάλιστα παρουσιάζει το IDyOT ενώ αντιμετωπίζει το φαινόμενο της λεξιλογικής αμφισημίας που προκύπτει από υπερφόρτωση ενός ηχητικού δείγματος που μπορεί να αντιστοιχεί σε δύο διαφορετικές λέξεις “passed” και “past”. Είναι αναμενόμενο πως η αναπαράσταση του τονισμού της ομιλίας που περιλαμβάνει το IDyOT θα παίξει σημαντικό ρόλο σε μελλοντικές μελέτες.

Για την παραγωγή σημασιολογικών αναπαραστάσεων από τη αναλυτική διαδικασία αυτή, είναι απαραίτητο να προστεθούν στην εκπαιδευτική διαδικασία αναπαραστάσεις σημασιολογικών ερμηνειών, σε αντιστοιχία με τους Eshghi et al. (Eshghi et al., 2013). Έτσι, η θεμελίωση του νοήματος στο IDyOT πετυχαίνεται μέσω της σύνδεσης μεταξύ σημασιολογικών συμβόλων και των αντίστοιχων επιφανειακών μορφών. Το πλαίσιο των πολλαπλών οπτικών γωνιών (multiple viewpoints) αρκεί για την αναπαράσταση των συνδέσεων αυτών και επομένως δεν απαιτείται χρήση επιπλέον μηχανισμού στο IDyOT.

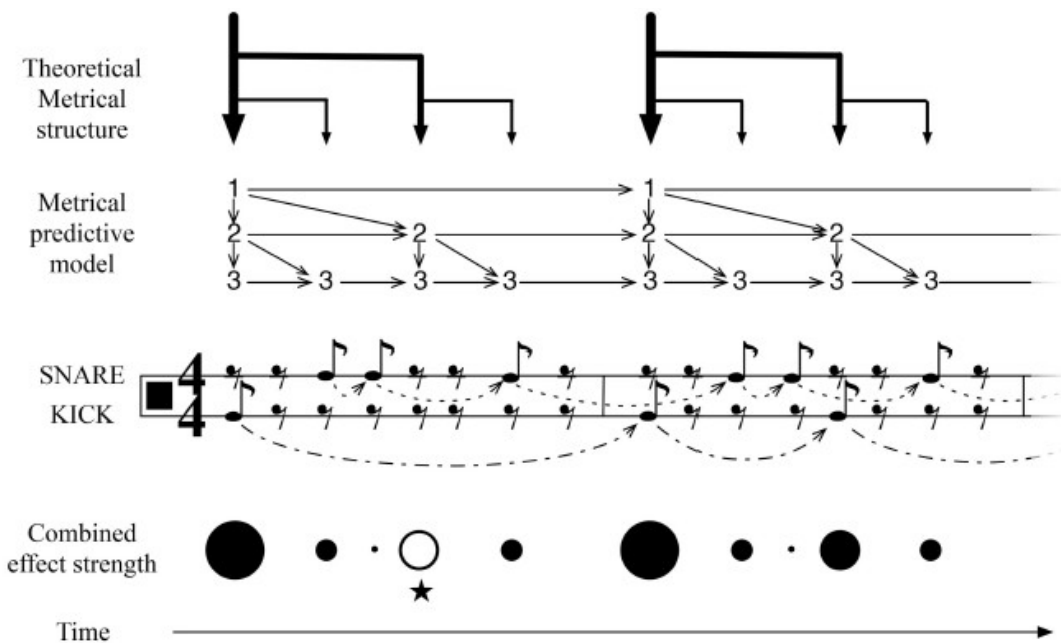


Διάγραμμα 6 – (α) Παράδειγμα ανάλυσης πρότασης από το IDyOT, όπου διαφαίνονται τα διαφορετικά επίπεδα συμβόλων και οι προβλέψεις του συστήματος. (β) Παράδειγμα παράλληλης ερμηνείας της ίδιας πρότασης, που παρουσιάζει λεξιλογική αμφισημία. Οι σκιασμένες γωνίες δηλώνουν διαφορετικές αναλύσεις. (Wiggins & Forth, 2015)

9.4. Συγχρονισμός (entrainment)

Ένας ακόμη τομέας έρευνας της γνωσιακής μουσικολογίας με τον οποίο σχετίζεται η θεωρία του IDyOT είναι ο συγχρονισμός ή entrainment, ο οποίος προσεγγίζεται από δύο οπτικές. Αρχικά, ορίζονται εννοιολογικοί χώροι για την αναπαράσταση του μουσικού μέτρου (J. C. Forth, 2012), που επιτρέπουν τον ακριβή προσδιορισμό μετρικών δομών, δηλαδή θεωρητικών μοτίβων συγχρονισμού που καθοδηγούν την προσοχή κατά την ακρόαση μουσικής (London, 2004). Η προσέγγιση αυτή μπορεί να θεωρηθεί ως μία κατωφερής διασαφήνιση της θεωρητικής έννοιας του μέτρου. Η δεύτερη προοπτική της θεωρίας του IDyOT για το συγχρονισμό είναι ανωφερής, και βασίζεται στον κατώτερο μηχανισμό των γεννητριών του συστήματος, ο οποίος μπορεί να μάθει ιεραρχικές αναπαραστάσεις του μουσικού ρυθμού μέσω έκθεσης στην στατιστική κανονικότητα της μουσικής και γενικότερης καθημερινής αντιληπτικής εμπειρίας.

Η βασική θέση που υποστηρίζεται είναι ότι η μουσική ακρόαση συντονίζεται από μοτίβα προσοχής, τα οποία προκύπτουν τόσο από εσωτερική παραγωγή όσο και επαγωγή από αντιληπτική πληροφορία, και επιπλέον ότι η ίδια διαδικασία ρυθμίζει την χρονική λειτουργία της νόησης εν γένει, που υποστηρίζεται από στοιχεία στη φυσική γλώσσα (J. Forth et al., 2016). Το IDyOT μπορεί να υπολογίσει προβλεπόμενες κατανομές σε πολλαπλά επίπεδα λεπτομέρειας σε σχέση με την αντιληπτική επιφάνεια, και γίνεται η υπόθεση ότι η απαίτηση για σύμπτωση των προβλέψεων μεταξύ διαφορετικών επιπέδων είναι αυτή που οδηγεί στην ανθρώπινη τάση για κυκλικά (αν και όχι απαραίτητα ισόχρονα) μέτρα στην μουσική και την ποίηση. Το σύστημα επομένως λειτουργεί με βάση μία κβαντισμένη εκδοχή του πραγματικού χρόνου, που του δίνει τη δυνατότητα να κάνει χρονικές προβλέψεις και να συγχρονίζεται με εξωτερικά ερεθίσματα.



Διάγραμμα 7 – Η δομή ενός απλού rock ρυθμού. Σύγκριση μεταξύ του μοντέλου του IDyOT και του θεωρητικού μοντέλου του μέτρου. Ο αστερίσκος δηλώνει την απουσία ενός προβλεπόμενου χτύπου.

9.5. Δημιουργικότητα

Η εξήγηση της μη συνειδητής δημιουργικότητας ήταν μία από τις αρχικές κινητοποιήσεις για την ανάπτυξη της θεωρίας του IDyOT, η οποία έγινε σταδιακά όλο και ευρύτερη και πολύπλευρη, ενσωματώνοντας όλα τα θέματα που παρουσιάστηκαν παραπάνω, όπως η

συνειδητή προσοχή, η αντίληψη σειριακών δεδομένων κ.α. Η δημιουργικότητα λοιπόν, σύμφωνα με τη θεωρία του IDyOT, είναι στενά συνδεδεμένη με την προβλεπτική ανάλυση. Η προβλεπτική ανάλυση αυτή διενεργείται από το μηχανισμό του συστήματος, ο οποίος επιχειρεί συνεχώς να ταιριάζει τις προβλέψεις του με το αντιληπτικό περιεχόμενο που συναντάται (μία ευέλικτη διαδικασία που περιλαμβάνει την πιθανότητα αλλαγών στη σημασιολογία). Τα οφέλη αυτής της προσέγγισης είναι η αποδοτικότητα και ευρωστία, που διαφαίνονται στην ικανότητα του συστήματος να προβλέπει σε πραγματικό χρόνο την πληροφορία που μεταφέρει ένας ομιλητής, να ανακατασκευάζει δυσνόητα μέρη της ροής δεδομένων, να συμπεραίνει τις προθέσεις ενός ομιλητή και πιθανότατα να αναμένει ολόκληρη τη συζήτηση, που καθιστούν τη συμπεριφορά του IDyOT μία ρεαλιστική προσομοίωση της ανθρώπινης συμπεριφοράς.

Όταν, ωστόσο, δεν υπάρχει αντιληπτικό περιεχόμενο ώστε να γίνει αντιστοίχιση με σύμβολα στη μνήμη του συστήματος, η πρόβλεψη συνεχίζεται, κινούμενη από την ίδια τη μνήμη. Έτσι οι γεννήτριες λαμβάνουν περιεχόμενο από την ήδη μαθημένη πληροφορία και κάνουν πιθανώς νέες προβλέψεις, ενώ ο μηχανισμός τμηματοποίησης μπορεί να χρησιμοποιηθεί για την επιλογή δομών από όλα τα επίπεδα αναπαράστασης. Οι ενδεχομένως νέες δομές αυτές περνούν στη συνειδητή προσοχή, μέσω του μηχανισμού του κατωφλιού, απ' όπου προστίθενται στη μνήμη και γίνονται διαθέσιμες για μελλοντική πρόβλεψη. Η (μη συνειδητή) δημιουργικότητα δηλαδή συνίσταται στην ίδια αντιληπτική διαδικασία του IDyOT, αλλά ωθείται από τις δομές της μνήμης και όχι αντιληπτικό περιεχόμενο, και όταν καινούργιες ή μη δομές περάσουν το πληροφοριακό κατώφλι περνούν στη συνείδηση, δημιουργώντας το 'Aha moment'. Σε σχέση με τους εννοιολογικούς χώρους, η δημιουργικότητα ταυτίζεται με την εξερεύνηση μίας νέας αλληλουχίας ή ενός νέου σημείου σε έναν χώρο οποιουδήποτε επιπέδου, τα οποία περιέχουν αρκετή πληροφορία για να γίνουν συνειδητά. Ο Wiggins προτείνει τη μελέτη της μικρής, καθημερινής δημιουργικότητας στα πλαίσια του συστήματος, που μπορεί να αποτελέσουν πάτημα για τη μελέτη της καλλιτεχνικής δημιουργικότητας (Wiggins, 2020).

Γ' ΜΕΡΟΣ: ΚΙΝΗΤΟΠΟΙΗΣΗ ΚΑΙ ΣΥΝΑΙΣΘΗΜΑ ΣΤΗΝ ΑΡΧΙΤΕΚΤΟΝΙΚΗ MICROPSI ΚΑΙ ΣΥΝΔΕΣΗ ΜΕ ΤΟ ΙΔΥΟΤ

10. Η θεωρία PSI και η γνωσιακή αρχιτεκτονική MicroPSI

10.1. MicroPSI

Η γνωσιακή αρχιτεκτονική Micro PSI του Bach επιχειρεί να αποτελέσει ένα πλαίσιο για την κατανόηση στον τομέας της Γενικής Τεχνητής Νοημοσύνης (Artificial General Intelligence – AGI) και βασίζεται σε μία δομή δημιουργίας και προσομοίωσης γνωσιακών πρακτόρων (2015). Η αρχιτεκτονική βασίζεται σε νευρο-συμβολικές αναπαραστάσεις, ενώ οι πράκτορες του MicroPSI είναι ιεραρχικά δίκτυα εξαπλούμενης ενεργοποίησης (spreading activation networks) τα οποία “πραγματοποιούν αντιληπτική μάθηση, κινητικό έλεγχο, σχηματισμό και ανάκληση μνήμης, λήψη αποφάσεων, προγραμματισμό και συναισθηματική διαρρύθμιση”.

Η γνωσιακή αρχιτεκτονική MicroPSI μοντελοποιεί, λοιπόν, τόσο ένα σύστημα κινητοποίησης όσο και ένα σύστημα συναισθημάτων, τα οποία συνδέονται στενά. Μπορούμε να ορίσουμε την ευφυΐα ως επίλυση προβλημάτων για την επίτευξη ορισμένων στόχων. Ωστόσο, η γενικότητα και ευελιξία του ανθρώπινου μυαλού προέρχεται σε μεγάλο βαθμό από την ικανότητα εντοπισμού και επιλογής κατάλληλων στόχων, επομένως ένα υπολογιστικό σύστημα το οποίο μοντελοποιεί την ανθρώπινη νόηση θα πρέπει να περιλαμβάνει αυτού του είδους την αυτονομία. Η λύση που δίνεται στα πλαίσια του MicroPSI δεν προϋποθέτει συγκεκριμένους στόχους, αλλά έναν ελάχιστο αριθμό συστημικών αναγκών (*needs*), οι οποίες σηματοδοτούνται στο γνωσιακό σύστημα ως *ορμές* (*urges*). Οι στόχοι καθορίζονται ως αποτέλεσμα της εκμάθησης της ικανοποίησης των αναγκών αυτών στο πλαίσιο ενός περιβάλλοντος. Οι ανάγκες του συστήματος αλλάζουν συνεχώς, επομένως το σύστημα θα χρειάζεται να επανεξετάζει τους σκοπούς και τη συμπεριφορά του συνέχεια, με αποτέλεσμα μία δυναμική ισορροπία δραστηριοτήτων. Η νόηση μπορεί να μεταβάλλει τους στόχους, τις προσδοκίες ανταμοιβής και τις προτεραιότητες του συστήματος, δεν μπορεί να επιδράσει άμεσα στις ίδιες τις ανάγκες.

Το μοντέλο αυτό της κινητοποίησης βασίζεται στη θεωρία PSI του Dörner, ενώ έχει επεκταθεί ώστε να εξηγεί αναλυτικότερα χαρακτηριστικά προσωπικότητας, αισθητική εκτίμηση και ρομαντική στοργή (Bach, 2012b; Dörner, 1999). Η τελευταία έκδοση της αρχιτεκτονικής (MicroPSI 2, 2012a) δεν υλοποιεί όλες τις πτυχές του συστήματος κινητοποίησης και των αναγκών στις οποίες βασίζεται, αλλά είναι αρκετά γενικό και υψηλού επιπέδου, ώστε να μπορεί προσαρμοστεί για εφαρμογές, ενώ έχει ήδη ενσωματωθεί σε άλλες γνωσιακές αρχιτεκτονικές όπως η OpenCog (Cai et al., 2012).

10.2. Συναίσθημα και νόηση

Είναι γενικά αποδεκτό στη γνωσιακή επιστήμη πως το συναίσθημα και το θυμικό (emotions and affect) παίζουν σημαντικό ρόλο στη διαρρύθμιση της νόησης του ανθρώπου. Ωστόσο, δεν υπάρχει συναίνεση σχετικά με το πως προκύπτουν τα συναισθήματα και μέσω ποιών μηχανισμών επηρεάζουν τις νοητικές διεργασίες. Η γνωσιακή αρχιτεκτονική MicroPSI επιχειρεί να δώσει μία λειτουργική, έστω και μερική, απάντηση στο ζήτημα αυτό (Bach, 2012b; Dörner, 2002). Η υπολογιστική μοντελοποίηση του συναισθήματος και του θυμικού βρίσκει σήμερα χρήση κατά κύριο λόγο σε πρακτικές εφαρμογές, όπως τη μοντελοποίηση συμπεριφοράς, την επικοινωνία ανθρώπου-υπολογιστή και κοινωνικές προσομοιάσεις. Τα μοντέλα που χρησιμοποιούνται για αυτούς τους σκοπούς είναι κατά κανόνα εξωτερικά και περιγραφικά μοντέλα συναισθηματικών πρακτόρων. Η αυτοαξιολόγηση των πρακτόρων, οι μηχανισμοί φιλτραρίσματος της πρόσβασης στη μνήμη, της αντίληψης και του ελέγχου της δράσης καθώς και η σχέση μεταξύ του συναισθήματος και της κινητοποίησης δεν περιλαμβάνονται σε αυτά τα απλοϊκά μοντέλα εμπορικής χρήσης. Για την βαθύτερη κατανόηση των παραπάνω μηχανισμών χρειαζόμαστε εσωτερικά, επεξηγηματικά μοντέλα, τα οποία εξηγούν λειτουργικά όλους τους μηχανισμούς και τις αλληλεπιδράσεις μεταξύ τους, που έχουν ως αποτέλεσμα την περίπλοκη συμπεριφορά των συναισθηματικών πρακτόρων.

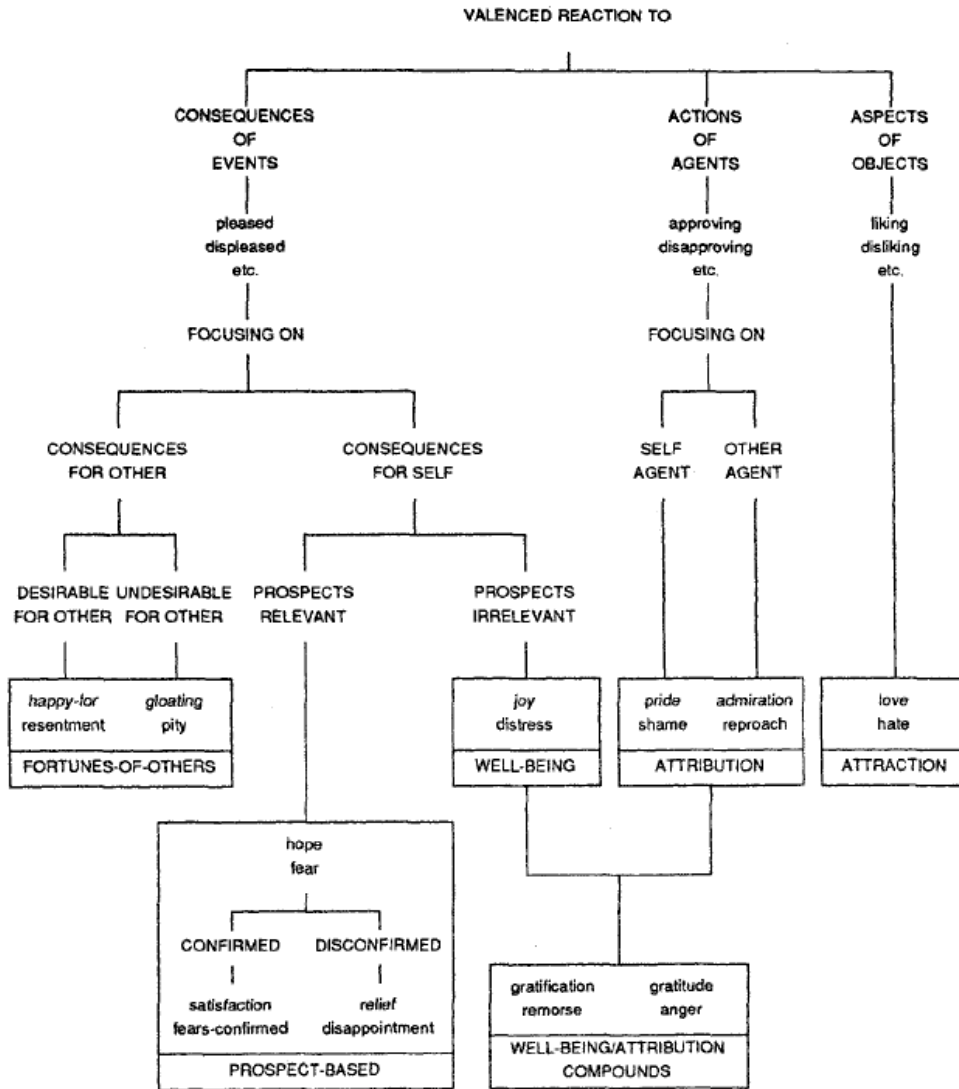
Στη σύγχρονη βιβλιογραφία των συνθετικών συναισθημάτων οι περιγραφικές θεωρίες είναι κυρίαρχες, και ιδιαίτερα οι θεωρίες *εκτίμησης (appraisal)*, σύμφωνα με τις οποίες τα εσωτερικά και εξωτερικά ερεθίσματα σε σχέση με έναν πράκτορα εκτιμώνται και δημιουργούν συμπεριφορικά και ψυχολογικά αποτελέσματα. Παρουσιάζονται ατομικές διαφορές μεταξύ ατόμων, ωστόσο η γενική έκφραση των συναισθημάτων γίνεται με παρόμοιο τρόπο. Με βάση τέτοιες θεωρίες έχουν αναπτυχθεί γενικές ταξινομήσεις συναισθηματικών καταστάσεων, όπως το μοντέλο Ortony-Clore-Collins, το οποίο αναπαριστά σε υψηλό επίπεδο τις εκτιμήσεις του συστήματος ενώ αντιμετωπίζει τα συναισθήματα ως αντιδράσεις με σθένος (valenced reactions) σε αποτελέσματα γεγονότων, δράσεις πρακτόρων ή πτυχές αντικειμένων, διακρίνοντας τα παραπάνω σε θεμιτά και μη θεμιτά, σε αυτά που αφορούν τον ίδιο πράκτορα ή άλλους, που συμβαίνουν ή είναι υποθετικά κτλ. (Ortony et al., 1988). Στο διάγραμμα 8 παρουσιάζεται η ταξινόμηση αυτή των συναισθημάτων υψηλού επιπέδου.

Με βάση το πλαίσιο αυτό, είναι εφικτό να καθοριστούν τα συναισθήματα σε μία τυπική γλώσσα η οποία περιγράφει:

1. Για *γεγονότα*: την επιθυμητότητά τους για τον ίδιο τον πράκτορα, την επιθυμητότητά τους για άλλους, την αξιολόγησή τους, την αρέσκειά τους, την πιθανότητα εμφάνισής τους, την σχετική προσπάθεια που χρειάζεται και αν τελικά υλοποιούνται
2. Για *πράκτορες*: το κατά πόσο είναι αξιέπαινοι, τη γνωστική τους συνάφεια, την παρέκκλιση των προσδοκιών τους

3. Για αντικείμενα: την ελκυστικότητα και οικειότητά τους

Το μοντέλο OCC προσφέρει, έτσι, μία καθαρή οντολογία η οποία προσφέρεται για ευθεία υλοποίηση εντός υπολογιστικών συστημάτων. Ωστόσο, οι εκ των έξω παρατηρήσιμες συμπεριφορικές επιπτώσεις των συναισθημάτων καλύπτουν μόνο ένα μικρό κομμάτι των λειτουργιών τους. Τα συναισθήματα και το θυμικό δίνουν εσωτερική ανατροφοδότηση στον πράκτορα, βοηθούν στην σκέψη, δομούν τις κοινωνικές αλληλεπιδράσεις, βελτιώνουν την επικοινωνία, προεγείρουν την ανάκληση της μνήμης μέσω συνδετιστικών στοιχείων, προκατειλημμένης αντίληψης και φιλτραρίσματος της πρόσβασης σε νοητικό περιεχόμενο σύμφωνα με τη δυναμική της εκάστοτε κατάστασης στην οποία βρίσκεται ο πράκτορας. Δηλαδή, τα συναισθήματα ρυθμίζουν τη νοητική επεξεργασία, επομένως για την βαθύτερη κατανόησή τους είναι απαραίτητη η πραγμάτωσή τους εντός μία γνωσιακής αρχιτεκτονικής.

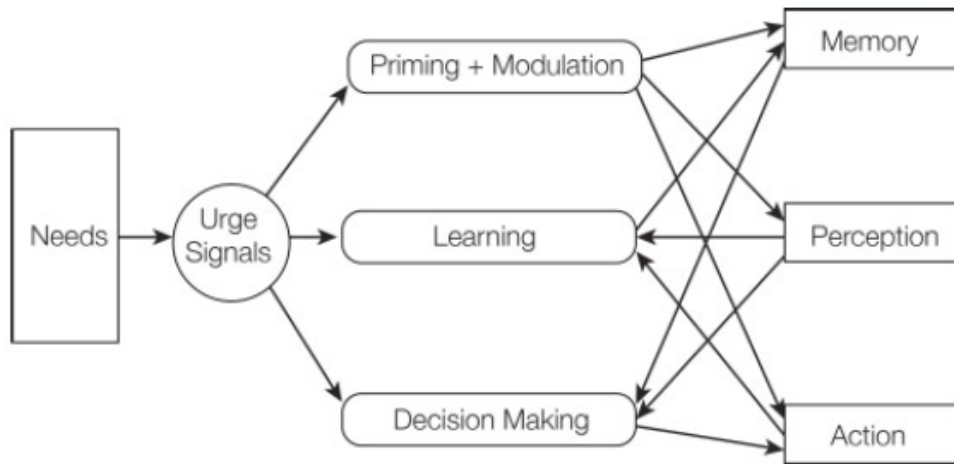


Διάγραμμα 8 - Η ταξινόμηση των συναισθημάτων στο μοντέλο OCC (Ortony et al., 1988)

10.3. Η λειτουργία της κινητοποίησης

Παρακάτω περιγράφεται η λειτουργία της κινητοποίησης στο MicroPSI ξεκινώντας από τις ανάγκες και φτάνοντας στην τελική συμπεριφορά. Οι ανάγκες ενός πράκτορα αντικατοπτρίζουν τις προϋποθέσεις για την ευημερία του οργανισμού, οι οποίες σχηματίστηκαν, στους ανθρώπους και τα υπόλοιπα ζώα της γης, μέσω εκατομμυρίων ετών εξέλιξης. Κάθε ανάγκη d σχετίζεται με μία μεταβλητή c_d , η οποία μεταβάλλεται μεταξύ μίας τιμής-στόχου v_d (στο οποίο η ανάγκη είναι πλήρως ικανοποιημένη) και ενός άκρου v_0 (στο οποίο το ίδιο στο σύστημα ίσως παύει να λειτουργεί). Όταν η τιμή μίας ανάγκης ξεφεύγει από την τιμή-στόχο και δεν είναι

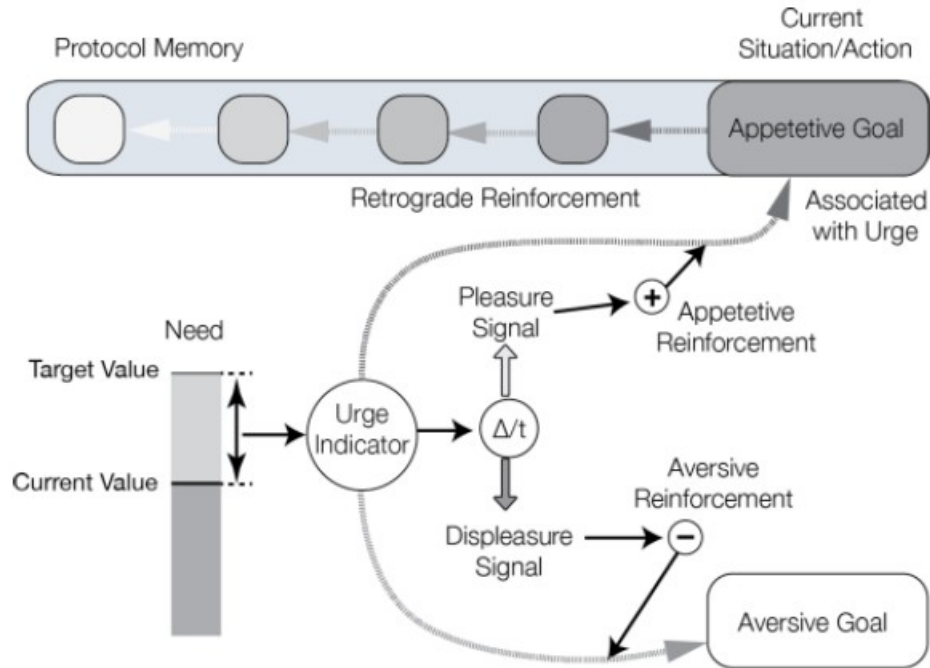
εφικτή αυτόνομη διαρρύθμιση (όπως ο μεταβολισμός), η ανάγκη σηματοδοτείται από έναν δείκτη ορμής. Η διαφορά μεταξύ της τρέχουσας τιμής και της τιμής-στόχου καθορίζει την ισχύ της ορμής (*strength of the urge*). Το πηλίκο της ισχύος μίας ορμής με την απόσταση της τρέχουσας τιμής και της κοντινότερης οριακής τιμής συνιστά την *επιτακτικότητα της ορμής* (*urgency of the urge*). Προσδιορίζεται ακόμη ένα εύρος στο οποίο γίνεται να βρίσκεται η τιμή μίας ανάγκης, τόσο πάνω όσο και κάτω από την τιμή-στόχο, αν και πρακτικά, στους ανθρώπους φαίνεται να υπάρχουν διαφορετικοί δείκτες ορμής και μηχανισμοί αντιμετώπισης για παρέκκλιση σε διαφορετικές κατευθύνσεις από την τιμή ικανοποίησης μίας ανάγκης (θερμότητα και ψύξη, σίτιση και κορεσμός, ανάπαυση και άσκηση).



Διάγραμμα 9 – Η επίδραση των δεικτών ορμής σε διάφορες λειτουργίες της αρχιτεκτονικής MicroPSI (Bach, 2015)

Οι δείκτες ορμών παίζουν σημαντικό ρόλο στην αρχιτεκτονική της νόησης του MicroPSI (Διάγραμμα 9). Πιο συγκεκριμένα, η ένταση των ορμών καθορίζει τη *διέγερση* (*arousal*), την *ταχύτητα εκτέλεσης* και την *ανάλυση της γνωσιακής επεξεργασίας*. Αλλαγές στους δείκτες ορμών υποδεικνύουν την ικανοποίηση ή επιδείνωση αντίστοιχων αναγκών. Γρήγορες αλλαγές προκαλούνται συνήθως από πράξεις του πράκτορα ή γεγονότα του περιβάλλοντος. Οι αλλαγές αυτές συνοδεύονται από σήματα ευχαρίστησης (στην περίπτωση ικανοποίησης) ή σήματα δυσαρέσκειας (στην περίπτωση επιδείνωσης). Τα σήματα αυτά χρησιμοποιούνται ως ενισχυτές για μάθηση μέσω κινητοποίησης (Διάγραμμα 10). Εντός του συστήματος, κάθε είδους σήμα συνδέεται σε έναν *συσχετιστή* (*associator*), ο οποίος συσχετίζει την αναπαράσταση του δείκτη ορμής με την τρέχουσα κατάσταση ή δράση. Επιπλέον, η μάθηση ενισχύει συνδέσεις μεταξύ της τρέχουσας κατάστασης και αυτών που προηγήθηκαν. Ο παραπάνω μηχανισμός, σε συνδυασμό με μηχανισμούς για λήθη και μνημονική παγίωση, οδηγεί στη μάθηση αλληλουχιών συμπεριφορών που καταλήγουν σε καταστάσεις-στόχους. Οι *στόχοι* (*goals*) είναι δράσεις ή καταστάσεις που επιτρέπουν την ικανοποίησης μίας ορμής (*appetitive goals*) ή

απειλούν την επιδείνωσή της (aversive goals). *Κίνητρο (motive)* ονομάζεται μία ορμή η οποία έχει συσχετιστεί με έναν συγκεκριμένο στόχο, και κάθε πράξη των πρακτόρων του MicroPSI κατευθύνεται στην επίτευξη ενός θετικού στόχου ή την αποφυγή ενός αρνητικού στόχου. Αυτός ο συσχετισμός μεταξύ ορμών και στόχων επιτρέπει στον πράκτορα να εντοπίσει πιθανούς τρόπους αντιμετώπισης, όποτε η ανάγκη προκύψει μελλοντικά, και να προεγείρει την αντίληψη και ανάκληση μνήμης του προς τις ενεργές ορμές. Σημαντικό είναι ακόμη ότι οι δείκτες ορμής σηματοδοτούν τις τρέχουσες ανάγκες του συστήματος, και επομένως επηρεάζουν την λήψη αποφάσεων.



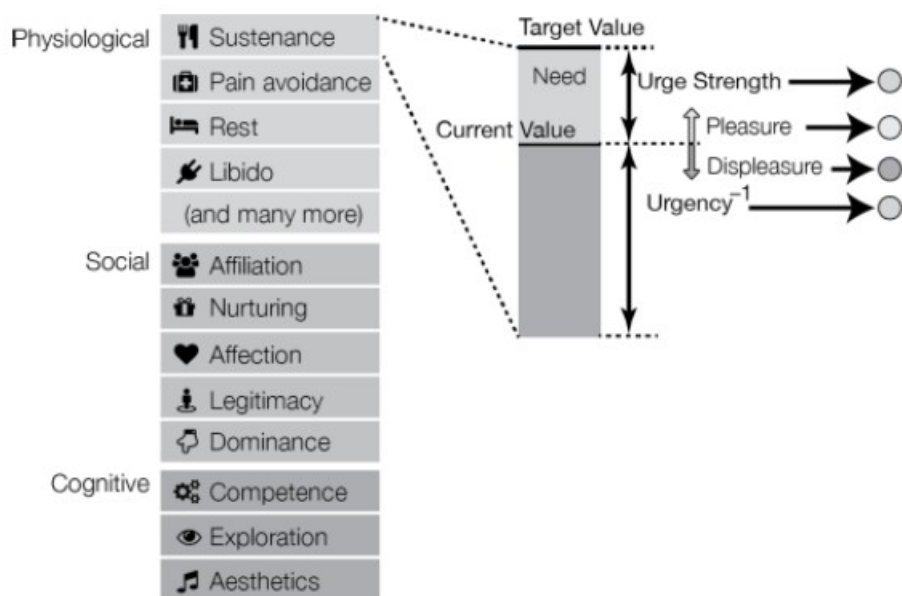
Διάγραμμα 10 – Η λειτουργία της μάθησης μέσω κινητοποίησης που επιτυγχάνεται με τον συνδυασμό δεικτών ορμών με στόχους (Bach, 2015)

10.4. Είδη αναγκών και γνωστικές ανάγκες

Σύμφωνα με την θεωρία PSI, οι ανάγκες ενός γνωστικού συστήματος χωρίζονται σε τρεις κατηγορίες: φυσιολογικές, κοινωνικές και γνωστικές, οι οποίες ενεργούν στο ίδιο επίπεδο και ταυτόχρονα, όχι ως ιεραρχία, όπως στην διάσημη πυραμίδα των συναισθημάτων του Maslow (Maslow & Frager, 1987). Για την καθιέρωση προτεραιοτήτων μεταξύ διαφορετικών ορμών, κάθε τιμή ανάγκης πολλαπλασιάζεται με ένα βάρος ($weight_d$), το οποίο εκφράζει τη σχετική σημασία της κάθε ανάγκης σε σχέση με τις υπόλοιπες. Επίσης, κάθε ανάγκη έχει μία *παράμετρο πτώσης (decay_d)* που καθορίζει πόσο γρήγορα το c_d κατευθύνεται προς το v_0 όταν η ανάγκη

αφήνεται ανικανοποίητη, επομένως και πόσο συχνά χρειάζεται να ανεφοδιάζεται. Ακόμα, υπάρχουν οι τιμές *αύξηση (gain_a)* και *απώλεια (loss_a)*, οι οποίες εκφράζουν πως κάθε ανάγκη αντιδρά σε ικανοποίηση ή επιδείνωση. Χρησιμοποιώντας το σύνολο των παραπάνω παραμέτρων {βάρος, πτώση, αύξηση, απώλεια} για κάθε ανάγκη ενός πράκτορα, καθίσταται δυνατή η μοντελοποίηση διαφορετικών ιδιοτήτων προσωπικότητας και ατομικών χαρακτηριστικών.

Οι φυσιολογικές ανάγκες ρυθμίζουν την βασική επιβίωση του οργανισμού και αντικατοπτρίζουν τις ανάγκες μεταβολισμού και της φυσικής ευημερίας του. Οι αντίστοιχοι δείκτες ορμής βασίζονται σε μετρήσεις της ιδιοδεκτικότητας (proprioception), του επιπέδου των ορμονών κτλ. Οι φυσιολογικές ανάγκες αποτελούνται μεταξύ άλλων από τη διατροφή, τη φυσική ακεραιότητα και αποφυγή πόνου, την ανάπαυση, την αποφυγή ακραίων θερμοκρασιών κ.α. Οι κοινωνικές ανάγκες καθορίζουν τη συμπεριφορά του πράκτορα απέναντι σε άλλους πράκτορες και ομάδες. Ικανοποιούνται και επιδεινώνονται από κοινωνικά σήματα και αντίστοιχες νοητικές αναπαραστάσεις, ωστόσο τα σήματα ευχαρίστησης και δυσαρέσκειας που λαμβάνονται από αυτές τις πηγές δεν είναι λιγότερο απτά για τον πράκτορα που τα βιώνει από τον φυσιολογικό πόνο. Επομένως, συχνά οι άνθρωποι είναι διατεθειμένοι να θυσιάσουν την ικανοποίηση φυσιολογικών αναγκών τους για να επιτύχουν κάποιο κοινωνικό στόχο (λήψη αναγνώρισης, διάσωση ενός παιδιού, κατάκτηση ενός συντρόφου, διατήρηση προσωπικής ακεραιότητας, αποφυγή απώλειας φήμης κτλ.). Ατομικές διαφορές μεταξύ των βαρών των κοινωνικών αναγκών πιθανώς οδηγούν σε ανθρώπους με διαφορετικά κοινωνικά χαρακτηριστικά (για παράδειγμα περισσότερο αλτρουιστές ή εγωιστές, εξωστρεφείς ή εσωστρεφείς, ρομαντικούς ή μη ρομαντικούς). Οι γνωστικές ανάγκες επιτρέπουν την επίλυση προβλημάτων, την απόκτηση ικανοτήτων, την εξερεύνηση, το παιχνίδι και τη δημιουργικότητα. Διαφορές στα βάρη των γνωστικών αναγκών πιθανώς επηρεάζουν την ευσυνειδησία ή τον αυθορμητισμό, την γνωστική δεκτικότητα σε καινούργιες ιδέες και τον ηδονισμό ενός πράκτορα.



Διάγραμμα 11 – Η δομή των αναγκών στο MicroPSI και χαρακτηριστικά μεταβολής και ρύθμισής τους (Bach, 2015)

Παρακάτω, αναλύονται βαθύτερα οι γνωστικές ανάγκες στη θεωρία PSI, δηλαδή *γνωστική ικανότητα (competence)*, *εξερεύνηση (exploration)* και *αισθητική (aesthetics)*, οι οποίες συνδέονται στενά με το γενικότερο θέμα της εργασίας και το σύστημα IDyOT.

Η ανάγκη για γνωστική ικανότητα μπορεί να σχετίζεται με συγκεκριμένες εργασίες, στην οποία περίπτωση μετρά την επιτυχία σε μεμονωμένες ικανότητες. Η επιτυχής εκτέλεση μίας ικανότητας και η απόκτηση νέων ικανοτήτων οδηγούν σε ικανοποίησή της, ενώ η αποτυχής εκτέλεση ή η προσδοκία αποτυχίας την επιδεινώνουν. Η ανάγκη της γνωστικής ικανότητας μπορεί επίσης να μετρά την επίδραση των δράσεων ενός πράκτορα στο περιβάλλον του, και να ικανοποιείται περισσότερο όσο μεγαλύτερο είναι το μέγεθος της παρατηρούμενης επίδρασης. Επιπροσθέτως, η ανάγκη της γενικής γνωστικής ικανότητας αποτελεί ένα σύνθετο μέτρο της ικανότητας ικανοποίησης αναγκών και απόκτησης ικανοτήτων σε εργασίες. Η ισχύς του δείκτη ορμής της ανάγκης αυτής χρησιμοποιείται ως ευρετικό στο συλλογισμό σχετικά με την γενική επίδοση του πράκτορα και για να προβλέψει την επιτυχία σε άγνωστες εργασίες. Χαμηλές τιμές γενικής γνωστικής ικανότητας αντιστοιχούν σε έλλειψη αυτοπεποίθησης. Προβλεπόμενη ή πραγματική αποτυχία σε εργασίες μπορεί να οδηγεί σε ορμή αναπλήρωσης της γνωστικής ικανότητας, που μπορεί να επιτευχθεί με την ενασχόληση με κάποια εργασία με υψηλή πιθανότητα επιτυχίας, δυναμική που ίσως σχετίζεται με την αναβλητικότητα.

Η ανάγκη για εξερεύνηση έχει να κάνει με την απόκτηση βεβαιότητας για αντικείμενα και διαδικασίες στο περιβάλλον. Η επιβεβαίωση για αβέβαιες προσδοκίες αντίληψης και δράσης αυξάνουν τη βεβαιότητα (και έτσι μειώνουν την ανάγκη για εξερεύνηση), ενώ παραβιάσεις προσδοκιών ή η έλλειψη διαθέσιμων προσδοκιών μειώνουν τη βεβαιότητα (και αυξάνουν την ανάγκη για σιγουριά, που αποκτάται μέσω εξερευνητικών δράσεων). Η ισχύς του δείκτη ορμής της εξερεύνησης χρησιμοποιείται ως ευρετικό για το βαθμό αβεβαιότητας σε μία δεδομένη κατάσταση.

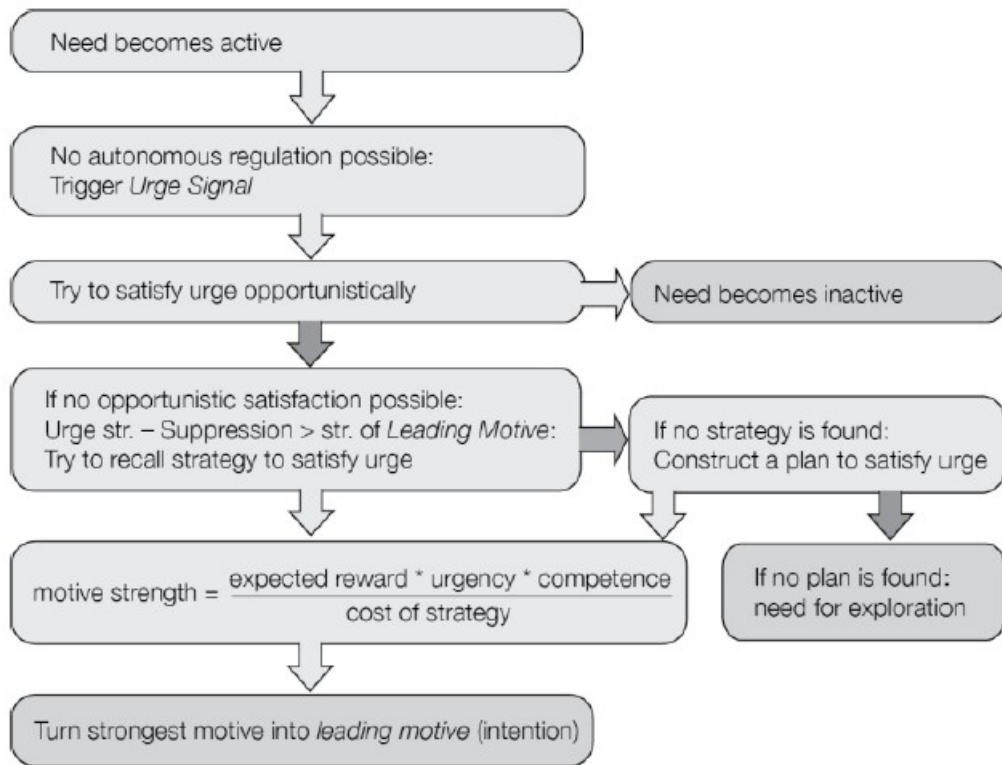
Τέλος, η ανάγκη για αισθητική διακρίνεται μεταξύ της αισθητική που προσανατολίζεται σε ερεθίσματα και της αφηρημένης αισθητικής. Η αισθητική που προσανατολίζεται σε ερεθίσματα είναι η ανάγκη για εγγενώς ευχάριστα ερεθίσματα, όπως αρμονικούς ήχους, συγκεκριμένα περιβαλλοντικά χαρακτηριστικά, απτικές αισθήσεις. Η θετική αντίδραση σε αυτά τα ερεθίσματα αποτελεί είτε επιφανόμενο της αισθητικής επεξεργασίας, είτε εξυπηρετεί έμμεσους σκοπούς (όπως την αναζήτηση εύφορης γης, τον εντοπισμό υγείων συντρόφων κτλ). Από την άλλη, η αφηρημένη αισθητική είναι η ανάγκη εξεύρεσης δομής σε νοητικές αναπαραστάσεις, και η αντικατάσταση των υπάρχοντων αναπαραστάσεων με πιο αποδοτικές. Η ίδια ανάγκη ευθύνεται για την ανακάλυψη της μαθηματικής κομψότητας, της μουσικής μορφής και της ευχαρίστησης που προκαλείται από την αναθεώρηση εννοιολογικών δομών.

10.5. Γνωσιακός κύκλος

Όπως αναφέρθηκε παραπάνω, όλη η συμπεριφορά των πρακτόρων του MicroPSI κατευθύνεται είτε στην ικανοποίηση μίας ανάγκης είτε την αποφυγή επιδείνωσης της. Ο γνωσιακός κύκλος του συστήματος εκκινείται από την απομάκρυνση κάποιας ανάγκης από την τιμή-στόχο της. Η ενεργοποίηση μίας ανάγκης οδηγεί, όταν δεν μπορεί να λυθεί από αυτόνομη διαρρύθμιση, στην ενεργοποίηση των δεικτών ορμής, που με τη σειρά τους σηματοδοτούν την ανάγκη στο αντιδραστικό κομμάτι του γνωστικού πράκτορα. Από παλαιότερες εμπειρίες, η εκάστοτε ορμή έχει συσχετιστεί με ποικίλες δράσεις και αντικείμενα που στο παρελθόν ικανοποίησαν ή επιδείνωσαν την αντίστοιχη ανάγκη. Μέσω εξαπλούμενης ενεργοποίησης στο δίκτυο αυτών των συσχετισμών, σχετικό περιεχόμενο στη μνήμη και την αντίληψη τονίζεται. Αν βρεθεί άμεσα ένας θετικός στόχος, ο οποίος δεν επεμβαίνει στην τρέχουσα δραστηριότητα, τότε η ορμή μπορεί να ικανοποιηθεί επιτυχώς χωρίς σημαντική χρήση της συνειδητής προσοχής.

Σε περίπτωση που δεν βρεθεί τέτοιος στόχος, ο πράκτορας επιχειρεί να καταστείλει τη νέα ορμή, ανάλογα με την ισχύ του τρέχοντος κυρίαρχου κινήτρου. Εφόσον η ορμή ξεπεράσει το κατώφλι επιλογής, η τρέχουσα συμπεριφορά παύει και ο πράκτορας προσπαθεί να ανακαλέσει μία εφαρμόσιμη στρατηγική, δηλαδή μία αλληλουχία συμπεριφορών που οδηγεί στην επίτευξη ενός θετικού στόχου, μέσω της προαναφερόμενης διαδικασίας εξαπλούμενης ενεργοποίησης της μνήμης. Όταν, ωστόσο, μία τέτοια αλληλουχία δεν βρεθεί αυτόματα, ο πράκτορας χρησιμοποιεί επιπλέον πόρους προσοχής και επιχειρεί να κατασκευάσει ένα σχέδιο, δηλαδή μία αλυσίδα που πιθανώς μπορεί να συνδέσει την τρέχουσα κατάσταση του κόσμου με έναν

από τους θετικούς στόχους. Αν και πάλι ο πράκτορας αποτύχει, τότε εγκαταλείπει προσωρινά την επιδίωξη ικανοποίησης της τρέχουσας ορμής και αυξάνει την ανάγκη του για εξερεύνηση, που έχει ως αποτέλεσμα συμπεριφορές συλλογής περισσότερων πληροφοριών και δοκιμή πειραματικών στρατηγικών. Όταν εντοπιστεί ένα σχέδιο, αυτό συνιστά ένα κίνητρο (motive, δηλαδή συνδυασμό ενεργής ορμής, συγκεκριμένου θετικού στόχου και αλληλουχίας δράσεων για την επίτευξη αυτού του στόχου). Η ισχύς του κινήτρου καθορίζεται συνεκτιμώντας την ανταμοιβή επίτευξης του στόχου, της επιτακτικότητας της ικανοποίησης της ανάγκης, της πιθανότητας επιτυχίας και διαιρώντας το αποτέλεσμα με το εκτιμώμενο κόστος υλοποίησης του σχεδίου. Το ισχυρότερο κίνητρο ανυψώνεται σε πρόθεση και καθορίζει τις δράσεις του πράκτορα. Όλη η παραπάνω διαδικασία αποτελεί τον τρόπο με τον οποίο οι πράκτορες του MicroPSI λαμβάνουν τις αποφάσεις τους.



Διάγραμμα 12 - Η ακολουθία ενεργειών (γνωσιακός κύκλος) που οδηγεί σε δράση στο σύστημα MicroPSI (Bach, 2015)

10.6. Διαρρυθμισμό της νόησης και συναίσθημα

Η ισχύς των αναγκών ενός πράκτορα δεν καθορίζει μόνο ποιους στόχους επιδιώκει, αλλά και πως τους επιδιώκει. Πιο συγκεκριμένα, η γνωστική και αντιληπτική επεξεργασία διαμορφώνονται από ένα σύνολο καθολικών διαρρυθμιστών (*global modulators*):

- Η *διέγερση (arousal)* αντικατοπτρίζει τον συνδυασμό της ισχύος και της επιτακτικότητας των αναγκών ενός πράκτορα, ενώ η υψηλή ενεργοποίησή της οδηγεί σε μεγαλύτερη δαπάνη ενέργειας σε δράσεις, ετοιμότητα, ισχυρότερες και γρηγορότερες αντιδράσεις.
- Το *σθένος (valence)* αναπαριστά μία ποιοτική εκτίμηση της τρέχουσας κατάστασης και ορίζεται ως η διαφορά του άθροισματος όλων των τρεχόντων σημάτων ευχαρίστησης με το άθροισμα όλων των σημάτων δυσaréσκειας, ξεκινώντας από ένα επίπεδο βάσης.
- Η *επιθετικότητα/υποβολή (aggression/submission)* καθορίζει τη στάση του πράκτορα απέναντι σε ένα παρακολουθούμενο αντικείμενο. Η επιθετικότητα ενεργοποιείται από την αντίδραση αρνητικού σθένους απέναντι σε έναν άλλο πράκτορα ο οποίος εμποδίζει κάποιο στόχο και αυξάνει την πιθανότητα σύγκρουσης (ενώ χαμηλή τιμή γνωστικής ικανότητας οδηγεί σε χαμηλή τιμή επιθετικότητας). Γενικά, ο διαρρυθμιστής αυτός καθορίζει την τάση προσέγγισης ή απομάκρυνσης από ένα παρακολουθούμενο αντικείμενο.

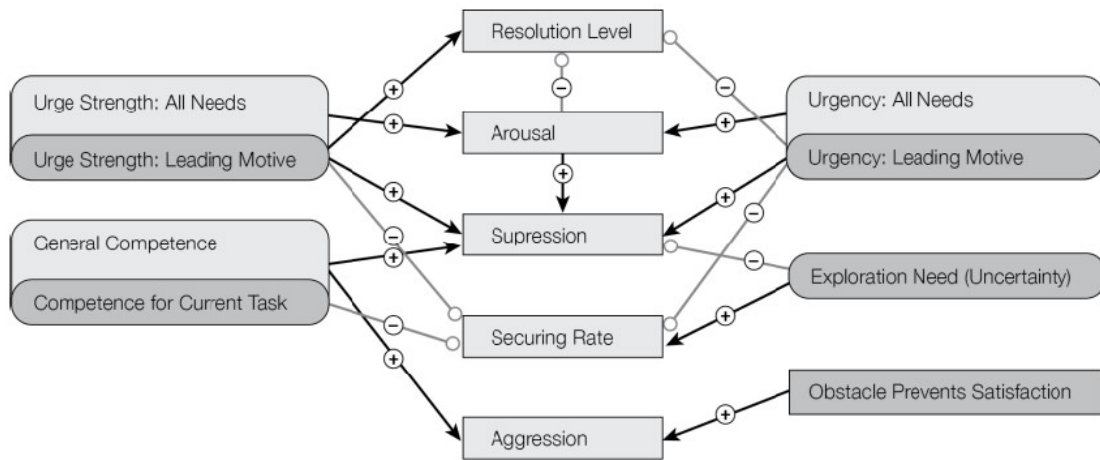
Ο συνδυασμός των παραπάνω διαρρυθμιστών καθορίζει τη *θυμική κατάσταση (affective state)* του πράκτορα, σχηματίζοντας έναν χώρο (σε αντιστοιχία με τους εννοιολογικούς χώρους), διαφορετικές περιοχές του οποίου αντιστοιχούν σε διαφορετικές συναισθηματικές διαμορφώσεις.

Εκτός από τους τρεις καθολικούς διαρρυθμιστές, το MicroPSI χρησιμοποιεί άλλους τρεις *διαρρυθμιστές προσοχής (attentional modulators)*:

- Το *επίπεδο ανάλυσης (resolution level)* αναπαριστά το βάθος της προσοχής στις λεπτομέρειες κατά τη διάρκεια γνωστικών και αντιληπτικών εργασιών. Όσο υψηλότερη είναι η τιμή του διαρρυθμιστή, τόσο πιο ακριβείς θα είναι οι λύσεις και αναπαραστάσεις στις οποίες φτάνει ο πράκτορας, ενώ όσο χαμηλότερη η τιμή του, τόσο πιο γρήγορη η απόκρισή του.
- Η *καταστολή (suppression)* σχετίζεται με την διαδικασία λήψης αποφάσεων, και αποτελεί μία ενίσχυση του δεσπόζοντος ενεργού κινήτρου (που συνιστά την πρόθεση του πράκτορα). Πιο αναλυτικά, η τιμή του διαρρυθμιστή καθορίζει ένα κατώφλι επιλογής, το οποίο προσδίδει ισχυρότερη συγκέντρωση στον τρέχοντα στόχο και στενότερη κατεύθυνση της προσοχής, αποτελώντας έτσι έναν μηχανισμό για την αποφυγή της ταλάντωσης μεταξύ ανταγωνιζομένων κινήτρων.

- Ο βαθμός ασφάλειας (*securing rate*) καθορίζει τη συχνότητα της πρόσληψης/ενημέρωσης πληροφορίας από το περιβάλλον. Δυναμικά περιβάλλοντα απαιτούν περισσότερους γνωστικούς πόρους για αντιληπτική επεξεργασία, ενώ στατικά περιβάλλοντα ελευθερώνουν πόρους για περίσκεψη. Επομένως, ο διαρρυθμιστής καθορίζει την κατεύθυνση της προσοχής: προς τα έξω, στο περιβάλλον ή προς τα μέσα, στη νοητική σκηνή.

Οι τιμές που λαμβάνουν οι διαρρυθμιστές εξαρτώνται από τη διαμόρφωση των ορμών και την αλληλεπίδραση μεταξύ των ίδιων των διαρρυθμιστών (βλ. διάγραμμα 13). Πιο συγκεκριμένα, η διέγερση καθορίζεται από την ισχύ και την επιτακτικότητα όλων των αναγκών. Υψηλή τιμή διέγερσης αυξάνει επίσης το επίπεδο ανάλυσης και την καταστολή. Το επίπεδο ανάλυσης αυξάνεται επίσης από την ισχύ του τρέχοντος μοτίβου, αλλά μειώνεται από την επιτακτικότητά του, επιτρέποντας ταχύτερες αντιδράσεις. Η καταστολή αυξάνεται από την ισχύ και την επιτακτικότητα του τρέχοντος ενεργού κινήτρου, ενώ μειώνεται από χαμηλή τιμή γενικής γνωστικής ικανότητας. Ο βαθμός ασφάλειας μειώνεται από την ισχύ και επιτακτικότητα του τρέχοντος κινήτρου, αλλά αυξάνεται με χαμηλή τιμή γνωστικής ικανότητας και υψηλή ανάγκη για εξερεύνηση (που ταυτίζεται με την εμπειρία της αβεβαιότητας). Η επιθετικότητα ενεργοποιείται από πράκτορες ή εμπόδια τα οποία αποτρέπουν την πραγματοποίηση ενός σημαντικού κινήτρου και μειώνεται από χαμηλή γνωστική ικανότητα.



Διάγραμμα 13 - Οι διαρρυθμιστές του *MicroPSI* και οι διαδράσεις μεταξύ τους (Bach, 2015)

Επιπλέον, κάθε διαρρυθμιστής έχει τουλάχιστον τέσσερις, σε γενικές γραμμές σταθερές, παραμέτρους, οι οποίες ευθύνονται για τις ατομικές διαφορές μεταξύ ατόμων: το *επίπεδο βάσης* (*baseline*) αποτελεί την προεπιλεγμένη τιμή του διαρρυθμιστή, το *εύρος* (*range*) καθορίζει τα άνω και κάτω όρια των αλλαγών στις τιμές του, η *πτητικότητα* (*volatility*) την αντίδραση σε αλλαγές και η *διάρκεια* (*duration*) το χρονικό ποσό μέχρι ο διαρρυθμιστής να επιστρέψει στο επίπεδο βάσης.

Σύμφωνα με τη θεωρία PSI, τα συναισθήματα μπορεί είτε να μην έχουν κατεύθυνση, είτε να αποτελούν αντίδραση με σθένος σε ένα αντικείμενο. Όταν ανήκουν στην κατηγορία χωρίς κατεύθυνση, μπορούν να περιγραφούν ως συγκεκριμένες διαμορφώσεις των τιμών των διαρρυθμιστών, σε συνδυασμό με το επίπεδο γνωστικής ικανότητας και αβεβαιότητας. Όταν ανήκουν στην κατηγορία κατευθυνόμενων συναισθημάτων, το αντικείμενο στο οποίο κατευθύνονται μπορεί να είναι μία συγκεκριμένη νοητική αναπαράσταση με περιεχόμενο σχετικό με την κινητοποίηση του πράκτορα, που συνδυάζεται με μία θυμική κατάσταση. Παραδείγματα μη κατευθυνόμενων συναισθημάτων αποτελούν η χαρά (θετικό σθένος, υψηλή διέγερση), η ευτυχία (θετικό σθένος, χαμηλή διέγερση) ή το άγχος (αρνητικό σθένος, υψηλή αβεβαιότητα, υποβολή και χαμηλή γνωστική ικανότητα). Παραδείγματα κατευθυνόμενων συναισθημάτων αποτελούν ο φόβος (αρνητικό σθένος κατευθυνόμενο σε έναν αρνητικό στόχο και χαμηλή γνωστική ικανότητα) και ο θυμός (αρνητικό σθένος κατευθυνόμενο σε έναν πράκτορα ο οποίος εμπόδισε έναν θετικό στόχο ή προκάλεσε έναν αρνητικό, υψηλή επιθετικότητα και διέγερση). Η ζήλια μπορεί να παρουσιαστεί είτε ως φόβος (κατευθυνόμενος στην απώλεια ενός ερωτικού δεσμού με υψηλή υποβολή) είτε ως θυμός (κατευθυνόμενος σε έναν πράκτορα ο οποίος εμποδίζει την ικανοποίηση κοινωνικών ή ερωτικών αναγκών). Τα συναισθήματα θεωρούνται, λοιπόν, αντιληπτικές μορφές (perceptual gestalt), οι οποίες εξαρτώνται από τον τομέα του χώρου των θυμικών καταστάσεων στον οποίο ανήκουν και από το σημαντικό για την κινητοποίηση αντικείμενο στο οποίο πιθανώς κατευθύνονται.

11. Πιθανές συνδέσεις των δύο μοντέλων

Το τελευταίο αυτό κεφάλαιο αποτελεί παράθεση ιδεών σχετικά με πιθανές συνδέσεις μεταξύ των μοντέλων IDyOT και MicroPSI και θα αποτελέσει θεωρητική βάση για μελλοντική υπολογιστική υλοποίηση κάποιου συνδυασμού των δύο γνωσιακών αρχιτεκτονικών. Καταρχάς, μπορούμε να λάβουμε μία οπτική υψηλού επιπέδου, προκειμένου να καταλάβουμε βαθύτερα τα ευρύτερα χαρακτηριστικά κάθε αρχιτεκτονικής και να εντοπίσουμε σημεία πιθανών συνδέσεων. Η αρχιτεκτονική IDyOT μοντελοποιεί την αντίληψη και νόηση σειριακών δεδομένων και συνδέεται με τη δημιουργικότητα. Θα μπορούσαμε, λοιπόν, να τη χαρακτηρίσουμε ως ένα μοντέλο χαμηλού (βάσει της λειτουργίας των γεννητριών) και μεσαίου επιπέδου (βάσει της λειτουργίας του οικουμενικού χώρου εργασίας, της μνήμης και των επακόλουθών τους), ενώ δεν δίνεται σαφής λειτουργική εξήγηση για την κινητοποίηση του συστήματος πέρα από το εξελικτικό επιχείρημα υπέρ του μηχανισμού. Η αρχιτεκτονική MicroPSI, αντιθέτως, έχει ως αντικείμενό της την κινητοποίηση και το συναίσθημα ενός πράκτορα ο οποίος δρα σε ένα σύνθετο περιβάλλον και σχετίζεται με την έρευνα της Γενικής Τεχνητής Νοημοσύνης (AGI). Συνιστά, έτσι, ένα μοντέλο υψηλού επιπέδου, που δεν προσδιορίζει λεπτομέρειες για τα χαμηλότερα χαρακτηριστικά της αρχιτεκτονικής, αλλά δρα σε ένα υψηλότερο επίπεδο αφάιρησης. Ο ίδιος ο Bach, μάλιστα, δηλώνει ότι η αρχιτεκτονική προσφέρεται για σύνδεση με άλλα μοντέλα (2015). Στο βιβλίο “Society of Mind”, ο Marvin Minsky κάνει τη διάκριση μεταξύ του A-brain και του B-brain, σύμφωνα με την οποία το A-brain συνδέεται με τον εξωτερικό κόσμο και με το B-brain, ενώ το B-brain δεν έχει άμεση επαφή με τον έξω κόσμο, ωστόσο χρησιμοποιεί τα δεδομένα που λαμβάνει από το A-brain για να βελτιστοποιήσει την απόδοσή του τελευταίου και να ενισχύσει τη λειτουργία του (Minsky & Lee, 1988). Οι παρακάτω ιδέες αποτελούν μία πρώιμη προσπάθεια για θεωρητική μίξη των αρχιτεκτονικών, ώστε το IDyOT να αποτελέσει το A-brain και το MicroPSI το B-brain.

Ο Bach παραθέτει ένα σύνολο αρχιτεκτονικών προϋποθέσεων που χρειάζεται να πληρούνται, για την εφαρμογή του συστήματος κινητοποίησης και αναφαινόμενων συναισθημάτων (2012b):

- Ένα σύνολο ορμών, τα οποία σηματοδοτούν ανάγκες του συστήματος. Το IDyOT δεν περιλαμβάνει ορμές, καθώς δεν διαθέτει σαφώς καθορισμένες ανάγκες.
- Ένας μηχανισμός επιλογής, ο οποίος προάγει μία ορμή σε κίνητρο του συστήματος. Επίσης λείπει στο IDyOT.
- Μία συνειρμική μνήμη, η οποία μπορεί να προεγερθεί από τα τρέχοντα ενεργά κίνητρα. Ούτε αυτή η λειτουργία εντοπίζεται στο IDyOT, στο οποίο οι προβλέψεις του συστήματος επηρεάζονται μόνο από τη δομή της μνήμης, επομένως το φαινόμενο της προέγερσης συναντάται μόνο σε σχέση με παλαιότερες εμπειρίες, και όχι σύμφωνα με κινητοποιητικά σχετικό περιεχόμενο.

- Μηχανισμοί εκτέλεσης δράσεων, οι οποίοι φέρουν πράξεις εις πέρας. Το IDyOT διαθέτει γεννήτριες που κάνουν συνεχείς προβλέψεις και μηχανισμούς βελτιστοποίησης των δομών της μνήμης (τελεστές στους γεωμετρικούς εννοιολογικούς χώρους). Ωστόσο, οι δυνατότητες αυτές είναι περιορισμένες σε σχέση με έναν γνωστικό πράκτορα ο οποίος δρα σε ένα σύνθετο περιβάλλον, όπως οι άνθρωποι, επομένως χρειάζονται επιπλέον προσθήκες.
- Ένα σύνολο διαρρυθμιστών οι οποίοι τροποποιούν την πρόσβαση στο μνημονικό περιεχόμενο και τον τρόπο με τον οποίο λειτουργεί η αντίληψη, η επιλογή και η εκτέλεση δράσεων. Και αυτό το χαρακτηριστικό λείπει από το IDyOT, το οποίο διαθέτει μόνο έναν τρόπο λειτουργίας.
- Ένας μηχανισμός ενισχυτικής μάθησης, ο οποίος δημιουργεί συσχετισμούς μεταξύ ορμών και γεγονότων, με βάση την εμπειρία. Η ενισχυτική μάθηση στο IDyOT βασίζεται μόνο σε στατιστικά μέτρα, όχι σε ορμές που αντικατοπτρίζουν ανάγκες.

Επιπλέον, ο Bach δίνει μία λίστα προϋποθέσεων για τη μνήμη, δηλαδή των δομών που η αρχιτεκτονική χρειάζεται να διαθέτει:

- Εικόνα κατάστασης (situation image) – Αποτελεί μοντέλο της τρέχουσας κατάστασης η οποία συμπεριλαμβάνει τον ίδιο τον πράκτορα. Στο IDyOT, το περιεχόμενο του οικουμενικού χώρου εργασίας μπορεί να αποτελέσει ανάλογο της εικόνας κατάστασης.
- Μακρόχρονη μνήμη (Long-term memory) – Περιλαμβάνει μοντέλο του εαυτού, δηλωτική και επεισοδιακή μνήμη. Στο IDyOT εντοπίζονται οι δύο τελευταίες, αλλά όχι το μοντέλο του εαυτού.
- Εσωτερική σκηνή (Inner stage) – Περιλαμβάνει υποθετικές αναπαραστάσεις, κυρίως προσδοκίες, στόχους και σχέδια. Το IDyOT παράγει προσδοκίες σε πολλαπλά επίπεδα αναπαράστασης, ωστόσο δεν διαθέτει ρητούς στόχους ή σχέδια.

Με βάση τις προϋποθέσεις για εφαρμογή του μοντέλου κινητοποίησης του MicroPSI σε άλλες αρχιτεκτονικές, και τις αποκλίσεις που παρουσιάζει το IDyOT από αυτές, μία πρόταση η οποία θα μπορούσε να συνδέσει τα μοντέλα είναι η εξής: από τις τρεις ομάδες αναγκών της θεωρίας PSI (φυσιολογικές, κοινωνικές, γνωστικές), οι γνωστικές ανάγκες προσφέρονται περισσότερο για υλοποίηση στο IDyOT, λόγω της αφηρημένης φύσης τους (δεν χρειάζονται ενσωμάτωση στον φυσικό κόσμο, ούτε περιβάλλον που περιλαμβάνει άλλους πράκτορες) και της σύνδεσής τους με τη δημιουργικότητα. Πιο αναλυτικά, η ανάγκη γνωστικής ικανότητας, εξερεύνησης και αισθητικής θα μπορούσαν να δράσουν σε υψηλό επίπεδο, ως κινητοποιητές και παραγωγείς συναισθημάτων, για τους μηχανισμούς χαμηλότερου επιπέδου που περιγράφονται στη θεωρία του IDyOT. Οι ανάγκες αυτές θα σηματοδοτούνταν στο σύστημα μέσω δεικτών ορμών και έτσι θα επηρέαζαν τη συμπεριφορά και τη μάθηση του συστήματος. Η επιρροή αυτή θα διεξαγόταν

από γνωστικούς διαρρυθμιστές (διέγερση, σθένος, επιθετικότητα, επίπεδο ανάλυσης, καταστολή, βαθμός ασφάλειας), που αποτελούν μηχανισμούς που αλλάζουν τον τρόπο με τον οποίο εκτελούνται γνωστικές διεργασίες. Οι διαρρυθμιστές αυτοί ορίζουν έναν χώρο θυμικών καταστάσεων, οι οποίες όταν κατευθύνονται σε ένα αντικείμενο με σημασία κινητοποίησης προάγονται σε συναισθήματα υψηλότερου επιπέδου.

Οι τρεις γνωστικές ανάγκες της θεωρίας PSI σχετίζονται στενά με έννοιες της θεωρίας του IDyOT. Πιο συγκεκριμένα, η γνωστική ικανότητα μπορεί να αποτελέσει έναν κινητοποιητή για καλύτερη απόδοση σε εξερευνητικές και αισθητικές δράσεις, ενώ μπορεί να αποτελέσει πηγή κινητοποίησης των λειτουργιών βελτιστοποίησης του συστήματος όταν λείπουν άλλοι κινητοποιητές. Η ανάγκη της γνωστικής ικανότητας συνιστά μία πιθανή εξήγηση της επιθυμίας των καλλιτεχνών να βελτιώνονται στην τεχνική τους κατάρτιση. Όσον αφορά την ανάγκη της εξερεύνησης, μπορούμε να την ταυτίσουμε με ώθηση προς την απόκτηση βεβαιότητας για αντικείμενα και διαδικασίες στο περιβάλλον του πράκτορα ενώ λειτουργεί και ως ευρετικό για την αβεβαιότητα μίας κατάστασης, σε άμεση αναλογία με την τιμή της εντροπίας στο IDyOT. Η ανάγκη αυτή ικανοποιείται από βέβαια γεγονότα, όπως για παράδειγμα την πλήρη ταυτοποίηση ενός αντικειμένου, μίας σκηνής ή ενός πλαισίου, τις εκπληρωμένες προσδοκίες και έναν μακρύ και μη-διακλαδιζόμενο ορίζοντα προσδοκιών. Επιδεινώνεται όταν ο πράκτορας συναντά άγνωστα αντικείμενα ή γεγονότα, όταν οι προσδοκίες του παραβιάζονται ή όταν ο τρέχον ορίζοντας προβλέψεων είναι υπερβολικά βραχύς ή διακλαδίζεται υπερβολικά. Τα σήματα αβεβαιότητας σταθμίζονται με την συνάφειά του αντικειμένου τους για την κινητοποίηση και ενεργοποιούν εξερευνητικές συμπεριφορές, εκτός αν ο πράκτορας έχει χαμηλή γνωστική ικανότητα για την ίδια την εργασία της εξερεύνησης (οι πράκτορες συνήθως στοχεύουν σε καταστάσεις με προσδόκιμο εξερεύνησης, αλλά εντός της ικανότητάς τους για επιτυχή εξερεύνηση). Σχετικά με την ανάγκη της αισθητικής, και οι δύο πτυχές της προσφέρονται για συνδυασμό με το IDyOT. Αναλυτικότερα, η αφηρημένη αισθητική έγκειται στην αναζήτηση δομών στις αναπαραστάσεις ενός πράκτορα, εργασία που είναι εφικτή με τους μηχανισμούς του IDyOT που σχετίζονται με τους εννοιολογικούς χώρους, τις φασματικές αναπαραστάσεις αλληλουχιών και τη μνημονική παγίωση. Η αισθητική που προσανατολίζεται σε ερεθίσματα έγκειται στην απόλαυση εγγενώς ευχάριστων ερεθισμάτων, και πιθανώς αποτελεί παράπλευρη συνέπεια άλλων γνωστικών μηχανισμών. Αυτή η ανάγκη ίσως αποτελεί εξήγηση της ευχαρίστησης που νιώθουν οι άνθρωποι όταν έρχονται σε επαφή με κάποια έργα τέχνης για πρώτη φορά, χωρίς να έχουν ήδη αναπαραστάσεις για αυτά, ή για τον γενικότερο καλλιτεχνικό τομέα από προηγούμενη εμπειρία.

Με βάση τα παραπάνω, μπορούμε να φανταστούμε τη δημιουργία ενός μουσικού πράκτορα, ο οποίος κινείται στον αφηρημένο κόσμο των μουσικών δομών (αφηρημένο με την έννοια ότι δεν περιλαμβάνει φυσικό χώρο ούτε άλλους πράκτορες). Ο πράκτορας αυτός θα αποτελούνταν σε θεμελιώδες επίπεδο από τις γεννήτριες του συστήματος IDyOT, και θα παρουσίαζε την ίδια στατιστική μάθηση, ικανότητα πρόβλεψης, λειτουργία της συνειδητής προσοχής και δομών της μνήμης καθώς και όλα τα αναδυόμενα από αυτά φαινόμενα (ικανότητα τεμαχιοποίησης, μη

συνειδητή δημιουργικότητα, κτλ). Ωστόσο, θα περιείχε συγχρόνως τις τρεις γνωστικές ανάγκες του συστήματος MicroPSI, οι οποίες θα καθόριζαν την κινητοποίηση, την ανάκληση μνήμης και τα συναισθήματα του πράκτορα. Μπορούμε να φανταστούμε πολλά ενδιαφέροντα πειράματα τα οποία θα μπορούσαν να γίνουν σε έναν τέτοιο πράκτορα, ανάλογα με το τι ερεθίσματα παρουσιάζονται στο περιβάλλον του (που ίσως να μπορούν να παραληφθούν αν μας ενδιαφέρει να δούμε τι δομές θα ανακαλύψει από μόνος του ο πράκτορας). Για παράδειγμα, θα μπορούσαμε να του παρουσιάσουμε μονοφωνική μουσική και να δούμε αν ανακαλύπτει την πολυφωνία από μόνος του. Θα μπορούσαμε να του παρουσιάσουμε μουσικά κομμάτια που θεωρούνται ευρέως “εγγενώς απολαυστικά”, και να παρατηρήσουμε αν ενεργοποιείται η ανάγκη της αισθητικής συγκεκριμένων ερεθισμάτων και αν αυτό έχει επίδραση στην ταχύτητα με την οποία ο πράκτορας ανακαλύπτει πιο περίπλοκες δομές. Ακόμα, θα μπορούσαμε να ερευνήσουμε τη συναισθηματική αντίδραση του πράκτορα σε μουσικές δομές και πως αυτή επηρεάζεται από τις ήδη υπάρχουσες δομές στη μνήμη του. Επιπλέον, θα μπορούσαμε να πειραματιστούμε με διαφορετικές τιμές των παραμέτρων των αναγκών και των διαρρυθμιστών, που αντιστοιχούν με διαφορετικές μουσικές προσωπικότητες, και να διαπιστώσουμε πως φανερώνονται οι διαφορές αυτές στη λειτουργία του πράκτορα. Φυσικά, ένας από τους σημαντικότερους ελέγχους για την ποιότητα του συστήματος ως γνωστικό μοντέλο της μουσικής νόησης είναι το μουσικό αποτέλεσμα που παράγει.

Τέλος, πρέπει να τονιστεί ότι η σύνδεση των δύο συστημάτων είναι ένα περίπλοκο εγχείρημα, και μένουν ακόμη πολλές λεπτομέρειες προς διευκρίνηση. Ωστόσο, το αποτέλεσμα αυτής της διαδικασίας ίσως να ρίξει φως στο πως λειτουργεί η ανθρώπινη αντίληψη και κινητοποίηση στον τομέα της δημιουργίας τέχνης, ποιοι μηχανισμοί πιθανώς ρυθμίζουν τη λειτουργία του πράκτορα και ποιος είναι ο ρόλος των συναισθημάτων στις γνωστικές αυτές εργασίες.

ΕΠΙΛΟΓΟΣ

Το θέμα της πτυχιακής αυτής εργασίας ήταν η γνωσιακή μουσικολογία και το υπολογιστικό σύστημα IDyOT. Στο πρώτο μέρος της εργασίας αναλύθηκαν διάφορα θέματα της γνωσιακής μελέτης της μουσικής μεταξύ των οποίων οι ιδιαιτερότητες της μουσικής που καθιστούν τη μοντελοποίησή της μοναδική, η σχέση μεταξύ μουσικής ομοιότητας και μνήμης, η νοητική αναπαράσταση του τόνου, η αντίληψη αλληλουχιών και η μουσική προσδοκία. Έγινε εισαγωγή στον κλάδο της υπολογιστική δημιουργικότητας, παρατέθηκαν διάφοροι ορισμοί της δημιουργικότητας και εισήχθησαν οι εννοιολογικοί χώροι. Παρουσιάστηκαν εν συντομία η θεωρία πληροφορίας και η μαθηματική διατύπωση των εννοιολογικών χώρων, ενώ εκτέθηκαν οι βασικές αρχές του τομέα της γνωσιακής αρχιτεκτονικής. Στο δεύτερο μέρος της εργασίας έγινε εκτενής παρουσίαση της λειτουργίας του συστήματος IDyOT ξεκινώντας από το κατώτερο επίπεδο και ανεβαίνοντας σταδιακά. Μετά από εισαγωγικά στοιχεία και πληροφορίες σχετικά με την επιστημονική προέλευση της θεωρίας, παρουσιάστηκαν οι γεννήτριες του συστήματος, ο ρόλος τους στην αναλογία της οπερατικής χορωδίας και η σύνδεσή τους με μεγέθη θεωρίας της πληροφορίας και τον οικουμενικό χώρο εργασίας. Ακολούθως, αναλύθηκαν ανώτερα χαρακτηριστικά της αρχιτεκτονικής όπως αναπαράσταση και μνήμη, η φασματική φύση των εννοιολογικών χώρων που χρησιμοποιεί στη δομή της μνήμης της, η διαδικασία της μνημονικής παγίωσης, ο ρόλος της πρόβλεψης και ο γνωσιακός κύκλος του συστήματος. Παρατέθηκαν επίσης στοιχεία για την γνωσιακή εγκυρότητα της θεωρίας και δείγματα από πεδία εφαρμογής της, όπως η τεμαχιοποίηση γλώσσας και μουσικής, το φαινόμενο της λεξιλογικής αμφισημίας, ο συγχρονισμός (entrainment) και η δημιουργικότητα. Το τρίτο μέρος της εργασίας είχε ως αντικείμενο της παρουσίαση της αρχιτεκτονικής MicroPSI, που εστιάζει σε λειτουργικές δομές κινητοποίησης και συναισθήματος μέσω της μοντελοποίησης θεμελιωδών αναγκών, ορμών που τις σηματοδοτούν, ειδικών δομών μνήμης και διαρρυθμιστών. Τέλος, θίχτηκε το ζήτημα πιθανού συνδυασμού των δύο αρχιτεκτονικών, το οποίο προσφέρεται για μελλοντική έρευνα.

Η εργασία, αναγκαστικά λόγω μεγέθους, περιελάμβανε μόνο μέρος της γνώσης της γνωσιακής μουσικολογίας που μπορεί να φανεί χρήσιμη για την υπολογιστική μοντελοποίηση της μουσικής νόησης, ενώ παραλήφθηκαν πολλά σημαντικά σχήματα και περαιτέρω εξηγήσεις για παρουσιαζόμενα θέματα, που θα έδιναν καθαρότερη εικόνα του κλάδου. Ακόμη, η μη χρήση υπολογιστικών μέσων (κώδικα) για την υλοποίηση των παρουσιαζόμενων γνωσιακών συστημάτων, ενδεχομένως καθιστά την εξήγηση του τρόπου λειτουργίας τους ελλιπή, και αποτελεί μελλοντική εργασία, μαζί με την πιθανότητα υλοποίησης του συνδυασμού των δύο μοντέλων. Είναι εμφανές στον συγγραφέα, πως η υπολογιστική προσέγγιση στη γνωσιακή μουσικολογία αποτελεί έναν τομέα που έχει μεγάλες προοπτικές να οδηγήσει σε σημαντική πρόοδο στην κατανόησή μας για τον τρόπο λειτουργίας του μυαλού, αλλά και να αναμειχθεί με τομείς όπως το Music Information Retrieval και τη ρομποτική με σκοπό να δημιουργήσει εργαλεία για χρήση σε κοινωνικά πλαίσια, βελτιώνοντας ανθρώπινες ζωές.

BIBΛΙΟΓΡΑΦΙΑ

- Amabile, T. M. (1983). *The Social Psychology of Creativity: A Componential Conceptualization*. 20.
- Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51(4), 355–365. <https://doi.org/10.1037/0003-066X.51.4.355>
- Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge University Press.
- Bach, J. (2012a). MicroPsi 2: The Next Generation of the MicroPsi Framework. In J. Bach, B. Goertzel, & M. Iklé (Eds.), *Artificial General Intelligence* (Vol. 7716, pp. 11–20). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35506-6_2
- Bach, J. (2012b). A Framework for Emergent Emotions, Based on Motivation and Cognitive Modulators: *International Journal of Synthetic Emotions*, 3(1), 43–63. <https://doi.org/10.4018/jse.2012010104>
- Bach, J. (2015). Modeling Motivation in MicroPsi 2. In J. Bieger, B. Goertzel, & A. Potapov (Eds.), *Artificial General Intelligence* (Vol. 9205, pp. 3–13). Springer International Publishing. https://doi.org/10.1007/978-3-319-21365-1_1
- Baddeley, A. D., & Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Bode, S., He, A. H., Soon, C. S., Trampel, R., Turner, R., & Haynes, J.-D. (2011). Tracking the Unconscious Generation of Free Decisions Using Ultra-High Field fMRI. *PLoS ONE*, 6(6), e21612. <https://doi.org/10.1371/journal.pone.0021612>
- Boden, M. A. (1998). *Creativity and artificial intelligence*. 10.
- Boden, M. A. (2009). *Computer Models of Creativity*. 12.

- Cai, Z., Goertzel, B., Zhou, C., Zhang, Y., Jiang, M., & Yu, G. (2012). Dynamics of a computational affective model inspired by Dörner's PSI theory. *Cognitive Systems Research*, 17–18, 63–80. <https://doi.org/10.1016/j.cogsys.2011.11.002>
- Cariani, P. (1999). Temporal Coding of Periodicity Pitch in the Auditory System: An Overview. *Neural Plasticity*, 6(4), 147–172. <https://doi.org/10.1155/NP.1999.147>
- Chella, A. (2015). A Cognitive Architecture for Music Perception Exploiting Conceptual Spaces. In F. Zenker & P. Gärdenfors (Eds.), *Applications of Conceptual Spaces* (pp. 187–203). Springer International Publishing. https://doi.org/10.1007/978-3-319-15021-5_10
- Chella, A., Coradeschi, S., Frixione, M., & Saf, A. (2004). *Perceptual Anchoring via Conceptual Spaces*.
- Chew, E. (2014). *Mathematical and Computational Modeling of Tonality* (Vol. 204). Springer US. <https://doi.org/10.1007/978-1-4614-9475-1>
- Cleary, J., & Witten, I. (1984). Data Compression Using Adaptive Coding and Partial String Matching. *IEEE Transactions on Communications*, 32(4), 396–402. <https://doi.org/10.1109/TCOM.1984.1096090>
- Colton, S., & Wiggins, G. A. (2012). Computational Creativity: The Final Frontier? *The Final Frontier*, 6.
- Conklin, D., & Witten, I. H. (1995). Multiple viewpoint systems for music prediction. *Journal of New Music Research*, 24(1), 51–73. <https://doi.org/10.1080/09298219508570672>
- Deliege, I. (1987). Grouping Conditions in Listening to Music: An Approach to Lerdahl & Jackendoff's Grouping Preference Rules. *Music Perception*, 4(4), 325–359. <https://doi.org/10.2307/40285378>
- Deliege, I., & Mélen, M. (2004). Cue abstraction in the representation of musical form. In *Perception and cognition of music* (pp. 374–397). Psychology Press.
- Dörner, D. (1999). *Bauplan für eine Seele*.

- Dörner, D. (2002). *Die Mechanik des Seelenwagens: Eine neuronale Theorie der Handlungsregulation*.
- Egermann, H., Pearce, M. T., Wiggins, G. A., & McAdams, S. (2013). Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. *Cognitive, Affective, & Behavioral Neuroscience*, 13(3), 533–553. <https://doi.org/10.3758/s13415-013-0161-y>
- Eshghi, A., Purver, M., & Hough, J. (2013). *Probabilistic induction for an incremental semantic grammar*. 11.
- Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, 7. <https://doi.org/10.3389/fnsys.2013.00068>
- Foerde, K., & Poldrack, R. A. (2009). *Procedural Learning in Humans*. 9.
- Forth, J., Agres, K., Purver, M., & Wiggins, G. A. (2016). Entraining IDyOT: Timing in the Information Dynamics of Thinking. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01575>
- Forth, J. C. (2012). *Cognitively-motivated geometric methods of pattern discovery and models of similarity in music*.
- Forth, J., Wiggins, G. A., & McLean, A. (2010). Unifying Conceptual Spaces: Concept Formation in Musical Creative Systems. *Minds and Machines*, 20(4), 503–532. <https://doi.org/10.1007/s11023-010-9207-x>
- Franklin, S., Madl, T., D’Mello, S., & Snider, J. (2014). LIDA: A Systems-level Architecture for Cognition, Emotion, and Learning. *IEEE Transactions on Autonomous Mental Development*, 6(1), 19–41. <https://doi.org/10.1109/TAMD.2013.2277589>
- Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. MIT Press.

- Gärdenfors, P. (2014). *Geometry of meaning: Semantics based on conceptual spaces*. The MIT Press.
- Gingras, B., Pearce, M. T., Goodchild, M., Dean, R. T., Wiggins, G., & McAdams, S. (2016). Linking melodic expectation to expressive performance timing and perceived musical tension. *Journal of Experimental Psychology: Human Perception and Performance*, 42(4), 594–609. <https://doi.org/10.1037/xhp0000141>
- Gobet, F., Lane, P., Croker, S., Cheng, P., Jones, G., Oliver, I., & Pine, J. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, 5(6), 236–243. [https://doi.org/10.1016/S1364-6613\(00\)01662-4](https://doi.org/10.1016/S1364-6613(00)01662-4)
- Goldwater, S., Griffiths, T. L., & Johnson, M. (2006). Contextual dependencies in unsupervised word segmentation. *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the ACL - ACL '06*, 673–680. <https://doi.org/10.3115/1220175.1220260>
- Grahn, J. A. (2012). Neural Mechanisms of Rhythm Perception: Current Findings and Future Perspectives. *Topics in Cognitive Science*, 4(4), 585–606. <https://doi.org/10.1111/j.1756-8765.2012.01213.x>
- Hansen, N. Chr., & Pearce, M. T. (2014). Predictive uncertainty in auditory sequence processing. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01052>
- Hélie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review*, 117(3), 994–1024. <https://doi.org/10.1037/a0019532>
- Honing, H. (2006). Computational Modeling of Music Cognition: A Case Study on Model Selection. *Music Perception*, 23(5), 365–376. <https://doi.org/10.1525/mp.2006.23.5.365>
- Honing, H. (Ed.). (2018). *The origins of musicality*. The MIT Press.
- Huron, D. B. (2006). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.

- Jacoby, N., Undurraga, E. A., McPherson, M. J., Valdés, J., Ossandón, T., & McDermott, J. H. (2019). Universal and Non-universal Features of Musical Pitch Perception Revealed by Singing. *Current Biology*, 29(19), 3229-3243.e12. <https://doi.org/10.1016/j.cub.2019.08.020>
- Jordanous, A. (2016). Four PPP Perspectives on computational creativity in theory and in practice. *Connection Science*, 28(2), 194–216. <https://doi.org/10.1080/09540091.2016.1151860>
- Jordanous, A., & Keller, B. (2016). Modelling Creativity: Identifying Key Components through a Corpus-Based Approach. *PLOS ONE*, 11(10), e0162959. <https://doi.org/10.1371/journal.pone.0162959>
- Juslin, P. N., & Sloboda, J. A. (Eds.). (2010). *Handbook of music and emotion: Theory, research, and applications*. Oxford University Press.
- Kemp, C., Perfors, A., & Tenenbaum, J. B. (2004). *Learning Domain Structures*. 7.
- Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, 105(31), 10687–10692. <https://doi.org/10.1073/pnas.0802631105>
- Kieras, D. E., & Meyer, D. E. (1995). *Predicting Human Performance in Dual-Task Tracking and Decision Making with Computational Models using the EPIC Architecture*.
- Kirby, S. (2000). Syntax Without Natural Selection: How Compositionality Emerges from Vocabulary in a Population of Learners. In C. Knight, M. Studdert-Kennedy, & J. Hurford (Eds.), *The Evolutionary Emergence of Language* (1st ed., pp. 303–323). Cambridge University Press. <https://doi.org/10.1017/CBO9780511606441.019>
- Laird, J. E., Newell, A., Rosenbloom, P. S., Nilsson, R. N., & Bobrow, D. G. (1987). *SOAR: An Architecture for General Intelligence*.

- Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the Growth of Knowledge* (1st ed., pp. 91–196). Cambridge University Press. <https://doi.org/10.1017/CBO9781139171434.009>
- Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, 10(2). <https://doi.org/10.1016/j.cogsys.2006.07.004>
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159. <https://doi.org/10.1037/0033-295X.106.1.119>
- Lerdahl, F. (1988). Tonal Pitch Space. *Music Perception*, 5, 315–350.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press.
- London, J. (2004). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press.
- MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. 640.
- Maslow, A. H., & Frager, R. (1987). *Motivation and personality* (3rd ed). Harper and Row.
- McAdams, S. (1987). Music: A science of the mind? *Contemporary Music Review*, 2(1), 1–61. <https://doi.org/10.1080/07494468708567053>
- Meyer, L. B. (1956). *Emotion and meaning in music* (Paperback ed., [Nachdr.]). Univ. of Chicago Press.
- Minsky, M. (1982). *Why People Think Computers Can't*. 13.
- Minsky, M., & Lee, J. (1988). *Society Of Mind*. Simon & Schuster. <https://books.google.gr/books?id=bLDLlIfRpdkC>
- Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.
- Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge University Press.

- Pearce, M., & Rohrmeier, M. (2012). Music Cognition and the Cognitive Sciences. *Topics in Cognitive Science*, 17.
- Pearce, M. T. (2005). *THE CONSTRUCTION AND EVALUATION OF STATISTICAL MODELS OF MELODIC STRUCTURE IN MUSIC PERCEPTION AND COMPOSITION*. 283.
- Pearce, M. T., Müllensiefen, D., & Wiggins, G. A. (2010). The Role of Expectation and Probabilistic Learning in Auditory Boundary Perception: A Model Comparison. *Perception*, 39(10), 1367–1391. <https://doi.org/10.1068/p6507>
- Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., & Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *NeuroImage*, 50(1), 302–313. <https://doi.org/10.1016/j.neuroimage.2009.12.019>
- Pinker, S. (2009). *How the mind works* (Norton pbk). Norton.
- Rhodes, M. (1961). An Analysis of Creativity. *The Phi Delta Kappan*, 42(7), 305–310.
- Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2006). Colour categories and category acquisition in Himba and English. In N. Pitchford & C. P. Biggam (Eds.), *Progress in Colour Studies* (pp. 159–172). John Benjamins Publishing Company. <https://doi.org/10.1075/z.pics2.14rob>
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. Pearson Education. <https://www.vlebooks.com/vleweb/product/openreader?id=none&isbn=9781292401171>
- Ruwet, N., & Everist, M. (1987). Methods of Analysis in Musicology. *Music Analysis*, 6(1/2), 3. <https://doi.org/10.2307/854214>
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species. *Current Biology*, 19(10), 831–836. <https://doi.org/10.1016/j.cub.2009.03.061>

- Shanahan, M. (2010). *Embodiment and the inner life Cognition and Consciousness in the Space of Possible Minds*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199226559.001.0001>
- Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shepard, R. N. (1982). Structural Representations of Musical Pitch. In *Psychology of Music* (pp. 343–390). Elsevier. <https://doi.org/10.1016/B978-0-12-213562-0.50015-2>
- Shiffrin, M., & Atkinson, R. C. (1968). Human Memory: A Proposed System and its Control Processes. *Human Memory*, 107.
- Taine, H., & Haye, T. D. (1872). *On Intelligence* (Issue v. 1). Holt & Williams.
<https://books.google.gr/books?id=gXhEAQAAMAAJ>
- Temperley, D. (2013). Computational Models of Music Cognition. In *The Psychology of Music* (pp. 327–368). Elsevier. <https://doi.org/10.1016/B978-0-12-381460-9.00008-0>
- Tulving, E. (1972). Organization of memory. *Episodic and Semantic Memory*.
<https://ci.nii.ac.jp/naid/10017350531/en/>
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4), 327–352.
<https://doi.org/10.1037/0033-295X.84.4.327>
- Varma, S. (2011). Criteria for the Design and Evaluation of Cognitive Architectures. *Cognitive Science*, 35(7), 1329–1351. <https://doi.org/10.1111/j.1551-6709.2011.01190.x>
- Vernon, D. (2014). *Artificial cognitive systems: A primer*. The MIT Press.
- Volk, A., Wiering, F., & van Kranenburg, P. (2011). Unfolding the potential of Computational Musicology. *Problems and Possibilities of Computational Humanities*. Thirteenth International Conference on Informatics and Semiotics in Organisations, Leeuwarden, The Netherlands.
- Wallas, G. (1926). *The art of thought*. Solis Press.

- Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7), 449–458. <https://doi.org/10.1016/j.knosys.2006.04.009>
- Wiggins, G. A. (2007). Models of musical similarity. *Musicae Scientiae*, 11(1_suppl), 315–338. <https://doi.org/10.1177/102986490701100112>
- Wiggins, G. A. (2012). The Mind's Chorus: Creativity Before Consciousness. *Cognitive Computation*, 4(3), 306–319. <https://doi.org/10.1007/s12559-012-9151-6>
- Wiggins, G. A. (2020). Creativity, information, and consciousness: The information dynamics of thinking. *Physics of Life Reviews*, 34–35, 1–39. <https://doi.org/10.1016/j.plrev.2018.05.001>
- Wiggins, G. A., & Forth, J. (2015). IDyOT: A Computational Theory of Creativity as Everyday Reasoning from Learned Information. In M. Schorlemmer & A. Smill (Eds.), *Computational Creativity Research: Towards Creative Machines* (Vol. 7, pp. 127–148). Atlantis Press. https://doi.org/10.2991/978-94-6239-085-0_7
- Wiggins, G. A., & Sanjekdar, A. (2019). Learning and Consolidation as Re-representation: Revising the Meaning of Memory. *Frontiers in Psychology*, 10, 802. <https://doi.org/10.3389/fpsyg.2019.00802>
- Wixted, J. T. (2004). The Psychology and Neuroscience of Forgetting. *Annual Review of Psychology*, 55(1), 235–269. <https://doi.org/10.1146/annurev.psych.55.090902.141555>
- Xiao, P., Toivonen, H., Gross, O., Cardoso, A., Correia, J., Machado, P., Martins, P., Oliveira, H. G., Sharma, R., Pinto, A. M., Díaz, A., Francisco, V., Gervás, P., Hervás, R., León, C., Forth, J., Purver, M., Wiggins, G. A., Miljković, D., ... Battersby, S. (2019). Conceptual Representations for Computational Concept Creation. *ACM Computing Surveys*, 52(1), 1–33. <https://doi.org/10.1145/3186729>