

Υψηλού επιπέδου υλοποίηση των αλγορίθμων HTM  
με έμφαση στη μαθηματική διατύπωση ως βάση για  
περαιτέρω μελέτη

Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών

Κωνσταντίνος Σαμαράς-Τσακίρης

2019-06-05

# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>3</b>
1.1	Στόχος της εργασίας . . . . .	3
1.2	Θεμελιώνοντας την έννοια της νοημοσύνης . . . . .	3
1.3	Γιατί μελετούμε την HTM; . . . . .	5
1.3.1	Φυσικοί αλγόριθμοι . . . . .	6
1.3.2	HTM ως μοντέλο του εγκεφαλικού νεοφλοιού . . . . .	6
1.4	Επιλογή της γλώσσας Julia . . . . .	8
1.4.1	Σύντομη παρουσίαση της γλώσσας Julia . . . . .	9
<b>2</b>	<b>Το μοντέλο HTM</b>	<b>12</b>
2.1	Το πρόβλημα της πρόβλεψης ακολουθιών . . . . .	12
2.2	Στοιχεία της Hierarchical Temporal Memory . . . . .	13
2.2.1	Αραιές Διανεμημένες Αναπαραστάσεις (SDR) . . . . .	14
2.2.2	Μοντέλο νευρώνα . . . . .	14
2.2.3	Μοντέλο δικτύου . . . . .	15
2.2.4	Encoders . . . . .	16
2.2.5	Decoders . . . . .	16
2.3	Βασικές λειτουργίες HTM . . . . .	16
2.3.1	Sequence Memory . . . . .	16
2.4	Αποτελέσματα - Προσομοιώσεις . . . . .	17

# Κεφάλαιο 1

## Εισαγωγή

### 1.1 Στόχος της εργασίας

Σε αυτήν την εργασία μελετάται μια αλγοριθμική θεωρία νοημοσύνης, η *Hierarchical Temporal Memory*. Βασικοί αλγόριθμοι της θεωρίας διατυπώνονται σε υψηλού επιπέδου γλώσσα με εκφραστική περιεκτικότητα, αποσκοπώντας στην καταπολέμηση της βασικής μορφής πολυπλοκότητας που απαντά σε αυτό το μοντέλο: την *εκφραστική πολυπλοκότητα* [34].

Με προγραμματιστική διατύπωση πιστή στη μαθηματική διατύπωση των αλγορίθμων, το έργο αυτό φιλοδοξεί να θεμελιώσει και να διευκολύνει την περαιτέρω μελέτη ενός συστήματος που, πέρα από το επιστημονικό του ενδιαφέρον, προσφέρεται για την αντιμετώπιση δύσκολων προβλημάτων τεχνητής νοημοσύνης. Η *Hierarchical Temporal Memory* βρίσκεται σε φάση ενεργούς έρευνας και ανάπτυξης. Καθώς οι αλγόριθμοι που την περιγράφουν εξελίσσονται, μια πλατφόρμα που επιτρέπει ταχύ πειραματισμό σε επεκτάσεις και εναλλακτικές ιδέες μπορεί να διαδραματίσει σημαντικό ρόλο στην περαιτέρω μελέτη της θεωρίας.

### 1.2 Θεμελιώνοντας την έννοια της νοημοσύνης

Κάθε φυσικό σύστημα καθορίζεται από τους περιορισμούς στα σύνορά του. Ομοίως ο άνθρωπος καθορίζεται από την αλληλεπίδραση με το περιβάλλον του. Στην προσπάθεια να κατανοήσουμε τον άνθρωπο, το πώς λειτουργεί, το γιατί δρα με συγκεκριμένο τρόπο, σταθμό αποτελεί η κατανόηση της συμπεριφοράς που καλούμε *νοημοσύνη*.

Η συμπεριφορά είναι παρατηρήσιμη και μας δίνει μια οπτική στην εσωτερική κατάσταση, στις κρυφές μεταβλητές ενός συστήματος, οπότε ίσως αποτελεί το σημείο όπου μπορούμε να πιάσουμε το νήμα της αναζήτησης. Είναι όμως ικανοποιητικό να χαρακτηρίσουμε τη νοημοσύνη συμπεριφορά;

Ο όρος νοημοσύνη χαίρει ευρείας ερμηνείας, ευρισκόμενος στο σταυροδρόμι πολλών επιστημονικών πεδίων, από την ψυχολογία μέχρι επιστήμη υπολογιστών [11]. Όλες μάλιστα οι πιο συγκεκριμένες ερμηνείες προσπίπτουν στο ιδιαίτερα ασαφές νόημα του όρου στην καθομιλουμένη. Ίσως μπορούμε να ξεμπλέξουμε για πρακτικούς σκοπούς αυτό το κουβάρι, παρατηρώντας ότι νοημοσύνη σίγουρα επιδεικνύουν ζωντανοί οργανισμοί ως εξελικτικό χαρακτηριστικό, ως εργαλείο στη διαρκή προσαρμογή τους στο επίσης δυναμικό τους περιβάλλον.

## Νοημοσύνη ως προσαρμογή

Η νοημοσύνη λοιπόν, ως προσαρμογή, προκύπτει από τη σχέση του οργανισμού με το περιβάλλον του. Αν ο οργανισμός προσαρμόζεται ταιριάζοντας υλικά του κατασκευάσματα στο περιβάλλον του, η νοημοσύνη επεκτείνει αυτήν τη δημιουργικότητα, επιτρέποντάς του να πειραματιστεί εικονικά. [2, σελ 3]. Ας ακολουθήσουμε όμως τη σκέψη του ψυχολόγου Jean Piaget στο ζήτημα.

Προσαρμογή είναι μια δυναμική διαδικασία αλληλεπίδρασης του οργανισμού με το περιβάλλον που περιλαμβάνει 2 στάδια: *αφομοίωση* και *συμβιβασμό*. Έστω ότι ο οργανισμός μπορεί να περιγραφεί με μια σειρά εσωτερικών μεταβλητών  $\{a, b, c\}$  και εξωτερικών στοιχείων του περιβάλλοντος  $\{x, y, z\}$ , που συνδέονται μεταξύ τους με κάποιες διαδικασίες, ορίζοντας ένα μοντέλο (schema):

$$a + x \rightarrow b$$

$$b + y \rightarrow c$$

$$c + z \rightarrow a$$

Η *αφομοίωση* συνίσταται στην ικανότητα του οργανισμού να ενσωματώνει τα στοιχεία του περιβάλλοντος στις εσωτερικές του καταστάσεις και να συνεχίζει αυτές τις διαδικασίες. Μια αλλαγή όμως στο περιβάλλον, έστω  $x \rightarrow x'$ , αποτελεί πρόκληση. Είτε ο οργανισμός δεν προσαρμόζεται, που σημαίνει ότι ο κύκλος σπάει και ο οργανισμός παύει να επιτελεί κάποια λειτουργία του, είτε προσαρμόζεται, τροποποιώντας υποχρεωτικά το μοντέλο του για να *συμβιβαστεί* με τη νέα εξωγενή πραγματικότητα ( $b \rightarrow b'$ ):

$$a + x' \rightarrow b'$$

$$b' + y \rightarrow c$$

$$c + z \rightarrow a$$

Σύμφωνα με αυτήν την περιγραφή, *προσαρμογή* είναι η ισορροπία της *αφομοίωσης* ( $a + x \rightarrow b \dots$ ) με το *συμβιβασμό* ( $x \rightarrow x' \Rightarrow b \rightarrow b'$ ).

Η προηγούμενη περιγραφή ισχύει εξίσου για τη *νοημοσύνη*. Νοημοσύνη είναι αφομοίωση, στο βαθμό που συμπεριλαμβάνει όλα τα εμπειρικά δεδομένα στη δομή της. Είναι όμως και συμβιβασμός, καθώς κατά τη διαρκή αφομοίωση αποκρίνεται στην πρόκληση των περιβαλλοντικών αλλαγών με τροποποίηση του μοντέλου που περιγράφει τον κόσμο.

Η διανοητική προσαρμογή λοιπόν, όπως κάθε προσαρμογή, συνίσταται από ένα μηχανισμό αφομοίωσης και συμπληρωματικού συμβιβασμού που διατηρούνται σε διαρκή ισορροπία. Ένα μυαλό προσαρμοσμένο στην πραγματικότητα είναι αυτό που δε δέχεται πια προκλήσεις στο νοητικό του μοντέλο για τον κόσμο, που δε χρειάζεται να τροποποιήσει περαιτέρω το μοντέλο αυτό για να εξηγήσει την εξελισσόμενη πραγματικότητα [2, σελ 5-7].

## Ένας πρακτικός ορισμός

Από το συμπεριφορικό ορισμό του Turing στο γνωστό "Turing test" μέχρι το μηχανιστικό ορισμό του Piaget, και με πολλές στάσεις ενδιάμεσα στο [Lenat 5] και στο [Minsky 4], η πρακτικότητα της έννοιας τίθεται υπό αμφισβήτηση.

Ένας άξονας του ορισμού είναι η σχέση της νοημοσύνης με τη λογική. Στο [35] ο Wang χωρίζει τα συλλογιστικά συστήματα σε 3 κατηγορίες:

- Αμιγώς αξιωματικά. Όλες οι λογικές προτάσεις προκύπτουν από τα αξιώματα, χαρακτηριστικό παράδειγμα η ευκλείδεια γεωμετρία.

- Μερικώς αξιωματικά. Η γνώση δεν επαρκεί σε όλες τις περιπτώσεις και υπάρχει μηχανισμός προσαρμογής, όπως στα ασαφή συστήματα.
- Μη αξιωματικά. Χτίζονται με βάση την υπόθεση ότι η γνώση ή οι πόροι δεν επαρκούν για οποιοδήποτε συλλογισμό.

Προτείνει λοιπόν ότι η απαίτηση νοημοσύνης ισοδυναμεί με την απαίτηση μη αξιωματικού συλλογιστικού συστήματος, λόγω της ανεπάρκειας πληροφορίας και πόρων.

Προσθέτοντας άλλο ένα έρεισμα στη συζήτηση, η νοημοσύνη συσχετίζεται άμεσα με τη δημιουργικότητα [15]. Δημιουργικότητα είναι η ικανότητα παραγωγής καινοτόμων και χρήσιμων ιδεών, επομένως συνδέεται με την τέλεση των προαναφερθέντων συμβιβασμών.

Σε αναζήτηση ενός χρήσιμου ορισμού στα πλαίσια αυτής της εργασίας, μπορούμε να στραφούμε στον ορισμό του [Wang 7]:

Νοημοσύνη είναι η ικανότητα ενός συστήματος επεξεργασίας πληροφορίας να προσαρμόζεται στο περιβάλλον του με ανεπαρκή γνώση και πόρους.

Αναφερόμαστε σε σύστημα επεξεργασίας πληροφορίας, για να μπορούμε να μελετήσουμε την εσωτερική του κατάσταση και αλληλεπίδραση με το περιβάλλον αφηρημένα, σε αντιδιαστολή με ένα πρόβλημα π.χ. ρομποτικής. Το σύστημα έχει μια γλώσσα εισόδου και εξόδου, με την οποία εκφράζονται τα ερεθίσματα του περιβάλλοντος και οι δράσεις του συστήματος. Το σύστημα συνήθως έχει κάποιο σκοπό για τον οποίο παράγει δράσεις σύμφωνα με τη γνώση του, και για την επεξεργασία του δαπανά περιορισμένους πόρους.

Η προσαρμογή μπορεί να ερμηνευθεί όπως προηγουμένως, ή πιο συνοπτικά ως ότι το σύστημα μαθαίνει από τις εμπειρίες του.

Ο περιορισμός της ανεπαρκούς γνώσης και πόρων σημαίνει ότι το σύστημα υπόκειται σε αυτές τις συνθήκες:

- *Περιορισμένο ως προς τους υπολογιστικούς του πόρους*
- *Λειτουργεί σε πραγματικό χρόνο*
- *Καλύπτει όλο το πεδίο των εκφράσιμων στη γλώσσα του εισόδων και εξόδων (δεν υπάρχουν άκυρες είσοδοι/έξοδοι)*

Σύμφωνα με αυτόν τον ορισμό νοημοσύνη είναι μια ισχυρή μορφή προσαρμογής.

Οι παραπάνω περιορισμοί μας οδηγούν προς την εξέταση *συστημάτων ροής (streaming)*, με *ανθεκτικότητα σε σφάλματα* και με *απόδοση που να κλιμακώνεται με τους διαθέσιμους πόρους* — υπό την προϋπόθεση της *διαρκούς προσαρμογής σε νέες συνθήκες* του περιβάλλοντος.

Η αυλαία σηκώνεται για να παρουσιαστεί το υπό μελέτη σύστημα: **Hierarchical Temporal Memory (HTM) της Numenta.**

### 1.3 Γιατί μελετούμε την ΗΤΜ;

Ο πρακτικός ορισμός της νοημοσύνης επιβάλλει περιορισμούς στο τι σύστημα θα θεωρήσουμε ότι επιδεικνύει χαρακτηριστικά νοημοσύνης. Για παράδειγμα, τα συστήματα εμπειρογνομώνων δεν πληρούν αρκετές από τις προδιαγραφές, ενώ πολλά συστήματα νευρωνικών δικτύων επιβλεπόμενης μάθησης που βρίσκονται τώρα σε χρήση επίσης προσπίπτουν τουλάχιστον στην προδιαγραφή της διαρκούς προσαρμοστικότητας.

### 1.3.1 Φυσικοί αλγόριθμοι

Ακολουθώντας τη λογική του Chazelle [34], η επιτυχία της φυσικής του 20ου αιώνα είναι σε μεγάλο βαθμό η επιτυχία της μαθηματικής έκφρασης. Με ένα μικρό σύνολο εξισώσεων μπορούμε να περιγράψουμε τι συμβαίνει στο φυσικό κόσμο. Αυτή η διαπίστωση αν μη τι άλλο σκιαγραφεί τις αρχές που διέπουν το φυσικό κόσμο: συμμετρία και κανονικότητα, η αμβροσία της συνήθους μαθηματικής διατύπωσης. Αν αυτή η παρατήρηση ήταν καθολική, τα ίδια εργαλεία θα επαρκούσαν για να περιγράψουν όλα τα επιστημονικά πεδία.

Η βιολογία διέπεται προφανώς από τους ίδιους φυσικούς νόμους και συνίσταται στην εφαρμογή τους επανειλημμένα στο βάθος των αιώνων. Παρόλο που οι αρχές είναι οι ίδιες, φαινομενολογικά η υπόθεση εργασίας είναι αντίστροφη: κάθε φαινόμενο είναι ειδικό και ξεχωριστό, αλλοιώσιμο υπό κάθε μετασχηματισμό, εκτός από ορισμένες περιπτώσεις που μπορούν να κατηγοριοποιηθούν μαζί. Γιατί οι αρχές επιτρέπουν καταστάσεις διακλάδωσης στα στοιχειώδη συστήματα που περιγράφουν και διάσπαση της συμμετρίας [27]. Έτσι, ο αναγωγισμός δε συνεπάγεται εποικοδομητισμό [3]. Με μια κομψή έκφραση: *η ιστορία είναι ο μεγάλος διασπαστής της συμμετρίας* [34].

Δίχως την απλοποιητική επιρροή της συμμετρίας, αυτή η οπτική γωνία έχει να αντιμετωπίσει μια μορφή πολυπλοκότητας διαφορετική από αυτήν που συνήθως ορίζουμε στην επιστήμη υπολογιστών: *εκφραστική πολυπλοκότητα*. Είναι ωφέλιμο αντιστοίχως να χρησιμοποιήσουμε και διαφορετική γλώσσα για τη μελέτη αυτών των φαινομένων: *τους (φυσικούς) αλγορίθμους* [34]. Παραδείγματα φυσικών αλγορίθμων εξερευνούνται στα [14], [18]

Συνδέοντας τον πρακτικό ορισμό της νοημοσύνης και τη χρησιμότητα των αλγορίθμων για να περιγράψουν *εκφραστικά πολύπλοκα* φαινόμενα, μπορούμε να επιχειρήσουμε τη μελέτη της νοημοσύνης με μια αλγοριθμική της θεώρηση. Αυτό είναι το βασικό κίνητρο για το μοντέλο που μελετά αυτή η εργασία, τη *Hierarchical Temporal Memory* (ΗΤΜ).

### 1.3.2 ΗΤΜ ως μοντέλο του εγκεφαλικού νεοφλοιού

#### Εποπτική εγκεφαλική ανατομία

Ο ανθρώπινος εγκέφαλος χωρίζεται σε διακριτές δομές με διαφορές τόσο στο μορφολογικό, όσο και στο λειτουργικό επίπεδο. Ένας χρήσιμος τέτοιος διαχωρισμός συνίσταται από τα εξής τμήματα:

- Πρόσθιος/διάμεσος εγκέφαλος, που περιλαμβάνει τα εγκεφαλικά ημισφαίρια, τους θαλάμους, τους υποκάμπους κ.α.
- Παρεγκεφαλίδα, που εντοπίζεται στο πίσω μέρος του κρανίου κάτω από τα ημισφαίρια
- Εγκεφαλικό στέλεχος, που εντοπίζεται επίσης προς τα πίσω και είναι η προέκταση του νωτιαίου μυελού

Σε αυτό το σημείο αξίζει μία πολύ συνοπτική επισκόπηση των εγκεφαλικών δομών, για την καλύτερη κατανόηση της ΗΤΜ.

Το εγκεφαλικό στέλεχος είναι το εξελικτικά αρχαιότερο τμήμα του εγκεφάλου και σχετίζεται με βασικές ομοιοστατικές λειτουργίες. Το μεταιχμιακό σύστημα στη βάση των ημισφαιρίων αναπτύχθηκε πριν από περίπου 250 εκατομμύρια χρόνια στα θηλαστικά και μία από τις βασικές του λειτουργίες είναι η ρύθμιση

των συναισθημάτων. Αυτές οι δομές έχουν μορφολογία πυρηνική, δηλαδή τα σώματα των νευρώνων τους συγκεντρώνονται σε σφαιροειδείς δομές από τις οποίες εκτείνονται οι άξονές τους.

Η επιφάνεια των εγκεφαλικών ημισφαιρίων είναι ο εγκεφαλικός φλοιός, με το μεγαλύτερο μέρος του να αποτελεί το νεοφλοιό, όπου οι νευρώνες (τα σώματα) είναι δομημένοι σε 6 επίπεδα, και σε μικρό μέρος τον αλλοφλοιό, που έχει 3 επίπεδα νευρώνων. Οι άξονες των νευρώνων φεύγουν από το επίπεδο του φλοιού σαν καλώδια σε πλακέτα. Ο φλοιός είναι το εξελικτικά πιο σύγχρονο τμήμα του εγκεφάλου, και ειδικά ο νεοφλοιός, που απαντά μόνο σε θηλαστικά. Εδώ εντοπίζονται λειτουργίες σχετικές με "ανώτερη συλλογιστική" και αφηρημένη σκέψη. Η παρεγκεφαλίδα επίσης έχει τη μορφή φλοιού, αλλά είναι διακριτή από τα ημισφαίρια. Σχετίζεται τουλάχιστον με τη ρύθμιση λεπτών κινήσεων. Και οι δύο αυτές δομές που οργανώνονται σε φλοιούς, αντί για πυρήνες, μοιράζονται ένα γεωμετρικό πλεονέκτημα: το μέγεθός τους μπορεί να κλιμακωθεί ευκολότερα. Αποτέλεσμα της κλιμάκωσης του μεγέθους τους είναι οι αναδιπλώσεις στην επιφάνεια του ανθρώπινου εγκεφάλου. Στον άνθρωπο ο νεοφλοιός αποτελεί περίπου τα 3/4 όλου του εγκεφάλου.

Στον τομέα της μηχανικής μάθησης εμφανίζονται 3 βασικά μοντέλα μάθησης: επιβλεπόμενη, μη επιβλεπόμενη και ενισχυτική. Υπάρχουν επιχειρήματα στη βιβλιογραφία [9] ότι 3 από τις εγκεφαλικές δομές που περιγράφηκαν εφαρμόζουν αυτά τα 3 μοντέλα αντίστοιχα: η παρεγκεφαλίδα, ο εγκεφαλικός φλοιός και τα βασικά γάγγλια.

Σημαντικά ανατομικά στοιχεία του φλοιού είναι η οργάνωση των νευρώνων (εννοώντας των σωμάτων των νευρώνων) σε επίπεδα και οι αναδρομικές συνδέσεις μεταξύ τους. Έχει επίσης παρατηρηθεί ότι η πλαστικότητα των νευρικών συνάψεων στο φλοιό ακολουθεί κανόνα Hebbian: ισχυροποιούνται όταν το προσυναπτικό ερέθισμα συσχετίζεται με μετασυναπτική δραστηριότητα και εξασθενούν αλλιώς, χτίζοντας την αιτιώδη σχέση μεταξύ προσυναπτικής και μετασυναπτικής δραστηριότητας. Διατυπώνεται έτσι η υπόθεση ότι ο φλοιός μαθαίνει με μη επιβλεπόμενο τρόπο να οργανώνει την εξωγενή και εσωτερική πραγματικότητα σε έννοιες, να δημιουργεί συμβολικές αναπαραστάσεις για μεταβλητές κατάστασης.

Με δεδομένη αυτήν τη βασική ανατομία, μπορούμε να διαπιστώσουμε τι σκοπεύει να μοντελοποιήσει η HTM.

## Στόχος της HTM

Η θεωρία της Hierarchical Temporal Memory αναπτύσσεται από την ερευνητική εταιρία Numenta, που δηλώνει το διττό της στόχο ως εξής: καταρχήν, την αλγοριθμική μοντελοποίηση της λειτουργίας του ανθρώπινου εγκεφαλικού νεοφλοιού, και ως συνέπεια τη μελέτη των εφαρμογών της θεωρίας τούτης ως σύστημα τεχνητής νοημοσύνης. Επομένως το μοντέλο που μελετά αυτή η εργασία δεν έχει σχεδιαστεί κατά κύριο λόγο ως σύστημα τεχνητής νοημοσύνης, αλλά ως *θεωρία της λειτουργίας του ανθρώπινου εγκεφαλικού νεοφλοιού, περιορισμένη από βιολογικά δεδομένα*. Παρόλα αυτά, εδώ δε θα συζητηθεί η νευροεπιστημονική πιστότητα του μοντέλου, μονάχα οι βιολογικές αρχές στις οποίες βασίζεται.

Σύμφωνα με τα παραπάνω ανατομικά στοιχεία, μοντελοποιώντας το νεοφλοιό η HTM δεν αποτελεί πλήρες μοντέλο του ανθρώπινου εγκεφάλου. Δεν προσφέρεται για παράδειγμα για μελέτη των συναισθημάτων ή των βασικών ομοιοστατικών μηχανισμών, αλλά μόνο των "ανώτερων συλλογιστικών". Η επιλογή αυτή δεν είναι τυχαία. Χάρη στην εξελικτική του νεότητα και γεωμετρική επεκτασιμότητα, ο νεοφλοιός δεν έχει υποστεί τόσες εξελικτικές βελτιστοποιήσεις, όσο τα υπόλοιπα

τμήματα του εγκεφάλου, και διατηρεί σε μεγάλο βαθμό κοινή μορφολογία σε όλη του την έκταση. Αυτή η παρατήρηση ενδεχομένως να καθιστά το πρόβλημα της διάκρισης των θεμελιωδών αρχών λειτουργίας του από τις εξελικτικές βελτιστοποιήσεις πολύ ευκολότερο, σε σχέση με άλλα τμήματα.

Ο δευτερεύων στόχος της HTM είναι αυτός με τον οποίο ασχολείται αυτή η εργασία. Συγκεκριμένα, η HTM μπορεί να χρησιμοποιηθεί για πρόβλεψη αιτιωδών ακολουθιών (πχ χρονοσειρών) και για αναγνώριση ανωμαλιών, ή γενικότερα για την αντιμετώπιση προβλημάτων πρόβλεψης ή κατηγοριοποίησης σε μη στάσιμες ροές δεδομένων. Έχει χρησιμοποιηθεί επιτυχώς ως τώρα (ως τεχνολογία βάσης κερδοσκοπικών επιχειρήσεων) για πρόβλεψη χρηματιστηριακών δεικτών και για έγκαιρη ανίχνευση ανωμαλιών σε κέντρα δεδομένων.

## 1.4 Επιλογή της γλώσσας Julia

Για την παρουσιαζόμενη υλοποίηση της HTM επιλέχθηκε η σχετικά καινούρια γλώσσα επιστημονικού προγραμματισμού Julia [26].

Προτού γίνει αυτή η επιλογή, δοκιμάστηκε η υλοποίηση της HTM σε Matlab. Όμως η Matlab αποτελεί ένα κλειστό οικοσύστημα, με λίγες προοπτικές για επαναχρησιμοποίηση και ευρύ απόηχο μιας τέτοιας δουλειάς, που εξαρχής στοχεύει στην υποβοήθηση περαιτέρω έρευνας. Η Matlab έχει μακρά ιστορία στο χώρο του επιστημονικού λογισμικού και γράφτηκε το 1984 στοχεύοντας ειδικά σε υπολογιστικούς επιστήμονες και όχι σε μηχανικούς λογισμικού. Πολλές σχεδιαστικές αποφάσεις ανακλούν αυτήν την εστίαση και δημιουργούν ένα περιβάλλον ανάπτυξης λογισμικού πιο δύσκολο σε σχέση με εναλλακτικές όπως η Python. Η γλώσσα διευκολύνει μεν τη χρήση γραμμικής άλγεβρας, αλλά δεν επιτρέπει συναρτησιακό προγραμματισμό, "lazy operations", ορισμό νέων τύπων δεδομένων και πολλά ακόμα στοιχεία απαραίτητα για την επιτυχία του κεντρικού στόχου αυτής της εργασίας: την εκφραστική απλότητα. Έτσι, η πρώτη υλοποίηση σε Matlab βοήθησε στην κατανόηση των αλγορίθμων, αλλά απέτυχε στον κεντρικό της στόχο.

Η Julia είναι ανοιχτό λογισμικό που ξεκίνησε από το JuliaLab του MIT το 2012 και αναπτύσσεται δημοσίως στο Github. Μόλις τον Αύγουστο του 2018 έφθασε στην πρώτη επίσημη έκδοσή της. Οι δημιουργοί της δηλώνουν ως κίνητρο για τη δημιουργία της την αντιμετώπιση του "προβλήματος των 2 γλωσσών" στην επιστημονική υπολογιστική: μία γλώσσα υψηλού επιπέδου, εύχρηστη αλλά όχι τόσο αποδοτική όπως η Python, χρησιμοποιείται αρχικά για την κατασκευή μιας πρωτότυπης λύσης· έπειτα το λογισμικό ξαναγράφεται σε μια πιο δύσκολη, αλλά αποδοτική γλώσσα, ενδεχομένως κατάλληλη για HPC ή για να αξιοποιήσει παραδοσιακές υπολογιστικές συστοιχίες, όπως η C++. Αυτή η ροή εργασίας είναι σύνθετη, αργή και αναποτελεσματική, απαιτώντας διπλή προσπάθεια και, συχνά, εξειδικευμένο προσωπικό. Η Julia επιδιώκει να αποτελέσει λύση σε αυτό το πρόβλημα, συνδυάζοντας την ευχρηστία της Python και την αποδοτικότητα της C++. Αν και νέα γλώσσα, έχει ήδη χτίσει ένα πλούσιο οικοσύστημα για επιστημονικό προγραμματισμό. Συμπεριλαμβάνει προφανώς τα βασικά όπως γραμμική άλγεβρα και αραιούς πίνακες. Σε μερικά πεδία όμως, όπως η αστρονομία, οι διαφορικές εξισώσεις και τα πολύπλοκα συστήματα, προσφέρει ήδη εξίσου πλήρεις ή πληρέστερες λύσεις από πιο παραδοσιακές γλώσσες, όπως η Python.

Μια επιτυχής υλοποίηση μεγάλου επιστημονικού λογισμικού που χαίρει ευρείας διαφήμισης για τη Julia είναι η Celeste [23]. Γραμμένο σε Julia, χρησιμοποιεί παραλληλισμό κοινής και κατανεμημένης μνήμης και μπόρεσε να αξιοποιήσει 8192 επεξεργαστικούς πυρήνες στον υπερυπολογιστή NERSC Cori.



### 1.4.1 Σύντομη παρουσίαση της γλώσσας Julia

Στην πράξη, η Julia καθιστά εύχρηστο ένα μικτό προγραμματιστικό μοντέλο προστακτικού και συναρτησιακού προγραμματισμού. Η δηλωτική φύση του συναρτησιακού προγραμματισμού μένει πιστή στη μαθηματική διατύπωση της θεωρίας και αυξάνει σημαντικά την εκφραστικότητα του κώδικα, προτρέποντας την αποσύνθεση πολύπλοκων ορισμών σε απλούστερους. Απελευθερώνει δε το πρόγραμμα από την έμμεση σειριοποίηση που επιβάλλει ο προστακτικός προγραμματισμός και διευκολύνει την αναδιάταξη του κώδικα, όπως και την κλιμάκωση της εκτέλεσής του σε περισσότερους υπολογιστικούς πόρους (“παραλληλοποίηση”). Ο προστακτικός προγραμματισμός οδηγεί σε ορισμένες περιπτώσεις σε πιο φυσική ή εύκολη διατύπωση της λύσης, ενώ σε άλλες λειτουργεί υπό μορφή βελτιστοποίησης.

Ο σχεδιασμός της Julia προσπαθεί να διευκολύνει τον προγραμματιστή, αν επιθυμεί να υλοποιήσει τη λύση του με τον πιο πρόχειρο τρόπο, και να του επιτρέψει να τη βελτιώσει και να την κάνει αποδοτική με περισσότερη προσπάθεια και φροντίδα. Πιστή σε αυτήν την αρχή, υιοθετεί προαιρετικό σύστημα τύπων (δε θα ήταν άστοχη η σύγκριση με typed λ-calculus).

Ένα πρόγραμμα Julia είναι κατά κανόνα ένα σύνολο ορισμών τύπων και τιμών. Στις τιμές μπορεί να αντιστοιχισθούν ονόματα, σύμβολα. Όταν δηλώνεται

```
julia> x= 1
1
```

η τιμή 1 αντιστοιχίζεται στο σύμβολο x. Έτσι ο τύπος του x είναι ο τύπος της τιμής του:

```
julia> x|> typeof
Int64

julia> x= 1.0
1.0

julia> x|> typeof
Float64
```

Σε όλην αυτήν την παρουσίαση δε χρησιμοποιήθηκε ο όρος “μεταβλητή”. Το x είναι απλώς ένα σύμβολο, που εδώ αντιστοιχίστηκε σε μια σταθερή, αμετάβλητη τιμή, και μετά επαναντιστοιχίστηκε σε μια διαφορετική σταθερή, αμετάβλητη τιμή.

Το x θα μπορούσε να αντιστοιχιστεί και σε μια όντως μεταβλητή τιμή, δηλαδή μία τιμή που επιτρέπεται να τροποποιηθεί:

```
julia> x= rand{Int8,5}
5-element Array{Int8,1}:
 31
  5
 86
 -3
 59

julia> x[5]= 0
0

julia> x
5-element Array{Int8,1}:
 31
  5
```

```
86
-3
0
```

Εξίσου, η τιμή που συμβολίζει το `x` μπορεί να είναι ένας τύπος

```
julia> x= Int64
Int64

julia> x|> typeof
DataType
```

ή μια συνάρτηση

```
julia> x= (i)-> i+1
#3 (generic function with 1 method)

julia> x|> typeof
getfield(Main, Symbol("##3#4"))

julia> x(5)
6
```

Ο μεταγλωττιστής εσωτερικά αποδίδει τύπο σε κάθε τιμή που απαντά στο πρόγραμμα. Ο προγραμματιστής δε χρειάζεται να συσχετίσει ευθέως σύμβολα με τύπους. Μπορεί όμως, αν θέλει, να χρησιμοποιήσει τύπους για ένα βασικό σκοπό: "πολλαπλή αποστολή" (multiple dispatch) μεθόδων συνάρτησης βάσει τύπων. Οι όροι "μέθοδος" και "συνάρτηση" σημαίνουν διαφορετικά πράγματα στη Julia. Παραπάνω ορίσαμε μία ανώνυμη συνάρτηση που υλοποιείται από μία μέθοδο. Θα μπορούσαμε όμως να ορίσουμε και συνάρτηση που να μην υλοποιείται από καμία μέθοδο:

```
julia> function myfun end
myfun (generic function with 0 methods)

julia> myfun()
ERROR: MethodError: no method matching myfun()
```

Ας προσθέσουμε 2 μεθόδους στη συνάρτηση, για να φανεί η "πολλαπλή αποστολή" κι η χρήση τύπων:

```
julia> myfun(i::Int)= print("I'm an Int")
myfun (generic function with 1 method)

julia> myfun(3)
I'm an Int

julia> myfun(i::Float64)= print("I'm a Double!")
myfun (generic function with 2 methods)

julia> myfun(3.0)
I'm a Double!

julia> myfun(3)
I'm an Int

julia> myfun()
ERROR: MethodError: no method matching myfun()
```

Μέθοδος λοιπόν είναι ο συνδυασμός μίας συνάρτησης και μιας πλειάδας (tuple) ορισμάτων. Αυτή η σχεδίαση επιτρέπει έναν πολυμορφισμό συγκρίσιμο με της C++.

Ο αναγνώστης θα κρίνει ο ίδιος την αποτελεσματικότητα των παραπάνω στον κορμό της εργασίας, όπου θα έρθει σε επαφή με αυτό το ιδίωμα γραφής.

## Κεφάλαιο 2

# Το μοντέλο HTM

### 2.1 Το πρόβλημα της πρόβλεψης ακολουθιών

Τα σύνορα της τεχνητής νοημοσύνης εκτείνονται πέρα από το παραδοσιακό πρόβλημα της εκμάθησης ενός στατικού συνόλου δεδομένων με επιβλεπόμενες μεθόδους, όπως γίνεται σε πολλές σύγχρονες εφαρμογές [22]. Η τεχνητή νοημοσύνη καλείται σήμερα να αξιοποιήσει τις καταγιστικές ροές δεδομένων που παρέχει το (ανθρωπογενές και μη) περιβάλλον, σε πραγματικό χρόνο, και δίχως την πολυτέλεια της προσήμανσης που απαιτεί η επιβλεπόμενη μάθηση — εφαρμογές IoT που απαιτούν δυναμική αλληλεπίδραση με το περιβάλλον τους έρχονται στο μυαλό [1].

Επομένως, καλούμαστε να μοντελοποιήσουμε έναν κόσμο που αλλάζει [25]. Το μοντέλο οφείλει ή να είναι γενικότερο από όλες τις δυνατές μεταβολές του κόσμου ή να αλλάζει μαζί του. Η έννοια της ροής δεδομένων που αναφέρθηκε υπονοεί την έννοια του χρόνου, της αλληλουχίας, που επιτρέπει στο μοντέλο να αντλήσει πληροφορία από την αιτιώδη σχέση. Την ερμηνεία του κόσμου με τη βοήθεια της συνεπαγωγής και της αναγνώρισης απαιτήσεων και συνεπειών.

Ορίζεται έτσι το πρόβλημα της εκμάθησης και πρόβλεψης ακολουθιών. Δηλαδή, ο πράκτορας που παρακολουθεί μια αλληλουχία δεδομένων (γεγονότων) καλείται να προβλέψει τη συνέχεια και να δράσει έτσι, ώστε βέλτιστα να ανταμειφθεί. Η ανταμοιβή, το κίνητρο και γενικότερα ο στόχος έχουν τεθεί από εξωτερικό παράγοντα και βρίσκονται εκτός του πλαισίου του προβλήματος. Άμεση εφαρμογή της πρόβλεψης είναι και η αναγνώριση ανωμαλιών.

Η εκμάθηση ακολουθιών είναι κλασικό πρόβλημα, για το οποίο έχουν αναπτυχθεί κλασικές λύσεις. Κυριότερο και ευρύτερα διαδεδομένο είδος μοντέλων, ειδικά πριν τη σύγχρονη επάνοδο των νευρωνικών δικτύων, είναι τα Hidden Markov Models. Τα κλασικά νευρωνικά δίκτυα προσφέρουν στο πρόβλημα τα νευρωνικά δίκτυα χρονικής καθυστέρησης (TDNN). Η ουσιαστική συνεισφορά των κλασικών νευρωνικών δικτύων επιτυγχάνεται όμως με τα ανάδρομα δίκτυα (RNN), που γενίκευσαν τις παλαιότερες προσθιοδρομικές (feedforward) αρχιτεκτονικές ακριβώς για να μπορούν να αντιμετωπίσουν εγγενώς προβλήματα ακολουθιών. Ιδιαίτερης μνείας χρήζει η δομή «μακράς βραχυπρόθεσμης μνήμης» (LSTM). Από το 2015 έχει χρησιμοποιηθεί με επιτυχία σε ποικίλες εφαρμογές πολύ πιο σύνθετες από την εισαγωγή που παρουσιάζεται εδώ, όπως ως τεχνητή νοημοσύνη που παίζει το παιχνίδι StarCraft 2 [AlphaStar 33].

Σε αυτόν το χώρο λύσεων υπάρχουν παρόλα αυτά σημεία για βελτίωση. Οι περισσότερες εφαρμογές, συμπεριλαμβανομένου του AlphaStar, βασίζονται σε επιβλεπόμενη και ενισχυτική μάθηση, αφήνοντας ευρύ πεδίο εξερεύνησης για μη επιβλεπόμενα μοντέλα. Καθώς ο όγκος των ροών δεδομένων αυξάνεται εκθετικά

[29], παρατηρείται αυξανόμενη ζήτηση για αλγορίθμους που προσαρμόζονται γρήγορα στις εξελισσόμενες στατιστικές των δεδομένων τους (συνεχής μάθηση), έχοντας πρόσβαση μόνο σε μικρό χρονικό παράθυρο πληροφορίας. Στην προσπάθεια αυτή αναπτύσσονται ενεργά τεχνικές για βελτίωση της αποδοτικότητας δειγμάτων της μάθησης [όπως 30], ή για παράκαμψη του προβλήματος, όπως η μεταφορική μάθηση (transfer learning) [32].

Στην ανάλυση των διαφόρων τρόπων ορισμού της τεχνητής νοημοσύνης, ο Wang [12] αναγνωρίζει τη μέθοδο «από τη δομή» για την έρευνα σε συστήματα που εμπνέονται ή μιμούνται το βασικό παράδειγμα ευφυούς συστήματος που επιλύει διαρκώς το παραπάνω πρόβλημα: τον εγκέφαλο, και ειδικά το φλοιό. Παρατηρείται μια τάση στο χώρο των τεχνητών νευρωνικών δικτύων ανασκόπησης της επαφής τους με τη βιολογική πραγματικότητα τα τελευταία χρόνια, όπως στο [16]. Το πιο ηχηρό παράδειγμα είναι ο μηχανισμός κάψουλας που πρότειναν ο Hinton και συνεργάτες [28], [31].

Σε αυτό λοιπόν το πλαίσιο, η μελέτη της HTM κρίνεται ιδιαίτερα καίρια.

## 2.2 Στοιχεία της Hierarchical Temporal Memory

Παρακάτω θα αναφερθούμε πάλι σε στοιχεία νευροεπιστήμης. Καθώς ο σκοπός της περιγραφής είναι η κατανόηση των αλγορίθμων HTM, πρέπει να γίνει αποδεκτή μια παρέκκλιση από την αυστηρότητα, χάριν απλότητας. Σε ό,τι βιολογικό στοιχείο αναφερθεί, ο αναγνώστης ας έχει υπόψιν ότι η πραγματικότητα είναι πάντα πιο πολύπλοκη και ότι αυτά που γράφονται εδώ δεν είναι ακριβώς ορθά. Είναι, όμως, *χρήσιμες προσεγγίσεις*.

Η κεντρική θέση στην οποία βασίζεται η Hierarchical Temporal Memory τμηματοποιεί τον εγκεφαλικό φλοιό σε ένα ψηφιδωτό βασικών μονάδων επεξεργασίας, των **“φλοιικών στηλών” (cortical columns)**, που έχουν την ίδια δομή και εκτελούν τους **ίδιους αλγορίθμους, αλλά σε διαφορετικά δεδομένα**. Η ιδέα αυτή έχει μακρά ιστορία στη νευροεπιστήμη, με μια συγκεντρωτική επισκόπηση από Defelipe, Markram κ.ά [13] να την οριοθετεί και να παρουσιάζει την ευρεία χρήση και κακομεταχείριση του όρου. Για παράδειγμα, η εισαγωγική πρόταση περί “ψηφιδωτού” παραπάνω πρέπει να αντιμετωπιστεί μόνο ως πρώτη προσέγγιση, γιατί οι στήλες δε φέρονται να έχουν γεωμετρικά σαφή σύνορα μεταξύ τους, αλλά διάχυτες ζώνες μετάβασης. Πρεσβευτής και βασικός αποκρυσταλλωτής της ιδέας είναι ο Mountcastle [8], με την ιδέα να χαίρει τόσο αποδοχής [20], όσο και κριτικής ως προς τη χρησιμότητά της [10]. Παρόλα αυτά, εδώ θα την υιοθετήσουμε. Έτσι, ο εγκεφαλικός φλοιός αποδομείται σε πλειάδα μονάδων επεξεργασίας κοινής αρχής, και οι αλγόριθμοι που θα παρουσιαστούν περιγράφουν τη λειτουργία της μονάδας.

Η φλοιική στήλη είναι λοιπόν ένας πληθυσμός νευρώνων με κοινή συνδεσμολογία: λαμβάνουν το σήμα εισόδου από κοινές πηγές και στέλνουν το σήμα εξόδου σε κοινούς παραλήπτες. Συχνά, τέτοιοι παραλήπτες είναι άλλες φλοιικές στήλες. Συντάσσονται έτσι επεξεργαστικές **ιεραρχίες**, με τα πρώτα στάδια της ιεραρχίας να δημιουργούν απλούστερα μοντέλα για τον κόσμο από τα μετέπειτα (κυρίως γιατί τα μετέπειτα συγκεντρώνουν περισσότερη πληροφορία).

Οι νευρώνες στους οποίους αναφερόμαστε παραπάνω, οι πυραμιδοειδείς νευρώνες, βρίσκονται κάθε στιγμή σε 1 από 2 καταστάσεις: ενεργοί ή ανενεργοί. Η ενεργοποίησή τους (“δυναμικό δράσης”) ερεθίζει άλλους νευρώνες με τους οποίους συνδέονται και μπορεί να τους οδηγήσει σε ενεργοποίηση. Μπορούμε λοιπόν να περιγράψουμε την κατάσταση του φλοιού κάθε στιγμή ως το σύνολο των νευρώνων

που είναι ενεργό. Προκύπτει ότι το σύνολο αυτό είναι πολύ μικρό ποσοστό του συνολικού νευρικού πληθυσμού, δηλαδή η ενεργοποίηση είναι **αραιή**, περίπου 2%.

Το μοτίβο ενεργοποίησης ενός πληθυσμού νευρώνων της ίδιας στήλης αποτελεί στο πλαίσιο της θεωρίας HTM τη δομή δεδομένων του εγκεφάλου και ονομάζεται **αραιή διανεμημένη αναπαράσταση (SDR)**. Κάθε αίσθηση στέλνει SDR στις στήλες που την επεξεργάζονται· κάθε στήλη στέλνει SDR στους μύες ή σε άλλες στήλες ως έξοδο. Η είσοδος σε ένα μοντέλο HTM πρέπει επομένως να μεταφράζει τη φυσική ποσότητα που θέλουμε να επεξεργαστούμε σε SDR, με τη διαδικασία να ονομάζεται **κωδικοποίηση**. Αντίστοιχα, η ερμηνεία ενός SDR εξόδου ονομάζεται **αποκωδικοποίηση**.

Από τους κοινούς αλγορίθμους που υλοποιεί κάθε φλοιική στήλη, η HTM αυτή τη στιγμή περιγράφει 2: τη **χωρική συγκέντρωση** και τη **χρονική μνήμη**. Η χρονική συγκέντρωση είναι επίσης διαδικασία που υποτίθεται, αλλά δεν έχει περιγραφεί επαρκώς και αποτελεί βασικό σημείο για περαιτέρω μελέτη.

### 2.2.1 Αραιές Διανεμημένες Αναπαραστάσεις (SDR)

Με βάση την παρατήρηση της αραιής πυροδότησης των νευρώνων στον εγκέφαλο σχεδιάζουμε ένα μεγάλο, αραιό, δυαδικό διάνυσμα που αναπαριστά την πυροδότηση ή μη κάθε νευρώνα της μελετούμενης περιοχής. Μπορεί να θεωρηθεί ότι κάθε περιοχή ανταλλάσσει τέτοια μεγάλα αραιά δυαδικά διανύσματα με τις επόμενες. Ουσιαστικά, αυτά τα αραιά διανύσματα αποτελούν τη “δομή δεδομένων” [6], [21] του εγκεφάλου. Το ιδιαίτερο χαρακτηριστικό αυτών των διανυσμάτων, λόγω αυτού που αναπαριστούν, είναι ότι κάθε ξεχωριστό bit τους φέρει σημασιολογικό νόημα – αντίθετα δηλαδή από την κωδικοποίηση ενός γράμματος σε κώδικα ASCII, όπου νόημα φέρει μόνο η πλήρης “λέξη”. Εξού και το “distributed” του ονόματος.

#### Χωρητικότητα SDR

Η χωρητικότητα ενός SDR με μέγεθος  $n$  και cardinality (πλήθος μονάδων)  $w$  είναι το πλήθος των διαφορετικών SDR με αυτή τη μορφή, δηλαδή οι συνδυασμοί των  $n$  ανά  $w$ :

$$\binom{n}{w} = \frac{n!}{w!(n-w)!}$$

#### Συνδυασμοί SDR

Διαφάνεια 9.

#### Σύνολα SDR

Διαφάνεια 10.

Επισημαίνεται μονάχα η **σχεδιαστική ελευθερία** που παρέχουν τα SDR. Όσο μεγαλύτερα SDR χρησιμοποιούνται, τόσο αυξάνεται ο χώρος των ενδεχομένων και τόσο πιο απίθανα γίνονται τα λανθασμένα ταιριάσματα.

### 2.2.2 Μοντέλο νευρώνα

Πρώτα απ' όλα: νευρώνα ονομάζουμε μια δομή που δέχεται σήματα από άλλους νευρώνες και, όταν αυτά πληρούν κάποιες προϋποθέσεις, ενεργοποιείται, για να στείλει σήματα στους επόμενους νευρώνες. Η επαφή 2 νευρώνων καλείται σύναψη.

Ο βιολογικός νευρώνας (συγκεκριμένα πυραμιδοειδής νευρώνας που απαντά συχνά στο neocortex) είναι ένα σύνθετο στοιχείο, με πολλούς χώρους και διανεμημένες λειτουργίες, που υλοποιούν τις λογικές πράξεις της άθροισης, του πολλαπλασιασμού, της χωρικής και χρονικής ολοκλήρωσης ταυτόχρονα και παράλληλα σε διαφορετικές ομάδες εισόδων. Η κεντρική δομή του νευρώνα, το σώμα, δέχεται ερεθίσματα από το προηγούμενο ιεραρχικά επίπεδο. Αν έρθουν αρκετά τέτοια ερεθίσματα σε σύντομο χρονικό διάστημα ο νευρώνας ενεργοποιείται. Οι πηγές των σημάτων που ο νευρώνας λαμβάνει στο σώμα του ονομάζονται *receptive field* του νευρώνα. Περιφερειακά του σώματος υπάρχουν τα δενδριτικά τμήματα, στα οποία υπάρχουν συνάψεις με νευρώνες του ίδιου επιπέδου, που παρέχουν την πληροφορία των συμφραζόμενων. Κάθε τμήμα αυτόνομα δέχεται σήματα και, αν αυτά είναι αρκετά, προκαλείται ένα NMDA spike. Σε αντίθεση με την ενεργοποίηση του σώματος, αυτό γενικά δεν αρκεί για να ενεργοποιήσει ολόκληρο το νευρώνα· θέτει όμως το νευρώνα σε κατάσταση “επιφυλακής”, διευκολύνοντας τη μετέπειτα ενεργοποίησή του από σωματικά ερεθίσματα. Ο νευρώνας επίσης λαμβάνει ερεθίσματα από την πλευρά του άξονα, από τους νευρώνες δηλαδή που “δέχονται” το σήμα του στην κανονική του πορεία. Αυτά αποτελούν σήματα ανάδρασης.

Μια απλοποιημένη εκδοχή αυτής της περιγραφής αποτελεί και ο υπολογιστικός νευρώνας του συστήματος που επιδεικνύεται. Έχει πάλι 3 περιοχές που συμπεριφέρονται διαφορετικά για να λαμβάνει εισόδους feedforward, context και feedback όπως ο βιολογικός. Επαρκής feedforward είσοδος αρκεί για να τον ενεργοποιήσει, ενώ επαρκής είσοδος context ή feedback τον θέτει στο λεγόμενο “predictive state”. Αν στη συνέχεια (στη συνέχεια... ακολουθία... συνεπαγωγή...) έρθει feedforward είσοδος, θα ενεργοποιηθεί πιο εύκολα και γρήγορα απ’ό,τι άλλοι νευρώνες με την ίδια είσοδο. Σε αντιπαράβολή, ο “νευρώνας” ενός κλασικού τεχνητού νευρωνικού δικτύου δεν έχει διακριτά λειτουργικά τμήματα και απλώς αθροίζει όλες τις εισόδους του.

### 2.2.3 Μοντέλο δικτύου

Και μόνο ο παραπάνω ανάδρομος ορισμός του νευρώνα παραπέμπει στο ρόλο του ως μονάδα δικτύου.

Η πρώτη οργανωτική δομή νευρώνων, τόσο στο neocortex, όσο και στο HTM, είναι το minicolumn. Συνοπτικά, ο ρόλος των minicolumns είναι να επιτρέπουν διαφορετικές εσωτερικές αναπαραστάσεις των εξωτερικών (feedforward) ερεθισμάτων, ανάλογα με τα συμφραζόμενα. Όλοι οι νευρώνες του minicolumn μοιράζονται *το ίδιο receptive field*, ενώ μεταξύ τους υπάρχει *inhibition*. Έτσι υλοποιείται μια αρχιτεκτονική “winner-takes-all”, όπου, αν κάποιοι νευρώνες είναι ήδη σε predictive κατάσταση όταν έρθει feedforward είσοδος που ερεθίζει το minicolumn, θα ενεργοποιηθούν πρώτοι και θα αποτρέψουν την ενεργοποίηση των υπολοίπων. Η παραπάνω περιγραφή βέβαια προτρέπει λίγο.

Πολλά minicolumns συνθέτουν ένα cortical layer, που αποτελεί τη λειτουργική μονάδα ενός συστήματος HTM. Κάθε minicolumn του cortical layer έχει διαφορετικές feedforward συνδέσεις προς την είσοδο του layer, ενώ διαφορετικά minicolumns μοιράζονται μεταξύ τους συνδέσεις που φέρουν πληροφορίες συμφραζόμενων. Σε πολλές περιπτώσεις μπορεί να υπάρχει ολικό inhibition μεταξύ των minicolumns. Αν βέβαια το cortical layer έχει κάποια τοπολογική διάταξη προς την είσοδό του<sup>1</sup>, όπως συχνά συμβαίνει στον εγκεφαλικό φλοιό, τότε οι inhibitory συνδέσεις είναι τοπικές και ένα minicolumn δεν ανταγωνίζεται κάποιο άλλο που βρίσκεται μακριά

<sup>1</sup>Δες το βίντεο “Topology (Episode 10)” της σειράς “HTM School” [24]

του. Τα λειτουργικά παραδείγματα παρακάτω αποτελούνται από 1 cortical layer. Ένα πιο σύνθετο σύστημα θα μπορούσε να ενώσει εν σειρά πολλά cortical layers.

Οι συνάψεις των HTM νευρώνων είναι επίσης πιστότερες στη βιολογία. Δε χρησιμοποιούν κάποιο βάρος που πολλαπλασιάζει την είσοδο, αλλά είναι δισταθείς, δηλαδή υπάρχουν ή όχι. Ένα μέγεθος όμως που χαρακτηρίζει κάθε σύναψη είναι η *μονιμότητά* της, που διαμορφώνεται δυναμικά από τη μαθησιακή διαδικασία (στην οποία δε θα αναφερθούμε παραπάνω εδώ) με αλγόριθμο τύπου Spike Timing Dependent Plasticity (STDP), που μπορεί ουσιαστικά να εμφανίσει και να εξαφανίσει υποθάλπουσες συνάψεις για να προσαρμόσει το σύστημα στα δεδομένα. Η μονιμότητα μιας σύναψης συγκρίνεται με ένα κατώφλι και, αν είναι πιο πάνω είναι ενεργή, αν είναι πιο κάτω είναι ανενεργή. Μια σύναψη με μονιμότητα λίγο πάνω από το κατώφλι μπορεί δηλαδή να απενεργοποιηθεί εύκολα.

### 2.2.4 Encoders

Ο εγκέφαλος έχει ως βασική διεπαφή με τον έξω κόσμο τον αμφιβληστροειδή χιτώνα του ματιού και τον κοχλία του αυτιού. Και τα 2 αυτά συστήματα μπορούμε να θεωρήσουμε ότι μετατρέπουν τα φυσικά ερεθίσματα του κόσμου σε μια πρώτη μορφή SDR, για να ερεθίσουν με τη σειρά τους το επόμενο επίπεδο νευρώνων. Το ρόλο αυτό στα συστήματα HTM καλύπτουν οι encoders.

Οι encoders μπορεί να είναι από πολύ απλά έως εξαιρετικά πολύπλοκα συστήματα (βλέπε αμφιβληστροειδή). Στη διαφάνεια 13 δίνεται το παράδειγμα ενός απλού encoder που κωδικοποιεί ακέραιους αριθμούς συγκεκριμένου εύρους. Μπορεί όμως να σχεδιαστεί και encoder που να κωδικοποιεί λέξεις ή και κείμενα, όπως κάνει η εταιρία cortical.io **semantic** στη διαφάνεια 14.

### 2.2.5 Decoders

## 2.3 Βασικές λειτουργίες HTM

Παρουσιάζονται ένα σύστημα που κανονικοποιεί τα SDR εισόδου, ο *Spatial Pooler*, και ένα σύστημα που μαθαίνει να αναγνωρίζει ακολουθίες απλών ερεθισμάτων, δηλαδή μια *μνήμη ακολουθιών*. Για το πρώτο προτείνονται τα σχετικά βίντεο από τη σειρά “HTM School” στο youtube, ενώ για το δεύτερο η αναφορά [2]. Λίγα λόγια για τη μνήμη ακολουθιών:

### 2.3.1 Sequence Memory

Στη *διαφάνεια 17* [19] βλέπουμε ένα cortical layer που αποτελείται από από minicolumns. Στη μέση (C) βλέπουμε πώς αντιδρά το σύστημα καθώς δέχεται για πρώτη φορά τις αλληλουχίες “ABCD” και “XBCY”. Παρατηρούμε ότι κάθε σύμβολο εισόδου ενεργοποιεί συγκεκριμένα minicolumns (αυτά που διεγείρονται περισσότερο), και ότι σε κάθε τέτοιο minicolumn ενεργοποιούνται *όλοι* οι νευρώνες.

Ας επικεντρωθούμε στο τι συμβαίνει όταν λαμβάνεται η είσοδος “Α”. Ενεργοποιούνται τα minicolumns που φαίνονται, αλλά κάποιοι από τους ενεργοποιημένους νευρώνες έχουν context συνάψεις (ενεργές ή μη) με κάποιους νευρώνες που ενεργοποιεί η είσοδος “Β”. Αυτό φαίνεται στη διαφάνεια 18 Α. Αφού το σύστημα δει την αλληλουχία Α -> Β μερικές φορές, συνάψεις αυτού του τύπου ενεργοποιούνται μέσω της διαδικασίας εκμάθησης, ακόμα κι αν στην αρχή ήταν ανενεργές. Έτσι, όταν ξαναέρθει η είσοδος “Α”, συγκεκριμένοι νευρώνες των στηλών που



αντιστοιχούν στο “B” τίθενται σε predictive state. Όταν λοιπόν όντως έρχεται η είσοδος “B” αμέσως μετά αυτοί οι νευρώνες ενεργοποιούνται πρώτοι και μπλοκάρουν μέσω inhibitory συνδέσεων τους υπόλοιπους του ιδίου minicolumn. Έτσι καταλήγουμε στην ακολουθία της κάτω σειράς (D) της διαφάνειας 17. Τα παραπάνω σημαίνουν ότι το σύστημα *προβλέπει την επόμενη είσοδό του*, που είναι ο βασικός στόχος του προβλήματος εκμάθησης ακολουθιών.

Ο μηχανισμός αυτός οδηγεί σε δύο παρατηρήσεις: η εσωτερική αναπαράσταση κάθε συμβόλου είναι *διαφορετική ανάλογα με τα συμφραζόμενα* και το σύστημα μπορεί να πραγματοποιεί *πολλαπλές προβλέψεις ταυτόχρονα*.

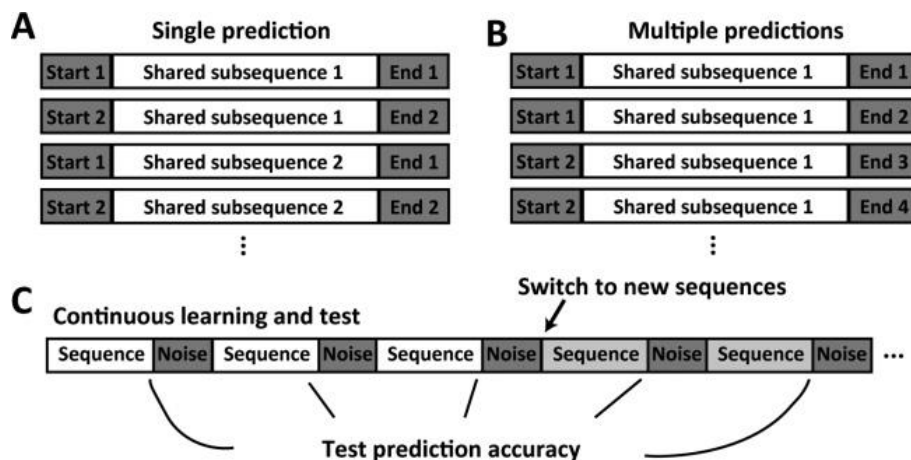
Το πρώτο φαίνεται στην κάτω σειρά (D) από τους διαφορετικούς νευρώνες που ενεργοποιούνται για να συμβολίσουν την ίδια είσοδο “B”, ανάλογα αν το προηγούμενο σύμβολο ήταν “A” ή “X”, οπότε προκύπτει η εσωτερική αναπαράσταση B’ ή B” αντίστοιχα. Το δεύτερο σημείο φαίνεται στη διαφάνεια 18 B, όπου δίνεται η είσοδος “B”, χωρίς να έχει προηγηθεί “A” ή “X”. Το σύστημα τότε προβλέπει την ένωση των 2 καταστάσεων C’ και C”, οι οποίες ενεργοποιούνται παράλληλα όταν όντως έρχεται η είσοδος C. Εκείνη τη στιγμή λοιπόν το σύστημα σωστά προβλέπει ότι η συνέχεια της ακολουθίας θα είναι ή “D” ή “Y”.

## 2.4 Αποτελέσματα - Προσομοιώσεις

Το HTM δοκιμάστηκε τόσο σε τεχνητά, όσο και σε πραγματικά δεδομένα για να αξιολογηθεί η δυνατότητα του να παρέχει ακριβείς προβλέψεις. Παρακάτω θα αναλυθούν δύο παραδείγματα και θα δοθούν τα συγκριτικά διαγράμματα των επιδόσεων με άλλα νευρωνικά δίκτυα.

### Πρόβλεψη συμβόλου ακολουθίας

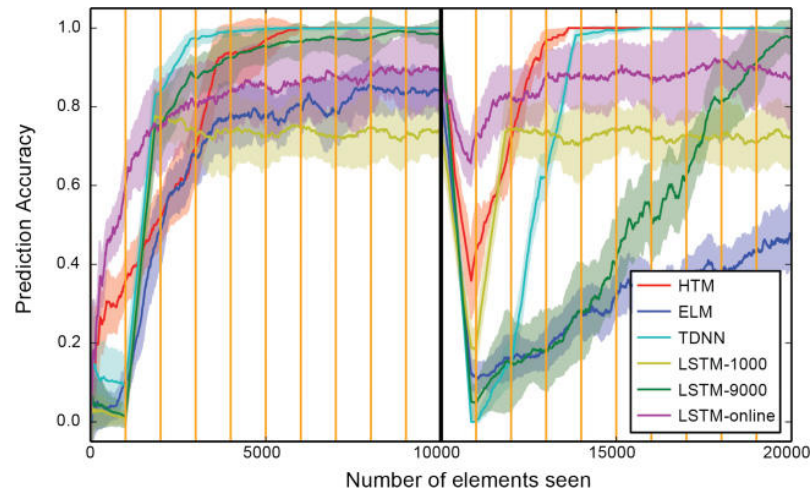
Οι αλγόριθμοι του Hierarchical Temporal Memory (HTM) δοκιμάστηκαν σε διαφορετικά προβλήματα [17], [19], ώστε να αξιολογηθεί η επίδοση του συστήματος σε σχέση με άλλες υλοποιήσεις. Στο πρώτο παράδειγμα έχουν δημιουργηθεί δύο ακολουθίες από σύμβολα, σύμφωνα με την εικόνα (2.1). Στόχος είναι η πρόβλεψη του τελευταίου συμβόλου.



Σχήμα 2.1: Ακολουθίες συμβόλων για την προσομοίωση

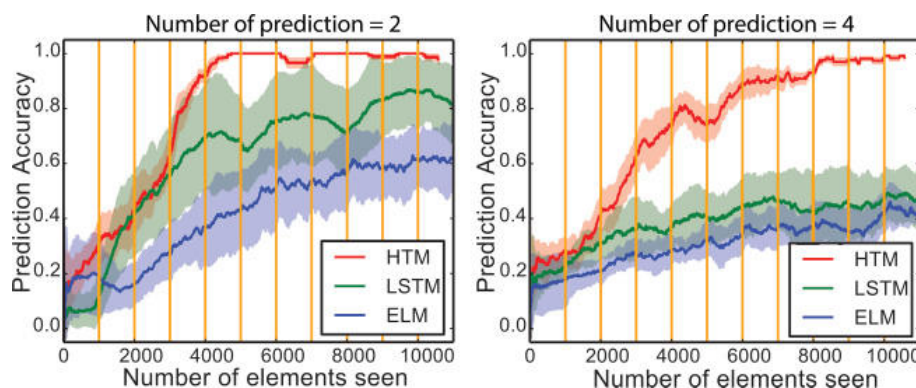
Οι ακολουθίες φτιάχτηκαν έτσι ώστε να χρειάζεται ένα βάθος μνήμης τουλάχιστον 2 παρελθοντικών συμβόλων για την επιτυχή πρόβλεψη. Το πρώτο set ακολουθιών

έχει δημιουργηθεί για να ελεγχθεί η δυνατότητα απλής πρόβλεψης, καθώς κάθε ακολουθία προσδιορίζεται πλήρως από τα πεδία “Start” και “Shared Subsequence”. Αντιθέτως, στο δεύτερο set δύο ακολουθίες μπορεί να έχουν ίδια τα παραπάνω πεδία. Σε αυτή την περίπτωση το νευρωνικό δίκτυο θα πρέπει να επιτύχει την ορθή πρόβλεψη όλων των πιθανών καταλήξεων, δηλαδή την ταυτόχρονη πολλαπλή πρόβλεψη. Τέλος, ανάμεσα στις ακολουθίες προστέθηκε ένα σύμβολο θορύβου, ώστε να εξεταστεί η ανθεκτικότητα του συστήματος σε αυτόν.



Σχήμα 2.2: Απόδοση διαφορετικών υλοποιήσεων για απλή πρόβλεψη

Στο σχήμα (2.2) δίνεται η συγκριτική απόδοση των διαφορετικών υλοποιήσεων νευρωνικών δικτύων για το πρώτο set ακολουθιών. Τα δίκτυα που υλοποιούνται είναι το HTM, το ELM, το TDNN και το LSTM. Ο προσδιορισμός στο LSTM αφορά το πλήθος των δειγμάτων που χρησιμοποιούνται για retraining σε κάθε κατακόρυφη κίτρινη γραμμή του διαγράμματος. Στο στοιχείο 10000 που σημειώνεται με κάθετη κατακόρυφη γραμμή, αλλάζουν αμοιβαία τα δύο τελευταία σύμβολα κάθε ακολουθίας και μελετάται η προσαρμογή του δικτύου στο νέο input space. Βλέπουμε ότι το HTM πετυχαίνει αρχικά σε έναν ικανοποιητικό αριθμό δειγμάτων ακρίβεια 100%. Το πιο σημαντικό είναι, βέβαια, ότι μέσω online training αναπροσαρμόζεται ταχύτερα από κάθε άλλη υλοποίηση στο νέο χώρο ακολουθιών εισόδου.

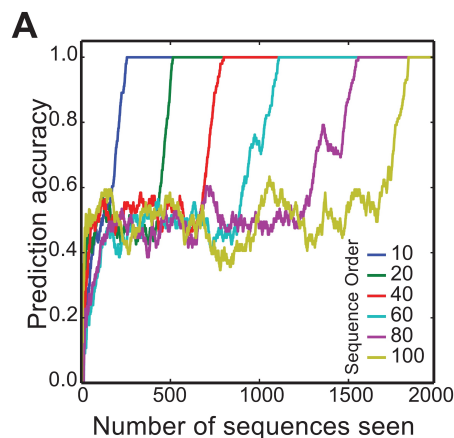


Σχήμα 2.3: Απόδοση δικτύων για πολλαπλές προβλέψεις

Οι πολλαπλές προβλέψεις εξετάζονται από το δεύτερο set και τα συγκριτικά αποτελέσματα για διπλή και τετραπλή πρόβλεψη δίνεται στο σχήμα (2.3). Όπως φαίνεται, το HTM είναι το μοναδικό σύστημα που πετυχαίνει μετά από κάποιο

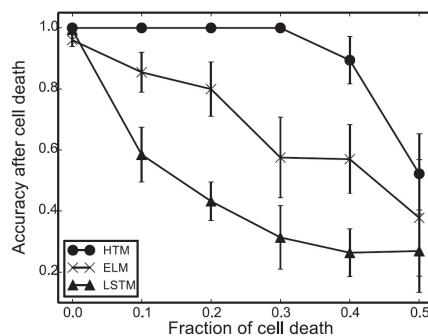
διάστημα ακρίβεια σχεδόν 100%. Αυτό απορρέει από την αναπαράσταση μέσω SDR και τη δυνατότητα του Temporal Pooler για πολλαπλές προβλέψεις.

Στη συνέχεια, μελετήθηκε η απόδοση του HTM για ακολουθίες διαφορετικού μήκους. Ένα από τα σημαντικότερα στοιχεία του Temporal Pooler είναι η δυνατότητα αναπροσαρμογής online του «βάθους» μνήμης που χρησιμοποιείται για την πρόβλεψη. Έτσι, όπως παρατηρούμε και στο διάγραμμα (2.4), πετυχαίνει πρόβλεψη ακόμα και τάξης 100, που σημαίνει ότι απαιτείται μνήμη 99 συμβόλων για να προβλεφθεί επιτυχώς το τελευταίο. Το πλήθος των ακολουθιών που απαιτείται για να εκπαιδευτεί το σύστημα και να πετύχει ακρίβεια κοντά στο 100% αυξάνει γραμμικά με την τάξη της ακολουθίας.



Σχήμα 2.4: Απόδοση HTM συναρτήσει της τάξης της ακολουθίας

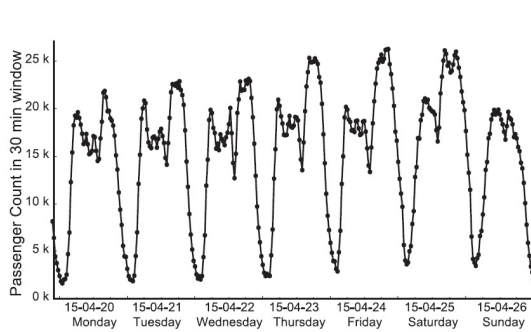
Τέλος, τα διαφορετικά συστήματα εξετάστηκαν και ως προς την ανθεκτικότητα σε καταστροφή του δικτύου μέσω νεκρών νευρώνων. Το HTM έμεινε ανεπηρέαστο ακόμα για 30% νεκρούς νευρώνες, συνθήκες κατά τις οποίες το ELM και το LSTM είχαν υποβαθμιστεί αρκετά.



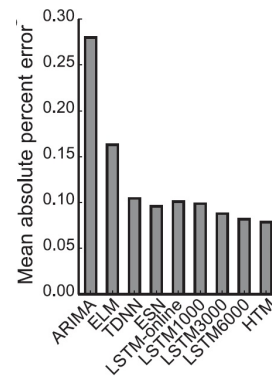
Σχήμα 2.5: Ανθεκτικότητα υλοποιήσεων σε νεκρούς νευρώνες

### Πρόβλεψη της ζήτησης taxi στη Νέα Υόρκη

Το Hierarchical Temporal Memory δοκιμάστηκε και σε πραγματικά streaming δεδομένα. Πιο συγκεκριμένα, ζητήθηκε η πρόβλεψη της ζήτησης στα taxi της Νέας Υόρκης 2.5 ώρες πριν την πραγματική μέτρηση. Η ζήτηση ποσοτικοποιήθηκε ως η συνολική εξυπηρέτηση πελατών σε ένα χρονικό παράθυρο 30 λεπτών. Εξετάστηκαν διάφοροι τύποι νευρωνικών δικτύων, καθώς και το στατιστικό μοντέλο ARIMA.



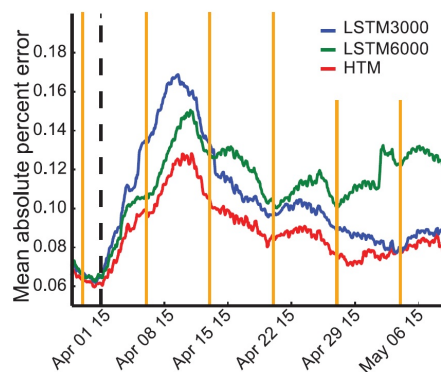
(α') Διάγραμμα ζήτησης



(β') Μέσο απόλυτο σφάλμα

Σχήμα 2.6: Πρόβλεψη ζήτησης των taxi

Στη συνέχεια, τα δεδομένα εισόδου τροποποιήθηκαν τεχνητά ανεβάζοντας ή μειώνοντας τη ζήτηση κατά 20% σε διαφορετικές χρονικές περιόδους εντός κάθε μέρας. Στόχος ήταν να μελετηθεί η προσαρμογή του δικτύου στις νέες (τεχνητές) συνήθειες των επιβατών taxi της Νέας Υόρκης. Η αλλαγή πραγματοποιήθηκε την 1η Απριλίου και εντός 2 εβδομάδων το σύστημα HTM αναπροσάρμοσε το μοντέλο του, ώστε να πετυχαίνει εκ νέου σωστές προβλέψεις. Σε αντίθεση, το LSTM είτε απέτυχε να επανέρθει, είτε χρειάστηκε πολύ περισσότερο διάστημα για retraining των παραμέτρων του.



Σχήμα 2.7: Προσαρμογή δικτύων σε μεταβολή της ζήτησης

# Βιβλιογραφία

- [1] M. Mohammadi, A. Al-Fuqaha, S. Sorour και M. Guizani, «Deep Learning for IoT Big Data and Streaming Analytics: A Survey,» *IEEE Communications Surveys Tutorials*, τόμ. 20, αρθμ. 4, σσ. 2923–2960, Fourthquarter 2018, ISSN: 1553-877X. doi: 10.1109/COMST.2018.2844341.
- [2] J. Piaget, *The Origins of Intelligence in Children*, σύνταξη υπό M. Cook, σειρά The Origins of Intelligence in Children. New York, NY, US: W W Norton & Co, 1952, 419 **pagetotals**. doi: 10.1037/11494-000.
- [3] P. W. Anderson, «More Is Different,» *Science*, τόμ. 177, αρθμ. 4047, σσ. 393–396, 4 Αύγ. 1972, ISSN: 0036-8075, 1095-9203. doi: 10.1126/science.177.4047.393. pmid: 17796623.
- [4] M. Minsky, *Society Of Mind*. Simon and Schuster, 15 Μαρ. 1988, 342 **pagetotals**, ISBN: 978-0-671-65713-0.
- [5] D. B. Lenat και E. A. Feigenbaum, «On the Thresholds of Knowledge,» *Artificial Intelligence*, τόμ. 47, αρθμ. 1, σσ. 185–250, 1 Ιαν. 1991, ISSN: 0004-3702. doi: 10.1016/0004-3702(91)90055-0.
- [6] P. Kanerva, «Associative Neural Memories,» στο, M. H. Hassoun, επιμελητής, New York, NY, USA: Oxford University Press, Inc., 1993, κεφ. Sparse Distributed Memory and Related Models, σσ. 50–76, ISBN: 0-19-507682-6.
- [7] P. Wang, «On the Working Definition of Intelligence,» 1995.
- [8] V. B. Mountcastle, «The columnar organization of the neocortex,» *Brain: A Journal of Neurology*, τόμ. 120 ( Pt 4), σσ. 701–722, Απρ. 1997, ISSN: 0006-8950. doi: 10.1093/brain/120.4.701. pmid: 9153131.
- [9] K. Doya, «What are the computations of the cerebellum, the basal ganglia and the cerebral cortex?» *Neural Networks: The Official Journal of the International Neural Network Society*, τόμ. 12, αρθμ. 7-8, σσ. 961–974, Οκτ. 1999, ISSN: 1879-2782. pmid: 12662639.
- [10] J. C. Horton και D. L. Adams, «The Cortical Column: A Structure without a Function,» *Philosophical Transactions of the Royal Society B: Biological Sciences*, τόμ. 360, αρθμ. 1456, σσ. 837–862, 29 Απρ. 2005, ISSN: 0962-8436. doi: 10.1098/rstb.2005.1623. pmid: 15937015.
- [11] S. Legg και M. Hutter, «A Collection of Definitions of Intelligence,» στο *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*, IOS Press, 2007, σσ. 17–24.
- [12] P. Wang, «What Do You Mean by “AI”?», **presented at** Proceedings of the 2008 Conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference, IOS Press, 20 Ιούν. 2008, σσ. 362–373, ISBN: 978-1-58603-833-5. διεύθυν.: <http://dl.acm.org/citation.cfm?id=1566174.1566207> (επίσκεψη 28/05/2019).



- [13] J. Defelipe, H. Markram και K. S. Rockland, «The Neocortical Column,» *Frontiers in Neuroanatomy*, τόμ. 6, 2012, ISSN: 1662-5129. doi: 10 . 3389 / fnana .2012 .00022.
- [14] R. Paton, H. Bolouri, W. M. L. Holcombe, J. H. Parish και R. Tateson, *Computation in Cells and Tissues: Perspectives and Tools of Thought*. Springer Science & Business Media, 14 Μαρ. 2013, 349 **pagetotals**, ISBN: 978-3-662-06369-9.
- [15] M. Benedek, E. Jauk, M. Sommer, M. Arendasy και A. C. Neubauer, «Intelligence, Creativity, and Cognitive Control: The Common and Differential Involvement of Executive Functions in Intelligence and Creativity,» *Intelligence*, τόμ. 46, σσ. 73–83, 1 Σεπτ. 2014, ISSN: 0160-2896. doi: 10 .1016/j .intell .2014 .05 .007.
- [16] Y. Bengio, D.-H. Lee, J. Bornschein, T. Mesnard και Z. Lin, «Towards Biologically Plausible Deep Learning,» 13 Φεβ. 2015. arXiv: 1502 .04156 [cs]. διεύθν.: <http://arxiv.org/abs/1502.04156> (επίσκεψη 28/05/2019).
- [17] A. Lavin και S. Ahmad, «Evaluating Real-Time Anomaly Detection Algorithms – The Numenta Anomaly Benchmark,» στο *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*, Δεκ. 2015, σσ. 38–44. doi: 10 .1109/ICMLA .2015 .141.
- [18] A. Adamatzky, επιμελητής, *Advances in Physarum Machines: Sensing and Computing with Slime Mould*, Emergence, Complexity and Computation, Springer International Publishing, 2016, ISBN: 978-3-319-26661-9. διεύθν.: <https://www.springer.com/gp/book/9783319266619> (επίσκεψη 29/05/2019).
- [19] Y. Cui, S. Ahmad και J. Hawkins, «Continuous online sequence learning with an unsupervised neural network model,» *Neural Computation*, 2016.
- [20] P. Haueis, «The Life of the Cortical Column: Opening the Domain of Functional Architecture of the Cortex (1955–1981),» *History and Philosophy of the Life Sciences*, τόμ. 38, αρθμ. 3, 2016, ISSN: 0391-9714. doi: 10 .1007/s40656 -016 -0103 -4. pmid: 27325058.
- [21] J. Hawkins και S. Ahmad, «Why Neurons Have Thousands of Synapses, a Theory of Sequence Memory in Neocortex,» *Frontiers in Neural Circuits*, τόμ. 10, σ. 23, 2016, ISSN: 1662-5110. doi: 10 . 3389 / fncir .2016 .00023. διεύθν.: <http://journal.frontiersin.org/article/10.3389/fncir.2016.00023>.
- [22] Y. LeCun, «Deep learning,» CERN talk, 2016, διεύθν.: <https://indico.cern.ch/event/510372/attachments/1245509/1840815/lecun-20160324-cern.pdf>.
- [23] J. Regier, K. Pamnany, R. Giordano, R. Thomas, D. Schlegel, J. McAuliffe και Prabhat, «Learning an Astronomical Catalog of the Visible Universe through Scalable Bayesian Inference,» 10 Νοέ. 2016. arXiv: 1611 .03404 [astro-ph, stat]. διεύθν.: <http://arxiv.org/abs/1611.03404> (επίσκεψη 29/05/2019).
- [24] M. Taylor. (2016). HTM School. Video lectures on YouTube, διεύθν.: <https://www.youtube.com/playlist?list=PL3yXMgtrZmDqhsFQzwUC9V8MeeV0Q7eZ9>.

- [25] C. M. Vineyard και S. J. Verzi, «Overcoming the Static Learning Bottleneck - the need for adaptive neural learning,» στο *2016 IEEE International Conference on Rebooting Computing (ICRC)*, Οκτ. 2016, σσ. 1–3. DOI: 10.1109/ICRC.2016.7738692.
- [26] J. Bezanson, A. Edelman, S. Karpinski και V. Shah, «Julia: A Fresh Approach to Numerical Computing,» *SIAM Review*, τόμ. 59, αρθμ. 1, σσ. 65–98, 1 Ιαν. 2017, ISSN: 0036-1445. DOI: 10.1137/141000671.
- [27] K. Brading, E. Castellani και N. Teh, «Symmetry and Symmetry Breaking,» στο *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, επιμελητής, Winter 2017, Metaphysics Research Lab, Stanford University, 2017. διεύθν.: <https://plato.stanford.edu/archives/win2017/entries/symmetry-breaking/> (επίσκεψη 29/05/2019).
- [28] S. Sabour, N. Frosst και G. E. Hinton, «Dynamic Routing Between Capsules,» στο *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan και R. Garnett, επιμελητές, Curran Associates, Inc., 2017, σσ. 3856–3866. διεύθν.: <http://papers.nips.cc/paper/6975-dynamic-routing-between-capsules.pdf> (επίσκεψη 30/05/2019).
- [29] V. Losing, B. Hammer και H. Wersing, «Incremental On-Line Learning: A Review and Comparison of State of the Art Algorithms,» *Neurocomputing*, τόμ. 275, σσ. 1261–1274, 31 Ιαν. 2018, ISSN: 0925-2312. DOI: 10.1016/j.neucom.2017.06.084.
- [30] O. Nachum, S. (Gu, H. Lee και S. Levine, «Data-Efficient Hierarchical Reinforcement Learning,» στο *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi και R. Garnett, επιμελητές, Curran Associates, Inc., 2018, σσ. 3303–3313. διεύθν.: <http://papers.nips.cc/paper/7591-data-efficient-hierarchical-reinforcement-learning.pdf> (επίσκεψη 30/05/2019).
- [31] S. Sabour, N. Frosst και G. Hinton, «Matrix Capsules with EM Routing,» **presented at** 6th International Conference on Learning Representations, ICLR, 2018.
- [32] P. Xiong, Y. Zhu, Z. Sun, Z. Cao, M. Wang, Y. Zheng, J. Hou, T. Huang και Z. Que, «Application of Transfer Learning in Continuous Time Series for Anomaly Detection in Commercial Aircraft Flight Data,» στο *2018 IEEE International Conference on Smart Cloud (SmartCloud)*, Σεπτ. 2018, σσ. 13–18. DOI: 10.1109/SmartCloud.2018.00011.
- [33] O. Vinyals, I. Babuschkin, J. Chung, M. Mathieu και D. Silver. (2019). AlphaStar: Mastering the Real-Time Strategy Game StarCraft II, διεύθν.: <https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/>.
- [34] B. Chazelle. (). Natural Algorithms and Influence Systems, διεύθν.: <https://cacm.acm.org/magazines/2012/12/157889-natural-algorithms-and-influence-systems/abstract> (επίσκεψη 28/05/2019).
- [35] P. Wang, «Cognitive Logic versus Mathematical Logic,»