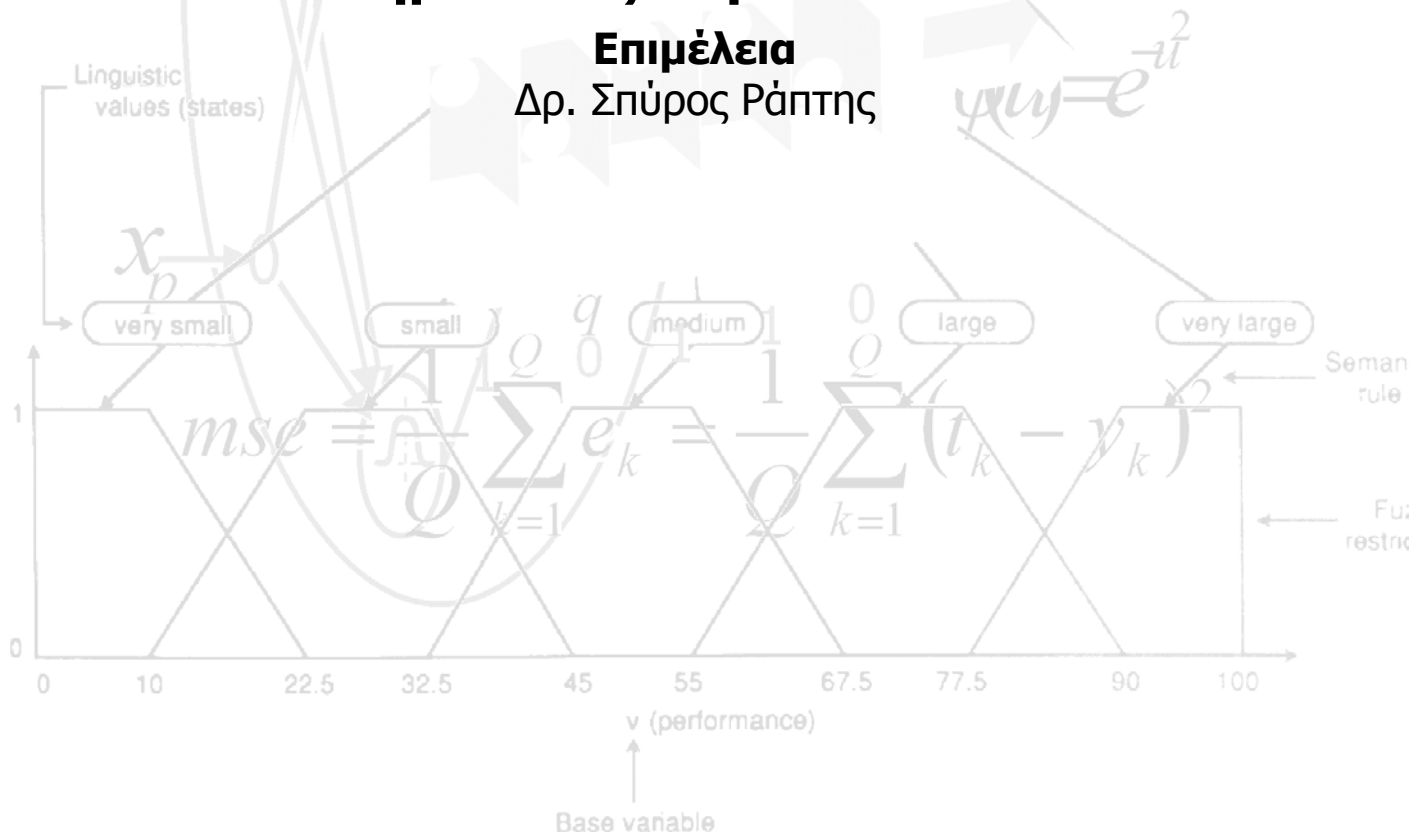


Υπολογιστική Νοημοσύνη

- Ασαφής Λογική
- Νευρωνικά Δίκτυα
- Γενετικοί Αλγόριθμοι

Επιμέλεια

Δρ. Σπύρος Ράπτης



Σκοπός

Σκοπός των σημειώσεων αυτών σίγουρα δεν είναι η εξαντλητική μελέτη των θεμάτων που θίγονται αφού κάτι τέτοιο είναι τελείως έξω τόσο από τους περιορισμούς σε χώρο όσο και από τους σκοπούς του μαθήματος.

Η έμφαση των σημειώσεων δόθηκε στην εξοικείωση με τις έννοιες της υπολογιστικής νοημοσύνης, την κατανόηση των προβλημάτων στα οποία απευθύνεται και την γενική μεθοδολογία που υιοθετεί για την επίλυσή τους. Παρουσιάζονται τρεις προσεγγίσεις: η ασαφής λογική, τα νευρωνικά δίκτυα και οι γενετικοί αλγόριθμοι. Η «κεντρική ιδέα» καθεμιάς αλλά και οι συγκεκριμένες ανάγκες που οδήγησαν στην διαμόρφωσή τους αποτελούν κύριους άξονες της παρουσίασης.

Θα πρέπει να καταστεί σαφές ότι η υπολογιστική νοημοσύνη δεν σκοπεύει να αντικαταστήσει και να παραγκωνίσει τις «κλασσικές» μεθοδολογίες επίλυσης προβλημάτων. Απλά παρέχει μια εναλλακτική μεθοδολογία που η πρακτική έχει δείξει ότι είναι πιο ευέλικτη και καταλληλότερη για συγκεκριμένες κατηγορίες προβλημάτων. Ο ρόλος της, δηλαδή, είναι πιο συχνά συμπληρωματικός παρά ανταγωνιστικός. Καταξιωμένες θεωρίες και μέθοδοι που χαίρουν ευρείας αποδοχής διατηρούν αμείωτη την αξία τους και (όχι σπάνια) είναι προτιμότερες από εκείνες που προσφέρει η υπολογιστική νοημοσύνη.

Οι σημειώσεις έχουν γραφεί με τέτοιο τρόπο ώστε να ελαχιστοποιηθούν κατά το δυνατό οι όποιες προαπαιτήσεις. Έτσι, οι μαθηματικοί φορμαλισμοί έχουν περιοριστεί στο ελάχιστο. Παρόλα αυτά κάποια στοιχειώδης τουλάχιστον επαφή με θέματα λογικής, συνολοθεωρίας, κλπ. θα διευκολύνουν τον αναγνώστη να εξοικειωθεί ευκολότερα και σε μεγαλύτερο βαθμό με τις παρουσιαζόμενες έννοιες.

*Για οποιαδήποτε επικοινωνία σχετικά με αυτές τις σημειώσεις, χρησιμοποιήστε την ηλεκτρονική διεύθυνση: **spy@i1sp.gr***

Περιεχόμενα

Σκοπός	2
Εισαγωγή	5
Θεωρία Συστημάτων	5
«Παραδοσιακή» Προσέγγιση	5
Τεχνητή Νοημοσύνη	6
Περί Αβεβαιότητας	6
Τα Κίνητρα της Μετάβασης	7
Υπολογιστική Νοημοσύνη.....	8
Ασαφής Λογική	11
Εισαγωγή	11
Από τη Θέση στην Αντίθεση — Από το A στο $\acute{o}\chi\iota-A$	11
Ο Ασαφής Κόσμος.....	11
Παράδοξα Κλασσικής Λογικής	12
Ασαφή Σύνολα έναντι Κλασσικών Συνόλων	12
Ορισμός Συνόλου.....	12
Πράξεις σε Ασαφή Σύνολα	14
Ασαφής Λογική	15
Γλωσσικές Μεταβλητές.....	15
Ασαφείς Προτάσεις	16
Εξαγωγή Συμπεράσματος από Ασαφείς Προτάσεις.....	16
Ασαφή Συστήματα	17
Παράδειγμα Εξαγωγής Ασαφούς Συμπεράσματος	18
Νευρωνικά Δίκτυα	21
Εισαγωγή	21
Η Δομή του Εγκεφάλου	21
Βιολογικά Νευρωνικά Δίκτυα Έναντι Υπολογιστικών Συστημάτων	21
Τεχνητά Νευρωνικά Δίκτυα	22
Ιδιότητες Τεχνητών Νευρωνικών Δικτύων	22
Το Μοντέλο του Νευρώνα.....	24
Αρχιτεκτονικές Νευρωνικών Δικτύων	25
Δίκτυα Προστροφολόγησης Ενός Στρώματος	25
Δίκτυα Προστροφολόγησης Πολλών Στρωμάτων.....	25
Δίκτυα Ανατροφολόγησης	26
Συνήθεις Αρχιτεκτονικές Νευρωνικών Δικτύων	27
Δίκτυα Perceptron.....	27
Γραμμικά Νευρωνικά Δίκτυα	29
Αλγόριθμοι Εκπαίδευσης Νευρωνικών Δικτύων.....	29
Αλγόριθμος Εκπαίδευσης για Δίκτυα Perceptron Ενός Στρώματος.....	29
Αλγόριθμος Εκπαίδευσης για Γραμμικά Νευρωνικά Δίκτυα.....	31
Κανόνας Μάθησης Widrow-Hoff	32
Αλγόριθμος Ανάστροφης Διάδοσης Σφάλματος	33
Βελτιώσεις στον Αλγόριθμο Ανάστροφης Διάδοσης Σφάλματος	34
Γενετικοί Αλγόριθμοι	37
Εισαγωγή	37
Βελτιστοποίηση	37
Πλεονεκτήματα Γενετικών Αλγορίθμων.....	37
Παραδοσιακές Μέθοδοι Αναζήτησης	38
Γενετικοί Αλγόριθμοι Έναντι Παραδοσιακών Μεθόδων	39
Η Λειτουργία των Γενετικών Αλγορίθμων	39
Κωδικοποίηση Παραμέτρων	40
Επιλογή Μεγέθους Πληθυσμού και Αρχικοποίηση	40
Γενετικοί Τελεστές	40
Συνάρτηση Αποτίμησης	41
Βρόχος Εκτέλεσης Γενετικού Αλγορίθμου	42
Παράδειγμα Γενετικού Αλγορίθμου	43
Πίνακας Ορολογίας Γενετικών Αλγορίθμων	46

Υβριδικά Μοντέλα	47
Εισαγωγή	47
Μοντέλα	47
Νευροασαφή Μοντέλα	48
Ασαφή-Γενετικά Μοντέλα	48
Νευρωνικά-Γενετικά Μοντέλα	49
Άλλα Μοντέλα	49
Βραχυγραφίες	50
Προτεινόμενη Βιβλιογραφία	51
Κατάλογος Σχημάτων και Πινάκων	52

Εισαγωγή

Θεωρία Συστημάτων

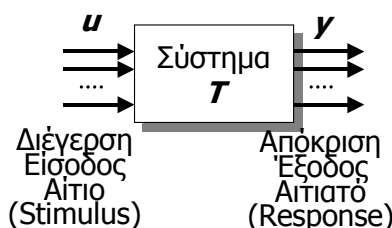
Η Θεωρία
Συστημάτων

Η **Θεωρία Συστημάτων** προσφέρει μια ενοποιημένη αντιμετώπιση φυσικών ή τεχνητών φαινομένων, διεργασιών, δομών κ.λπ. Είναι μια καταξιωμένη και ευρέως χρησιμοποιούμενη θεωρία με εφαρμογή σε ευρύ φάσμα επιστημών και ειδικοτήτων.

Ως σύστημα μπορεί να θεωρηθεί μια σφαίρα που αναπηδά στο έδαφος, ένας άνθρωπος που ταξιδεύει σε ένα τρένο, ένα κτίριο που υφίσταται τις επιδράσεις ενός σεισμού, ένας ρομποτικός βραχίονας που επιτελεί κάποιο έργο, αλλά ακόμη και ένα σύνολο ανθρώπων, ένα χρηματοοικονομικό σύστημα, το Χρηματιστήριο, ή και μια ολόκληρη κοινωνία.

Το κοινό στοιχείο σε όλα αυτά τα παραδείγματα είναι η ιδέα μιας οντότητας που επηρεάζεται από το περιβάλλον της αλλά και επιδρά σε αυτό. Δηλαδή, η θεωρία συστημάτων βασίζεται στο τρίπτυχο:

- ❑ Είσοδος (διέγερση, αίτιο — stimulus)
- ❑ Σύστημα
- ❑ Έξοδος (απόκριση, αιτιατό — response)



Σχήμα 1. Αφαιρετική θεώρηση ενός συστήματος

Στο Σχήμα 1, παρουσιάζεται μια αφαιρετική αναπαράσταση ενός συστήματος. Το σχήμα αυτό θα χρησιμοποιηθεί ως βάση σε όλη τη περαιτέρω παρουσίαση.

Οικογένειες
προβλημάτων
συστημάτων

Με βάση το τρίπτυχο αυτό μπορούμε να θεωρήσουμε τις παρακάτω οικογένειες προβλημάτων:

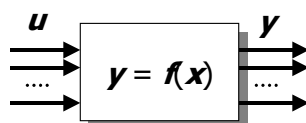
- ❑ **Προβλήματα Ανάλυσης:**
Υπολογισμός εξόδου, δεδομένων της εισόδου και του συστήματος
- ❑ **Προβλήματα Σύνθεσης:**
Υπολογισμός συστήματος, δεδομένων της εισόδου και της εξόδου
- ❑ **Προβλήματα Μέτρησης:**
Υπολογισμός εισόδου, δεδομένων του συστήματος και της εξόδου

«Παραδοσιακή» Προσέγγιση

Ακριβή μαθηματικά
μοντέλα

Για πολλές δεκαετίες, η επιστήμη ρητά υπαγόρευε την χρήση **μοντέλων**. Για την μελέτη ενός συστήματος, απαιτούνταν η κατάστρωση ενός μαθηματικού μοντέλου που να περιγράφει επακριβώς τον τρόπο εξάρτησης των εξόδων από τις εισόδους.

Σε αναλογία με το Σχήμα 1, μπορούμε να παραστήσουμε την «βασισμένη-σε-μοντέλα» θεώρηση με το Σχήμα 2.



Σχήμα 2. Σύστημα βασισμένο σε μαθηματικά μοντέλα

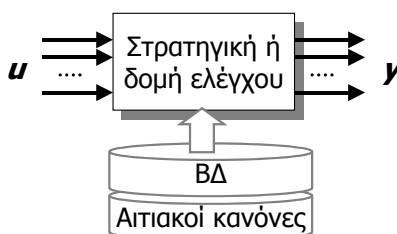
Όμως, ακόμη και το φαινομενικά απλούστατο σύστημα μιας σφαίρας που αναπηδά ελαστικά στο έδαφος μπορεί να παρουσιάσει αυξημένα επίπεδα δυσκολίας εάν συνυπολογιστούν εξαντλητικά όλοι παράγοντες που το επηρεάζουν όπως, για παράδειγμα, η κατεύθυνση και ένταση του αέρα τη στιγμή του πειράματος, οι ακριβείς γεωγραφικές συντεταγμένες του τόπου διεξαγωγής του πειράματος (αφού επηρεάζουν την τιμή της επιτάχυνσης της βαρύτητας), κλπ.

Έτσι, όταν η ακριβής γνώση των συσχετίσεων δεν είναι διαθέσιμη ή εγείρει αυξημένη πολυπλοκότητα, καταφεύγουμε σε ένα σύνολο από **υποθέσεις** για τα άγνωστα χαρακτηριστικά του προβλήματος και συχνά ένα σύνολο από **απλοποιητικές παραδοχές** για τον περιορισμό της πολυπλοκότητας.

Τεχνητή Νοημοσύνη

Η **Τεχνητή Νοημοσύνη** (TN — Artificial Intelligence, AI) ξεκίνησε με στόχο την κατανόηση των γνωστικών ικανοτήτων του ανθρώπου. Η γνώση αυτή θα οδηγούσε στη σχεδίαση μηχανών που προσομοιώνουν την ανθρώπινη διαδικασία συμπερασμού και λήψης αποφάσεων. Η TN προσέλκυσε έντονο ενδιαφέρον και έτυχε ισχυρής χρηματοδότησης.

Σε αναλογία με το Σχήμα 1, μπορούμε να παραστήσουμε ένα σύστημα TN με τον τρόπο που φαίνεται στο Σχήμα 3.



Σχήμα 3. Ένα τυπικό σύστημα Τεχνητής Νοημοσύνης

Εργαλεία TN

Βασικά εργαλεία της TN αποτελούν οι **λειτουργίες χειρισμού συμβόλων** καθώς και η **λογική πρώτης τάξης**.

Αποτελέσματα TN

Η εκτεταμένη έρευνα στην περιοχή της TN οδήγησε σε σημαντικά αποτελέσματα όπως:

- ❑ Τα έμπειρα συστήματα
- ❑ Τη θεωρία παιγνίων
- ❑ Την αυτόματη απόδειξη θεωρημάτων
- ❑ Τα συστήματα επεξεργασίας φυσικής γλώσσας, κλπ.

Αδυναμίες TN

Όμως, η TN δεν στάθηκε στο ύψος των αρχικών προσδοκιών της. Δεν συνεισέφερε σημαντικά στη λύση ρεαλιστικών προβλημάτων σε τομείς όπως η αναγνώριση και δεν βελτίωσε σημαντικά την κατανόηση της ανθρώπινη διαδικασίας συλλογισμού, διαμόρφωσης εννοιών, αναγνώρισης προτύπων, κ.λπ. Επιπρόσθετα, η TN είναι αφιλόξενη σε διαδικασίες που εμπλέκουν αριθμητικά δεδομένα. Τέλος, οι δυνατότητες χειρισμού της αβεβαιότητας, της ανακρίβειας και του θορύβου που παρέχει είναι μάλλον περιορισμένες.

Περί Αβεβαιότητας

Η **αβεβαιότητα** είναι μια βασική έννοια στη σχεδίαση συστημάτων. Εμφανίζεται σε διάφορες μορφές που μπορεί να συνυπάρχουν σχεδόν σε κάθε ρεαλιστικό πρόβλημα. Χονδρικά, μπορούμε να διακρίνουμε τις παρακάτω:

❑ Στοχαστική αβεβαιότητα

Οφείλεται στην τυχαιότητα. Αυτή η μορφή της αβεβαιότητας είναι γνωστή και αντιμετωπίζεται από τα στοχαστικά μαθηματικά.

Κλασσικό παράδειγμα στοχαστικής αβεβαιότητας είναι η ρίψη του ζαριού: δεν μπορούμε να γνωρίζουμε με βεβαιότητα το αποτέλεσμα εκ των προτέρων.

❑ Πληροφοριακή αβεβαιότητα

Οφείλεται στους περιορισμούς στους οποίους υπόκεινται τόσο ο άνθρωπος όσο και ο υπολογιστής σε σχέση με την ποσότητα της πληροφορίας που μπορούν να χειριστούν ταυτόχρονα.

Σε αυτήν την κατηγορία μπορούμε να εντάξουμε και την αβεβαιότητα που οφείλεται σε σφάλματα μετρήσεων (δεν υπάρχει μετρητικό όργανο με άπειρη ακρίβεια) αλλά και σε σφάλματα αναπαράστασης (π.χ. αριθμός χρησιμοποιούμενων λέξεων για την αναπαράσταση πραγματικών αριθμών στον υπολογιστή). Τέτοια σφάλματα εισάγουν πληροφοριακή αβεβαιότητα σε ένα σύστημα.

❑ Γλωσσική αβεβαιότητα

Αποτελεί το πιο «γνώριμο» είδος αβεβαιότητας αφού το βιώνουμε και το χειριζόμαστε καθημερινά. Λέξεις όπως «μικρό» ή «μεγάλο», «πολύ» ή «λίγο» κ.λπ. εξαρτώνται από το αντικείμενο, το άτομο που τις χρησιμοποιεί, τη χρονική στιγμή, τα συμφραζόμενα κ.λπ.

Τα Κίνητρα της Μετάβασης

Η «κλειστή» αντιμετώπιση των προβλημάτων με την επιβολή ακριβών μαθηματικών μοντέλων είχε ως άμεσο αποτέλεσμα την κλιμάκωση της πολυπλοκότητας στους περισσότερους γνωστικούς κλάδους. Επιπρόσθετα, με την εξέλιξη της τεχνολογίας νέες ανάγκες εμφανίστηκαν και νέα προβλήματα προέκυψαν, τα οποία όμως παρουσιάζουν σημαντικά αυξημένη δυσκολία.

Οι παραδοσιακές τεχνικές και μεθοδολογίες αποδείχθηκαν ανίκανες να αντιμετωπίσουν αποτελεσματικά διάφορες κατηγορίες προβλημάτων, πολλά από τα οποία, παραδόξως, ο άνθρωπος καθημερινά αντιμετώπιζε και έλυνε με ιδιαίτερη ευκολία. Χαρακτηριστικά τέτοια παραδείγματα αποτελούν τα προβλήματα **αναγνώρισης**. Καθημερινές τετριμμένες διεργασίες όπως η αναγνώρισης ομιλίας ή η αναγνώρισης μιας γνώριμης μορφής σε ένα άγνωστο περιβάλλον (παραδείγματα που ο άνθρωπος επιλύει ασυναίσθητα σε κλάσματα του δευτερολέπτου), συνεχίζουν να θεωρούνται δυσεπίλυτα, ακόμη και σήμερα, ακόμη και για τα πιο εξελιγμένα υπολογιστικά συστήματα.

Άρση του δόγματος της ακρίβειας

Έτσι, γρήγορα το «δόγμα της ακρίβειας» που θεωρούταν αναπόσπαστο στοιχείο κάθε επιστημονικής προσέγγισης, έγινε αντιληπτό ότι όχι μόνο είναι κάτι μη επιτεύξιμο (τουλάχιστον όχι σε αυθαίρετο βαθμό) αλλά και **«εκτός θέματος»**· η περιγραφή της φύσης σε «κλειστή μορφή» ούτε είναι δυνατή αλλά ούτε αποτελεί και το στόχο. Ο στόχος είναι η σχεδίαση συστημάτων που να επιδεικνύουν ευστάθεια και ευρωστία συμπεριφοράς, κάποιο βαθμό νοημοσύνης και τελικά συνολική ικανότητα.

Σχηματοποιώντας τα παραπάνω, θα μπορούσαμε να πούμε ότι από την *ακριβή επίλυση μιας προσέγγισης του προβλήματος* που ήταν το ζητούμενο μιας επιστήμης επικεντρωμένης στα μοντέλα όπου το ακριβές πρόβλημα μετασχηματίζεται (μέσω απλοποιήσεων και παραδοχών) σε μια προσέγγισή του η οποία επιδέχεται κλειστής μοντελοποίησης, περάσαμε στην *προσεγγιστική επίλυση του ακριβούς προβλήματος*. Απευθυνόμαστε ευθέως στο ίδιο το πρόβλημα για το οποίο αναζητούμε προσεγγιστικές λύσεις.

Επιστροφή στη φύση και τον άνθρωπο

Υπό το πρίσμα των παραπάνω, για μια ακόμη φορά, η επιστήμη στράφηκε στη φύση και τον άνθρωπο, ως ζωντανά παραδείγματα «συστημάτων» που επιδεικνύουν ενδογενή ικανότητα επίλυσης ιδιαίτερα πολύπλοκων προβλημάτων και την «ύψιστη» μορφή νοημοσύνης: την ανθρώπινη νοημοσύνη. Εξάλλου αυτή είναι και η μόνη νοημοσύνη που μπορεί ο άνθρωπος να αντιληφθεί· κάθε άλλη μορφή νοημοσύνης κάθε άλλο παρά τέτοια θα φαινόταν στον άνθρωπο αφού θα αδυνατούσε να την αντιληφθεί. Συνεπώς, η προσπάθεια για εμφύτευση νοημοσύνης στις μηχανές δεν θα μπορούσε παρά να δανειστεί μεθόδους, τεχνικές και δομές από τη φύση και από τον άνθρωπο. Το δάνειο αυτό αφορά τόσο σε δομικά όσο και σε λειτουργικά πρότυπα.

Δομικά και λειτουργικά πρότυπα

Τα **δομικά πρότυπα** αφορούν στην φυσική (βιολογική) οργάνωση των μονάδων στις οποίες αποδίδεται νοημοσύνη. Τα **λειτουργικά πρότυπα** αφορούν στην «φαινομενική» επεξεργασία της πληροφορίας: πώς δηλαδή αυτή παρατηρείται από έξω, ανεξάρτητα από τις δομές που την «υποστηρίζουν».

Υπολογιστική Νοημοσύνη

Σε ένα κόσμο τεράστιας πολυπλοκότητας και ποικιλίας, τα γεγονότα ποτέ δεν επαναλαμβάνονται πανομοιότυπα. Δεν είναι όμως και τελείως διαφορετικά. Υπάρχει ένας κρίκος συνέχεις, ομοιότητας και προβλεψιμότητας που μας επιτρέπει να γενικεύουμε, συχνά με επιτυχία, από παρελθούσα εμπειρία προς μελλοντικά γεγονότα.

Λαμβάνοντας αυτό υπόψη, είναι εύκολο να δικαιολογήσουμε την ολοένα και αυξανόμενη απήχηση που έχουν οι αριθμητικές μέθοδοι αντιμετώπισης προβλημάτων του πραγματικού κόσμου· προβλήματα στα οποία η Τεχνητή Νοημοσύνη συχνά απέτυχε ακόμη και να απευθυνθεί.

Η **παρεμβολή** και η **μάθηση μέσω παραδειγμάτων** συντελούν στην «διαμόρφωση» μοντέλων με βάση αποκλειστικά και μόνο μια συλλογή από ζεύγη εισόδων / εξόδων. Τα ζεύγη αυτά αποτελούν ένα τρόπο περιγραφής της επιθυμητής συμπεριφοράς του υπό σχεδίαση συστήματος εναλλακτικό των μαθηματικών μοντέλων.

Σε ένα πλαίσιο σαν το παραπάνω, γεννήθηκε και αναπτύχθηκε η υπολογιστική νοημοσύνη (YN — Computational Intelligence, CI). Στην YN συγκαταλέγονται τρεις βασικές μεθοδολογίες:

- ❑ Τα Ασαφή Συστήματα
- ❑ Τα Νευρωνικά Δίκτυα
- ❑ Οι γενετικοί Αλγόριθμοι

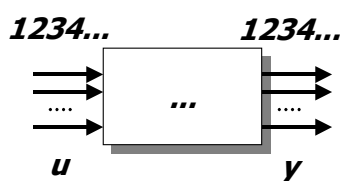
Αν ο άνθρωπος αλλά και η φύση μπορούν (εντελώς αφαιρετικά αλλά και με πλήρη σεβασμό προς τις αναλογίες) να θεωρηθούν ως συστήματα που επεξεργάζονται πληροφορίες, καθεμία από τις παραπάνω μεθοδολογίες έχουν εφελκύσει τη βάση τους από διαφορετικές πτυχές αυτών των «προτύπων επεξεργασίας».

Έτσι, η ασαφής λογική έχει δανειστεί λειτουργικά πρότυπα από την ανθρώπινη διαδικασία συλλογισμού και εξαγωγής συμπεράσματος, τα νευρωνικά δίκτυα έχουν δανειστεί το δομικό πρότυπο του ανθρώπινου εγκεφάλου, και οι γενετικοί αλγόριθμοι την ιδέα της εξέλιξης της φύσης, μεγάλο μέρος της οποίας ολοκληρώνεται στη θεωρία του Δαρβίνου για την εξέλιξη των ειδών.

Κοινά χαρακτηριστικά και των τριών μεθοδολογιών αποτελούν η ικανότητα για μάθηση και προσαρμογή αλλά και η ανεξαρτησία τους από μοντέλα. Γι αυτό το τελευταίο σημείο, καλούνται **εκτιμητές άνευ μοντέλου** (model-free estimators).

Στο σημείο αυτό θα πρέπει να τονίσουμε ότι αυτή η «ανεξαρτησία από μοντέλα» αφορά σε μοντέλα του προβλήματος το οποίο καλούμαστε να επιλύσουμε. Και οι τρεις μεθοδολογίες, όπως και κάθε άλλη μεθοδολογία, προτείνουν συστήματα με καθορισμένη δομή τα οποία μπορούν (ως έχουν) να αντιμετωπίσουν κάθε πρόβλημα (μιας συγκεκριμένης κατηγορίας) χωρίς την ανάγκη προσαρμογής της δομής τους στις ιδιαιτερότητες του κάθε συγκεκριμένου προβλήματος. Τα συστήματα αυτά **έχουν** συγκεκριμένη δομή η οποία συχνά καλείται και μοντέλο (για παράδειγμα, τα μοντέλα νευρωνικών δικτύων) τα οποία όμως **δεν** αποτελούν μοντέλα του προβλήματος. Αποτελούν αρχιτεκτονικά πρότυπα που προτείνονται για την σχεδίαση νοημόνων «μηχανών».

Πριν την αναλυτική παρουσίαση κάθε μίας από τις τρεις μεθοδολογίες, αυτές παρουσιάζονται πολύ συνοπτικά στις παρακάτω παραγράφους. Η μορφή των συστημάτων αυτών είναι κοινή και αναπαρίσταται στο παρακάτω σχήμα σε αναλογία με το Σχήμα 1. Βασικό χαρακτηριστικό της είναι ότι τόσο οι εισοδοί όσο και οι εξοδοί βρίσκονται σε αριθμητική μορφή.

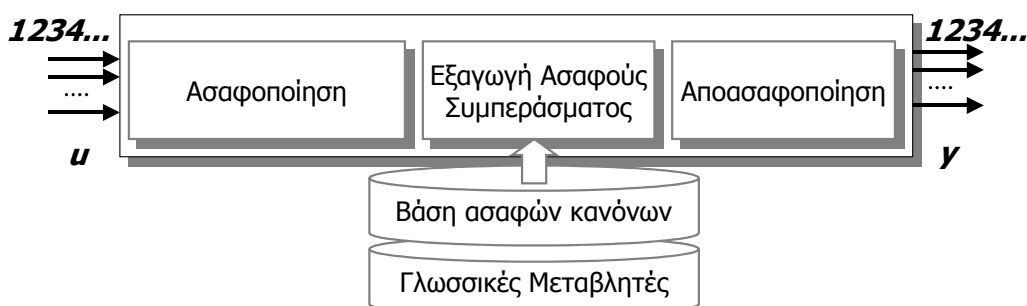


Σχήμα 4. Αφαιρετική αναπαράσταση ενός τυπικού συστήματος YN

Ασαφή Συστήματα

Τα Ασαφή Συστήματα (ΑΣ — Fuzzy Systems, FS) έχουν βασιστεί στην ανθρώπινη διαδικασία συλλογισμού. Μιμούνται, δηλαδή, τις ανθρώπινες διεργασίες εξαγωγής συμπεράσματος σε συνθήκες αβεβαιότητας. Αναπαριστώντας τη γνώση κατάλληλα και με διαφανή τρόπο μπορούν να αξιοποιήσουν την όποια διαθέσιμη ανθρώπινη εμπειρία στο τομέα του προβλήματος. Τέτοια γνώση μπορεί να ενσωματωθεί σε ένα ασαφές σύστημα με τη μορφή αιτιακών κανόνων διατυπωμένων σε σχεδόν φυσική γλώσσα.

Σε αναλογία με την αφαιρετική αναπαράσταση που δίνεται στο Σχήμα 4, ένα ασαφές σύστημα μπορεί να απεικονιστεί με τον παρακάτω τρόπο.

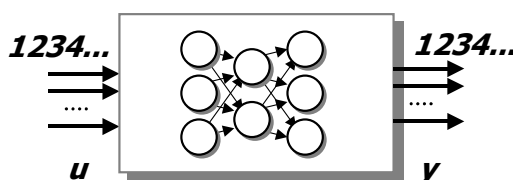


Σχήμα 5. Αφααιρετική αναπαράσταση ενός ασαφούς συστήματος

Νευρωνικά Δίκτυα

Τα Νευρωνικά Δίκτυα (ΝΔ — Neural Networks, NN) μιμούνται τη δομή του ανθρώπινου εγκεφάλου. Αποτελούν μια μαζική συνάθροιση απλούστατων και πυκνά αλληλοσυνδεδεμένων στοιχείων, των νευρώνων, σε αναλογία με τα βιολογικά νευρικά κύτταρα. Τα ΝΔ μπορούν να αξιοποιήσουν αριθμητικά δεδομένα σε μορφή ζευγών εισόδου / εξόδου, όπως παραδειγματικά πρότυπα, ιστορικά δεδομένα ή και μετρήσεις.

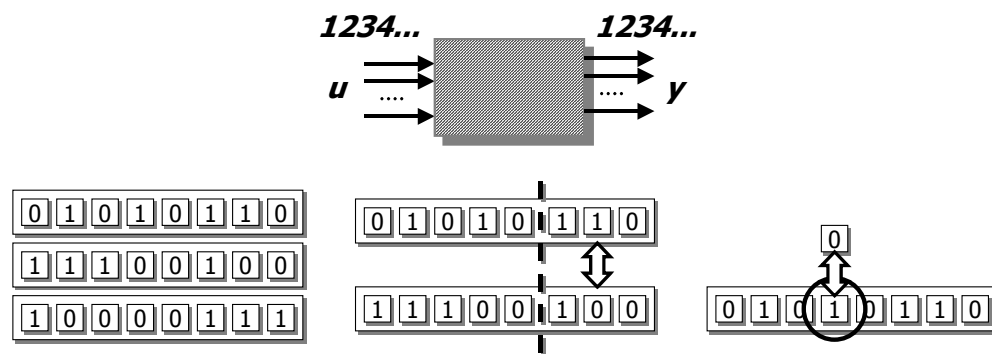
Σε αναλογία με την αφαιρετική αναπαράσταση που δίνεται στο Σχήμα 4, ένα νευρωνικό δίκτυο μπορεί να απεικονιστεί με τον παρακάτω τρόπο.



Σχήμα 6. Αναπαράσταση ενός νευρωνικού δικτύου

Γενετικοί Αλγόριθμοι

Οι Γενετικοί Αλγόριθμοι (ΓΑ — Genetic Algorithms, GA) μιμούνται την εξελικτική διαδικασία των βιολογικών συστημάτων, όπως αυτή παρουσιάζεται στη θεωρία της εξέλιξης των ειδών του Δαρβίνου. Αναπαριστούν τις πιθανές τιμές για τις παραμέτρους ενός προβλήματος με συμβολοσειρές αποτελούμενες συνήθως από δύο μόνο σύμβολα: το 0 και το 1. Οι τιμές αυτές των παραμέτρων (που αποτελούν και πιθανές λύσεις) αρχικοποιούνται με τυχαίες τιμές και υποβάλλονται σε μια διαδικασία εξέλιξης με τη βοήθεια τελεστών όπως η μετάλλαξη και η διασταύρωση σε μια προσπάθεια να εντοπιστούν συνεχώς βελτιωμένες νέες τιμές για τις παραμέτρους που, έπειτα από επαρκή εξέλιξη, θα αποτελέσουν και την λύση στο πρόβλημα.



Σχήμα 7. Αφαιρετική αναπαράσταση γενετικών αλγορίθμων

Ασαφής Λογική

Εισαγωγή

Η θεωρία συνόλων κατέχει κεντρική θέση στα μαθηματικά και βρίσκεται στη βάση της λογικής. Κι όμως, δεν ήταν λίγες οι φορές που η θεωρία αυτή αμφισβητήθηκε ως προς την ισχύ της ενώ διάφορα παράδοξα στα οποία οδηγείται έχουν παραμείνει ιστορικά. Αν και τα παράδοξα αυτά, όταν διατυπώθηκαν, είχαν θεωρηθεί σοφιστείες, όλα τους είχαν το κέντρο τους στο σημείο ανάμεσα στο 0 και το 1: τις δύο μοναδικές τιμές που προβλέπει η κλασσική συνολοθεωρία για να περιγράψει εάν ένα σημείο ενός πεδίου ορισμού ανήκει σε ένα ορισμένο σύνολο (1) ή όχι (0).

Από τη Θέση στην Αντίθεση — Από το *Α* στο *όχι-Α*

«Αυτός είναι ένας σωρός άμμου!»

Πολύ συχνά όμως η δυαδική αυτή λογική αδυνατεί να περιγράψει καθημερινά φαινόμενα και έρχεται σε αντίθεση με την «κοινή» λογική. Για παράδειγμα, δεν είναι καθόλου δύσκολο να χαρακτηρίσουμε μια μεγάλη ομάδα κόκκων άμμου ως ένα «σωρό άμμου»: είναι κάτι «προφανές» που με μια ματιά μπορούμε να διαπιστώσουμε. Ας θεωρήσουμε τώρα ότι αρχίζουμε να αφαιρούμε έναν-έναν τους κόκκους από το σωρό και σε κάθε βήμα θέτουμε την ερώτηση: *αυτό που απομένει είναι ένας «σωρός άμμου»;*

Σίγουρα, για ένα πολύ μεγάλο αριθμό από τέτοια βήματα η απάντηση θα παραμείνει ρητή: «ναι». Για πόσο όμως; Θα φτάσει κάποια στιγμή που αναγκαστικά θα πρέπει να απαντήσουμε «όχι» γιατί τον μόνο που θα έχει απομείνει είναι μια μικρή ομάδα κόκκων που σίγουρα δεν αποτελούν σωρό.

Και εδώ έρχεται το κρίσιμο ζήτημα: ποια στιγμή θα μεταβάλλουμε τη απάντησή μας από «ναι» σε «όχι»; Στο βήμα εκείνο, το μόνο που κάναμε ήταν να απομακρύνουμε ένα και μόνο κόκκο άμμου. Κι όμως αυτό το γεγονός αποτέλεσε την ειδοποιό διαφορά ώστε από ένα σωρό άμμου να περάσουμε στο μη-σωρό (μιλώντας αυστηρά). Είναι φανερό ότι μια τέτοια αντιμετώπιση δημιουργεί προβλήματα, ιδίως όταν ξεφεύγουμε από τον «ιδεατό» και πλήρως δομημένο χώρο των μαθηματικών για να προσανατολιστούμε σε πραγματικά προβλήματα.

Το βασικό θέμα στο παραπάνω πρόβλημα είναι η αυστηρότητα της κλασσικής συνολοθεωρίας στον ορισμό συνόλων: οι μόνες εναλλακτικές για την ομάδα των κόκκων άμμου είναι δύο: είτε ανήκουν στο σύνολο των «σωρών» είτε όχι. Το σύνολο των «σωρών» όπως και κάθε άλλο σύνολο έχει αυστηρά όρια και μια αμελητέα διαφοροποίηση μπορεί να αποβεί καθοριστική για τη συμμετοχή ενός στοιχείου στο σύνολο ή όχι.

Το σύνολο των «ψηλών ανθρώπων»

Ένα δεύτερο παράδειγμα μπορεί να βοηθήσει στην καλύτερη κατανόηση του θέματος αυτού. Ας θεωρήσουμε το σύνολο των «ψηλών ανθρώπων». Στην προσπάθειά μας να ορίσουμε το σύνολο αυτό, το πιο πιθανό είναι να καταφύγουμε σε μια στατιστική έρευνα για να εντοπίσουμε το μέσο ύψος των ανθρώπων και τότε θα λέγαμε ότι ένας άνθρωπος ανήκει στο σύνολο αυτό (δηλαδή είναι ψηλός) εάν και μόνο εάν το ύψος του είναι μεγαλύτερο από το μέσο όρο που υπολογίσαμε. Έτσι, αν το μέσο ύψος είναι 172 εκ. τότε όποιος είναι ψηλότερος από αυτό θεωρείται ψηλός ενώ όλοι οι υπόλοιποι όχι.

Και εδώ εμφανίζεται και πάλι το πρόβλημα των αυστηρών συνόρων ενός κλασσικού συνόλου: Ένας άνθρωπος με ύψος 171,9 δεν είναι «ψηλός» ενώ ένας άνθρωπος που είναι μόλις 1 χιλιοστό ψηλότερος είναι.

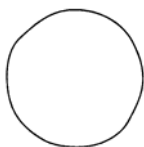
Είναι φανερό ότι μια τέτοια αυστηρή συνολοθεωρία παγιδευμένη στις μοναδικές εναλλακτικές του 0 και του 1, έχει πολύ συχνά δυσκολία στο να εκφράσει την πραγματικότητα.

Ο Ασαφής Κόσμος

Ο ασαφής κύκλος

Η δυσκολία αυτή μπορεί να γίνει αντιληπτή και από το παρακάτω παράδειγμα:

Ποια πρόταση από τις παρακάτω είναι πιο ακριβής στο να εκφράσει το διπλανό σχήμα;



(α) Το σχήμα πιθανώς να είναι ένας κύκλος

(β) Το σχήμα είναι ένας ασαφής κύκλος

Παράδοξα Κλασσικής Λογικής

Κατά καιρούς ένα σύνολο από «παράδοξα» έχουν διατυπωθεί. Τα παραδείγματα αυτά αναφέρονται ως παράδοξα επειδή η κλασσική λογική αδυνατεί να τα «περιγράψει». Το κοινό στοιχείο τους είναι ότι κατά την προσπάθεια περιγραφής του με μαθηματικό φορμαλισμό οι εξισώσεις που καταστρώνονται καταλήγουν σε σχέση της μορφής:

$$p = \text{όχι-}p$$

όπου το p όχι μόνο είναι περιορισμένο στο διάστημα $[0, 1]$ αλλά μπορεί να λάβει μόνο τις τιμές 0 ή 1. Προφανώς η λύση σε τέτοια προβλήματα είναι θεωρητικά το σημείο $\frac{1}{2}$ (αφού $\text{όχι-}p = 1-p$).

- Κάποιος Κρητικός έλεγε:
"Όλοι οι Κρητικοί λένε ψέματα"
Πρέπει να τον πιστέψουμε;
- Ο κουρέας του Russell έλεγε:
"Κουρεύω μόνο όσους δεν κουρεύονται μόνοι τους"
Ποιος κουρεύει τον κουρέα;
- Ένας κανόνας λέει:
"Όλοι οι κανόνες έχουν εξαιρέσεις"
Ισχύει αυτός ο κανόνας;

Διάφοροι εξέχοντες επιστήμονες έχουν κατά καιρούς εκφράσει την αδυναμία των μαθηματικών να απευθυνθούν σε ρεαλιστικά προβλήματα καθώς και την ενδογενή αβεβαιότητα των ρεαλιστικών προβλημάτων.

- A. Einstein
"Όταν οι νόμοι των μαθηματικών αναφέρονται στην πραγματικότητα δεν είναι βέβαιοι. Κι όταν είναι βέβαιοι δεν αναφέρονται στην πραγματικότητα."
- L. Zadeh
"Καθώς εγείρεται η πολυπλοκότητα, οι ακριβείς προτάσεις τείνουν να χάσουν την σημασία τους και οι σημαντικές προτάσεις τείνουν να χάνουν τη ακρίβειά τους."

Ασαφή Σύνολα έναντι Κλασσικών Συνόλων

Η θεμελίωση της Θεωρίας των Ασαφών Συνόλων οφείλεται στον Lotfi Zadeh που τα εισήγαγε με μια δημοσίευσή του το 1965 προκαλώντας ποικίλες αντιδράσεις. Στη θεωρία του, πρότεινε έναν διαφορετικό ορισμό για το σύνολο εισάγοντας την έννοια του **βαθμού συμμετοχής** ενός στοιχείου σε ένα σύνολο και επιτρέποντας στο βαθμό αυτό να λαμβάνει τιμές όχι μόνο 0 (για μη συμμετοχή) και 1 (για συμμετοχή) αλλά και όλες τις (άπειρες) ενδιάμεσες. Επέτρεψε έτσι την μερική συμμετοχή ενός στοιχείου σε ένα σύνολο.

Ο ορισμός του συνόλου αποτελεί τη βάση της συνολοθεωρίας η οποία με τη σειρά της στηρίζει τη μαθηματική λογική. Έτσι, η υιοθέτηση ενός ριζικά διαφορετικού ορισμού για το σύνολο, είχε αλυσιδωτές αντιδράσεις επηρεάζοντας καταλυτικά και όλες τις στηριζόμενες σε αυτό θεωρίες. Για παράδειγμα, το τρίπτυχο των βασικών αξιωμάτων της κλασσικής λογικής δεν είχε πλέον ισχύ στην νέα θεωρία των ασαφών συνόλων.

Ορισμός Συνόλου

Τα κλασσικά μαθηματικά, προσφέρουν τρεις τρόπους για τον ορισμό ενός συνόλου:

1. **Μέσω απαρίθμησης**
Ένα σύνολο ορίζεται παραθέτοντας εξαντλητικά όλα τα στοιχεία που το απαρτίζουν.
2. **Μέσω μιας κοινής ιδιότητας**
Για παράδειγμα, το σύνολο των άρτιων φυσικών αριθμών.
3. **Μέσω μιας χαρακτηριστικής συνάρτησης**
Για παράδειγμα, το σύνολο των ανθρώπων που έχουν ύψος πάνω από 172 εκατοστά.

Είναι σημαντικό να σημειωθεί ότι τα στοιχεία που ανήκουν σε ένα σύνολο, ανήκουν **πλήρως** σε αυτό (με βαθμό συμμετοχής 1) ενώ τα υπόλοιπα δεν ανήκουν «καθόλου» (βαθμός συμμετοχής 0). Έτσι, ο

ορισμός ενός συνόλου αποτελεί μια απεικόνιση από έναν πληθυσμό (όλα τα πιθανά στοιχεία) στις δύο τιμές $\{0,1\}$. Η απεικόνιση αυτή καλείται **χαρακτηριστική συνάρτηση** και για ένα κλασσικό σύνολο έχει τη μορφή:

$$X \rightarrow \{0,1\}$$

Συνάρτηση
συμμετοχής

Τα ασαφή σύνολα, αίρουν αυτόν τον περιορισμό, επεκτείνοντας την χαρακτηριστική συνάρτηση στην **συνάρτηση συμμετοχής**. Η συνάρτηση συμμετοχής λαμβάνει τιμές στο κλειστό διάστημα $[0,1]$:

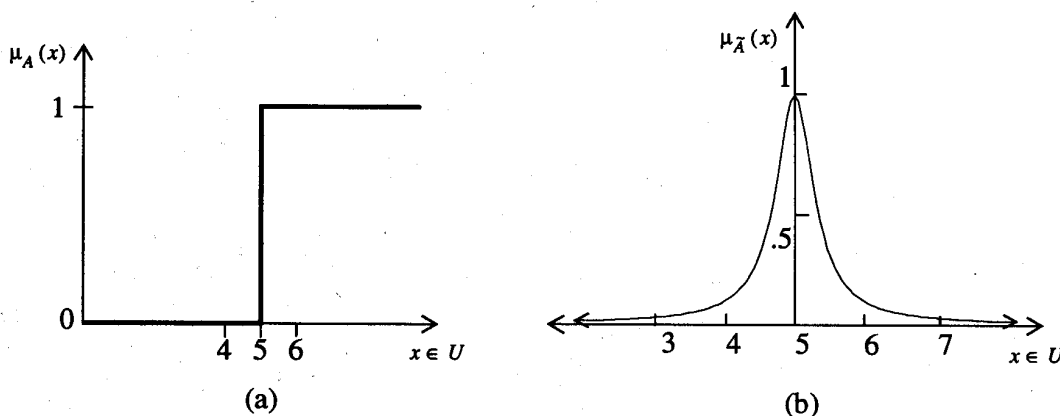
$$X \rightarrow [0,1]$$

Έτσι, η κλασσική λογική μπορεί να θεωρηθεί ως μια ειδική περίπτωση της ασαφούς λογικής αφού η δεύτερη την εμπεριέχει.

Ως παράδειγμα, ας θεωρήσουμε τα δύο παρακάτω σύνολα:

1. Το κλασσικό σύνολο των αριθμών που είναι μεγαλύτεροι του 5
2. Το ασαφές σύνολο των αριθμών που είναι κοντά στο 5

Η χαρακτηριστική συνάρτηση του πρώτου και η συνάρτηση συμμετοχής του δεύτερου δίνονται στο παρακάτω σχήμα.



(a) Το κλασσικό σύνολο των αριθμών που είναι μεγαλύτεροι του 5 (b) Το ασαφές σύνολο των αριθμών που είναι κοντά στο 5

Σχήμα 8. Παράδειγμα (a) κλασσικού και (b) ασαφούς συνόλου

Παράδειγμα

Ας θεωρήσουμε το πρόβλημα της σχεδίασης ενός ελεγκτή για ένα κλιματιστικό σύστημα. Ο ελεγκτής αυτός θα χρησιμοποιεί τις μετρήσεις ενός αισθητήρα για τη θερμοκρασία και με βάση αυτές θα ρυθμίζει την ένταση λειτουργίας του ώστε να διατηρεί σταθερή τη θερμοκρασία ενός δωματίου. Είναι αρκετά απλό να περιγράψουμε τον επιθυμητό αλγόριθμο λειτουργίας:

Όσο πιο υψηλή είναι η θερμοκρασία του δωματίου, τόσο πιο έντονα θα πρέπει να λειτουργήσει το κλιματιστικό ώστε να καταφέρει γρήγορα να την φέρει στα επιθυμητά επίπεδα. Όταν η θερμοκρασία φτάσει στα επιθυμητά επίπεδα, η λειτουργία του κλιματιστικού θα πρέπει να σταματήσει.

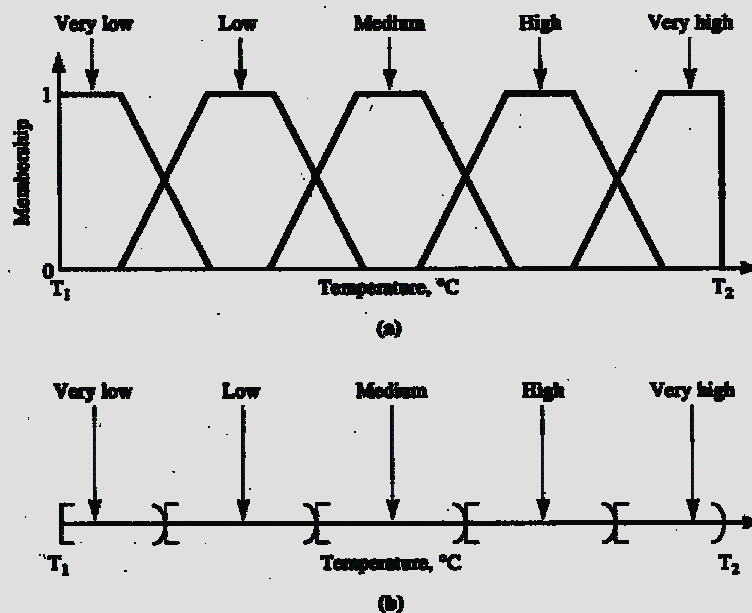
Η απλή αυτή λογική προϋποθέτει την κατάτμηση του πεδίου τιμών των θερμοκρασιών σε ένα αριθμό συνόλων καθένα από τα οποία θα αντιπροσωπεύει περιοχές θερμοκρασιών όπως «πολύ υψηλή», «υψηλή», κλπ.

Με τη βοήθεια αυτών των συνόλων μπορούμε να γράψουμε τους παρακάτω εμπειρικούς κανόνες:

1. Αν η θερμοκρασία είναι «πολύ υψηλή» τότε η ένταση λειτουργίας πρέπει να είναι «υψηλή»
2. Αν η θερμοκρασία είναι «υψηλή» τότε η ένταση λειτουργίας πρέπει να είναι «κανονική»
3. Αν η θερμοκρασία είναι «μεσαία» τότε η λειτουργία πρέπει να διακοπεί.

Στο παρακάτω σχήμα δίδονται οι δύο εναλλακτικές: ασαφή σύνολα και κλασσικά σύνολα. Από το

σχήμα αυτό γίνεται φανερή η διαφορά μεταξύ της ασαφούς και της κλασσικής προσέγγισης.



Στην κλασσική προσέγγιση ορίζουμε διαστήματα. Όταν η θερμοκρασία βρίσκεται στο διάστημα που είναι σημειωμένο με «very high», τότε θα πυροδοτηθεί ο κανόνας 1 ενώ οι κανόνες 2 και 3 δεν θα πυροδοτούνται καθόλου αφού η παρούσα θερμοκρασία δεν ανήκει καθόλου ούτε στο σύνολο των «υψηλών» θερμοκρασιών αλλά ούτε και στο σύνολο των «μεσαίων» θερμοκρασιών. Καθώς η θερμοκρασία πέφτει και όταν φτάσει στο όριο των δύο συνόλων, απότομα θα σταματήσει να πυροδοτείται ο κανόνας 1 ενώ θα ενεργοποιηθεί ο κανόνας 2.

Στην ασαφή προσέγγιση, στα σύνολα έχει επιτραπεί να «επεκταθούν». Έτσι μια θερμοκρασία μπορεί να ανήκει ταυτόχρονα και στο σύνολο των «πολύ υψηλών» θερμοκρασιών αλλά και στο σύνολο των «υψηλών» με διαφορετικό ίσως βαθμό συμμετοχής. Έτσι, σε κάποια δεδομένη στιγμή δύο ή και περισσότεροι κανόνες μπορεί να είναι ταυτόχρονα ενεργοί επηρεάζοντας τη συνολική απόκριση του συστήματος. Συνεπώς, η μετάβαση μεταξύ των διαφόρων περιοχών των θερμοκρασιών είναι πλήρως ομαλή.

Πράξεις σε Ασαφή Σύνολα

Οι πράξεις μεταξύ συνόλων όπως αυτές ορίζονται στην κλασσική λογική, μπορούν να επεκταθούν και στα ασαφή σύνολα. Έτσι, μπορούμε να ορίσουμε τις τρεις παρακάτω βασικές πράξεις.

Έστω δύο ασαφή σύνολα A και B . Τότε ορίζουμε τις εξής πράξεις:

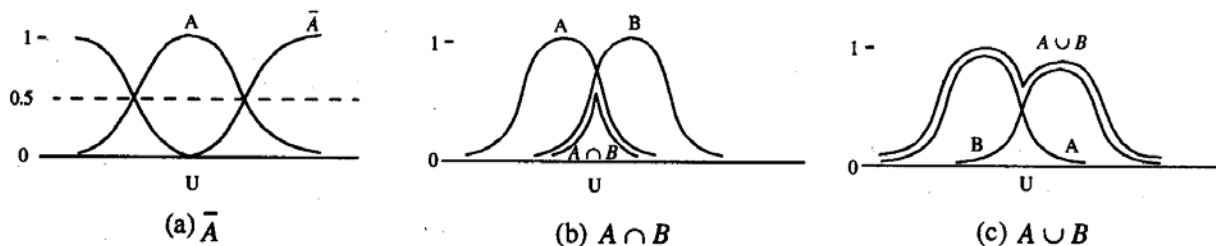
- **Συμπλήρωμα $\bar{A}$:**

$$\bar{A}(x) = 1 - A(x)$$
- **Ένωση $A \cup B$:**

$$(A \cup B)(x) = \max\{A(x), B(x)\}$$
- **Τομή $A \cap B$:**

$$(A \cap B)(x) = \min\{A(x), B(x)\}$$

Μια γραφική αναπαράσταση των πράξεων αυτών δίνεται στο παρακάτω σχήμα.



Σχήμα 9. Πράξεις ασαφών συνόλων. (α) Συμπλήρωμα, (β) τομή και (γ) ένωση

Οι παραπάνω πράξεις είναι κατάλληλα ορισμένες ώστε όταν τα σύνολα A και B είναι κλασσικά και όχι ασαφή να έχουν τα ίδια αποτελέσματα με τις αντίστοιχες πράξεις της κλασσικής συνολοθεωρίας.

Είναι σημαντικό επίσης να σημειωθεί ότι οι παραπάνω ορισμοί δεν είναι μοναδικοί. Κατά καιρούς έχουν προταθεί διάφοροι άλλοι αλλά αυτοί είναι οι απλούστεροι και πλέον χρησιμοποιούμενοι.

Ασαφής Λογική

Λογικές Πράξεις

Η ασαφής λογική θεμελιώνεται στις τρεις πράξεις:

- **Λογικό και** ($\wedge$, αντιστοιχεί σε τομή)
- **Λογικό ή** ($\vee$, αντιστοιχεί σε ένωση),
- **Λογική άρνηση** ($\neg$, αντιστοιχεί σε συμπλήρωμα)

Λογική

Λογική είναι η μελέτη των συλλογιστικών μεθόδων και της διαδικασίας εξαγωγής συμπερασμάτων από δεδομένα.

Προτασιακός Λογισμός

Ο κλασσικός **Προτασιακός Λογισμός** είναι η παραγωγή συλλογιστικής με το σχηματισμό προτάσεων από λογικές μεταβλητές οι οποίες (προτάσεις) αποτιμούνται ως *αληθείς* ή *ψευδείς*.

Ασαφής Λογική

Η **Ασαφής Λογική** «επανατοποθετεί» το θέμα αυτό επιτρέποντας σε μια πρόταση να λάβει οποιοδήποτε βαθμό αλήθειας μεταξύ του ψευδούς (0) και αληθούς (1). Η μελέτη της ασαφούς λογικής αναλύεται στα παρακάτω τρία συστατικά:

- Τις Γλωσσικές Μεταβλητές
- Τις Ασαφείς Προτάσεις
- Την Εξαγωγή συμπεράσματος από ασαφείς προτάσεις

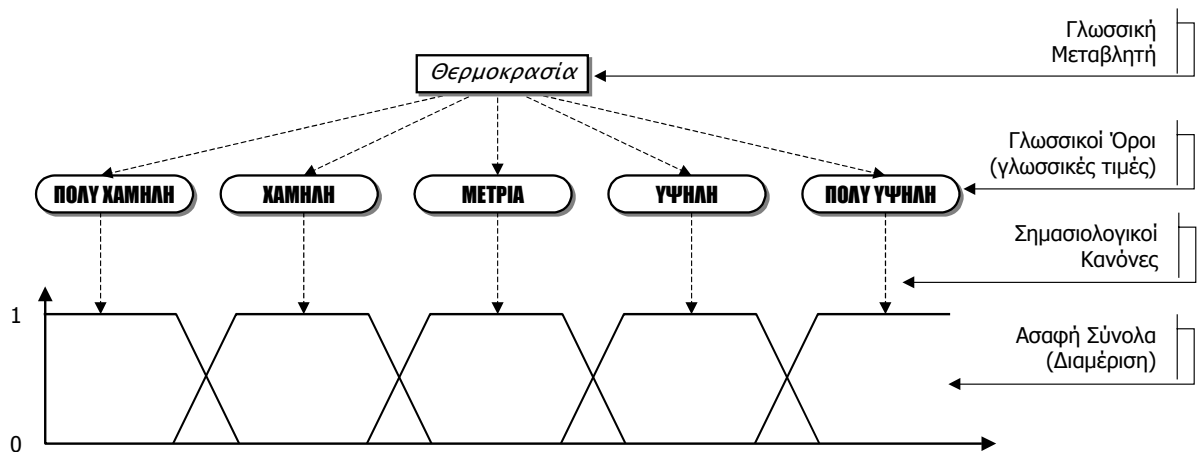
Καθένα από αυτά τα θέματα εξετάζονται ξεχωριστά στις επόμενες παραγράφους.

Γλωσσικές Μεταβλητές

Στο εσωτερικό ενός ασαφούς συστήματος, η διαδικασία συλλογισμού και εξαγωγής συμπεράσματος επεξεργάζεται ασαφή σύνολα και όχι αριθμούς: τόσο οι εμπειρικοί αιτιακοί κανόνες που καθορίζουν την συμπεριφορά του συστήματος όσο και οι ίδιες οι είσοδοι και έξοδοι είναι ασαφή σύνολα.

Γλωσσικοί όροι/ τιμές (linguistic terms/ values)

Στο παράδειγμα της προηγούμενης παραγράφου (ελεγκτής κλιματιστικού συστήματος), για να περιγράψουμε την επιθυμητή συμπεριφορά χρησιμοποιήσαμε λεκτικές περιγραφές όπως «μεσαία», «υψηλή», «πολύ υψηλή» κ.λπ. Οι όροι αυτοί ονομάζονται **γλωσσικοί όροι** και θα πρέπει, προφανώς, να αντιστοιχούν σε ρητά ορισμένες οντότητες ώστε να μπορέσει το σύστημα να τους ερμηνεύσει και να τους χρησιμοποιήσει. Στο παρακάτω σχήμα (Σχήμα 10) φαίνεται η αντιστοίχιση γλωσσικών όρων με κατάλληλα ορισμένα ασαφή σύνολα μέσω των λεγόμενων σημασιολογικών κανόνων.



Σχήμα 10. Ορισμός γλωσσικών όρων και γλωσσικών μεταβλητών

Συμπερασματικά, μπορούμε να πούμε ότι μέσω των γλωσσικών όρων δημιουργούμε μια μορφή «ορολογίας» κοινής για εμάς και το σύστημα. Έτσι, όταν στη συνέχεια γράψουμε έναν αιτιακό κανόνα της μορφής:

ΕΑΝ η *Θερμοκρασία* είναι **ΥΨΗΛΗ** ΤΟΤΕ η *Ένταση Λειτουργίας* είναι **ΥΨΗΛΗ**

η σημασία του κανόνα είναι σαφώς καθορισμένη.

Γλωσσική μεταβλητή
(linguistic variable)

Η μεταβλητή *θερμοκρασία* στον παραπάνω ορισμό λαμβάνει ως τιμές γλωσσικούς όρους. Έτσι, αναφερόμαστε πλέον στην **γλωσσική μεταβλητή** *θερμοκρασία* (linguistic variable).

Ασαφείς Προτάσεις

Η απλούστερη μορφή μιας **πρότασης** είναι η:

***V* είναι P**

όπου *V* γλωσσική μεταβλητή: **υποκείμενο** και
P γλωσσικός όρος: **κατηγόρημα**.

Βαθμός αλήθειας

Στην κλασσική λογική κάθε πρόταση αποτιμάται είτε ως αληθής (1) είτε ως ψευδής (0). Η δυαδικότητα αυτή αίρεται από την ασαφή λογική: στις προτάσεις αποδίδεται κάποιος **βαθμός αλήθειας** (ή, ισοδύναμα, ψεύδους). Έτσι, η αποτίμηση μια ασαφούς πρότασης μπορεί να λάβει οποιαδήποτε τιμή στο διάστημα [0,1].

Υποθετικές
προτάσεις

Μια μορφή πιο σύνθετων προτάσεων αποτελούν οι **υποθετικές προτάσεις**:

ΑΝ *A* είναι **A ΤΟΤΕ *B* είναι **B**** ($A \Rightarrow B$)

όπου *A, B* γλωσσικές μεταβλητές και
A, B γλωσσικοί όροι

Εξαγωγή Συμπεράσματος από Ασαφείς Προτάσεις

"Τρόπος του θέτειν"
(modus ponens)

Ένας τυπικός κανόνας εξαγωγής συμπεράσματος στην κλασσική λογική, είναι ο «**Τρόπος του θέτειν**» (modus ponens). Με βάση κάποιον κανόνα (γνώση) και κάποιο γεγονός (κάποια δεδομένα εισόδου, παρατήρηση) εξαγεί συμπέρασμα (αποφαινεται για την έξοδο):

$(A \Rightarrow B) \wedge A \Rightarrow B$

Λεκτικά, αυτό μπορεί να αποδοθεί ως εξής:

Αν (1) *γνωρίζω ότι το A συνεπάγεται το B* και (2) *παρατηρώ το A*,
τότε *συμπεραίνω το B*.

Ο κανόνας αυτός, αν και φαινομενικά απλοϊκός, αποτελεί βασικό εργαλείο για τη διαδικασία συλλογισμού και εξαγωγής συμπεράσματος.

Το πρώτο μέρος του κανόνα αυτού αποτελείται από (1) μια υποθετική πρόταση ($A \Rightarrow B$) και (2) μια απλή δηλωτική πρόταση A . Όπως προαναφέρθηκε, στην κλασσική λογική κάθε πρόταση μπορεί να αποτιμηθεί με δύο και μόνο τιμές: 0 ή 1. Αυτό σημαίνει ότι εάν δεν παρατηρήσουμε (μετρήσουμε) **ακριβώς** το A τότε δεν μπορούμε να συμπεράνουμε τίποτε για το B , ακόμη και εάν αυτό που θα παρατηρήσουμε είναι (αυθαίρετα) κοντά στο A .

«Γενικευμένος
τρόπος του θέτειν»
(generalised modus
ponens)

Υιοθετώντας βαθμούς αλήθειας, η ασαφής λογική ξεφεύγει από τους περιορισμούς της κλασσικής λογικής: μια πρόταση μπορεί να ισχύει εν μέρει. Έτσι, απαιτείται η επέκταση του «Τρόπου του θέτειν» ώστε να μπορέσει να αξιοποιήσει το χαρακτηριστικό αυτό προβαίνοντας σε συμπεράσματα ακόμη και όταν οι υποθέσεις δεν ισχύουν πλήρως. Το αποτέλεσμα είναι ο «**Γενικευμένος τρόπος του θέτειν**» (generalised modus ponens):

$$(A \Rightarrow B) \wedge A' \Rightarrow B'$$

ή

ΚΑΝΟΝΑΣ: ΑΝ A είναι A ΤΟΤΕ B είναι B

ΓΕΓΟΝΟΣ: A είναι A'

ΣΥΜΠΕΡΑΣΜΑ: B είναι B'

Λεκτικά, ο κανόνας αυτός μπορεί να αποδοθεί ως:

Αν (1) γνωρίζω ότι το A συνεπάγεται το B και (2) παρατηρώ κάτι «κοντά» στο A , τότε συμπεραίνω κάτι «κοντά» το B .

Όσο πιο «κοντά» είναι το A' στο A , τόσο πιο «κοντά» θα είναι το B' στο B .

Ασαφής Συστήματα

Μηχανή εξαγωγής
ασαφούς
συμπεράσματος

Η καρδιά ενός ασαφούς συστήματος είναι η **μηχανή εξαγωγής ασαφούς συμπεράσματος**. Η μηχανή αυτή χρησιμοποιεί τις μεθόδους που περιγράφηκαν προηγουμένως για να παράγει εξόδους με βάση τις εισόδους, τις γλωσσικές μεταβλητές που έχουν οριστεί και τους κανόνες που έχουν εισαχθεί.

Όλη η επεξεργασία σε ένα ασαφές σύστημα πραγματοποιείται χρησιμοποιώντας ασαφή σύνολα. Όμως, ένα ασαφές σύστημα καλείται να λειτουργήσει σε ένα περιβάλλον όπου οι εισοδοί του προέρχονται από κάποιους αισθητήρες ή άλλες μετρητικές συσκευές που μπορούν να παρέχουν μόνο βαθμωτά μεγέθη (απλές τιμές) και οι έξοδοί του θα πρέπει να προωθηθούν σε κάποιες άλλες συσκευές που δέχονται μόνο βαθμωτές τιμές.

Ασαφοποίηση και
αποασαφοποίηση

Συνεπώς, κάποια μετατροπή από βαθμωτά μεγέθη σε ασαφή σύνολα απαιτείται κατά την είσοδο των δεδομένων και η ανάστροφη μετατροπή κατά την έξοδο. Τις μετατροπές αυτές αναλαμβάνουν δύο αντίστοιχα υποσυστήματα, ονομαστικά το **υποσύστημα ασαφοποίησης** (fuzzification) και το **υποσύστημα αποασαφοποίησης** (defuzzification).

Για τη μετατροπή ενός ασαφούς συνόλου σε βαθμωτό, το υποσύστημα αποασαφοποίησης μπορεί να χρησιμοποιήσει διάφορες εναλλακτικές μεθόδους. Η πιο απλή και κοινή είναι η αποασαφοποίηση με τη μέθοδο του κέντρου βάρους.



Σχήμα 11. Ένα τυπικό ασαφές σύστημα

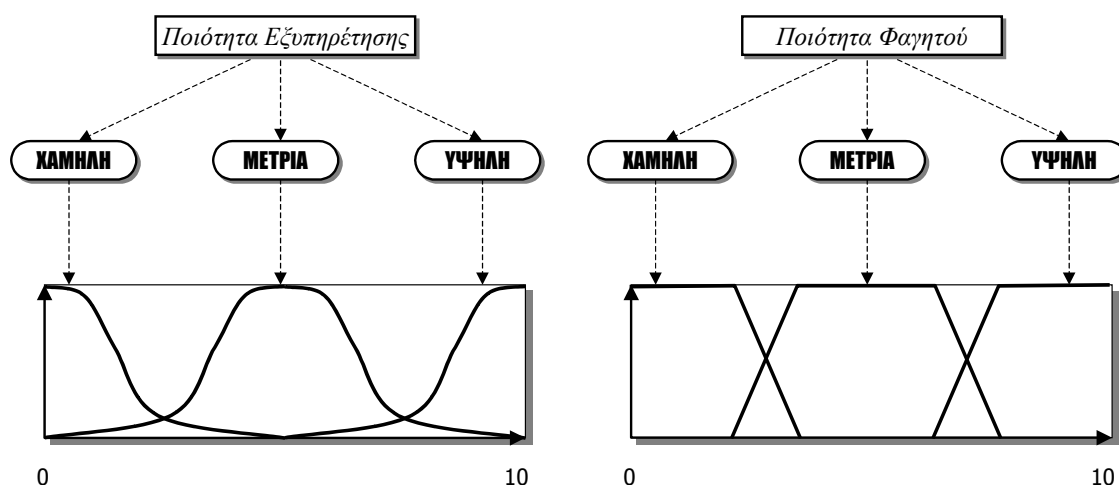
Παράδειγμα Εξαγωγής Ασαφούς Συμπεράσματος

Ας θεωρήσουμε το παρακάτω απλοϊκό παράδειγμα¹:

Επιθυμούμε να σχεδιάσουμε ένα σύστημα που να αποφασίζει αυτόματα για το φιλοδώρημα που πρέπει να αφήσουμε σε ένα εστιατόριο με βάση (α) την ποιότητα της εξυπηρέτησης και (β) την ποιότητα του φαγητού. Θεωρούμε ότι η ποιότητα εξυπηρέτησης και φαγητού (είσοδοι) περιγράφονται από ένα βαθμό στην κλίμακα 0-10, ενώ το φιλοδώρημα (έξοδος) μπορεί να λάβει τιμές από 0% ως 25% του ποσού του λογαριασμού.

Η απόφαση για το ύψος του φιλοδωρήματος είναι σχετικά απλή και βασίζεται σε απλούς κανόνες. Είναι, για παράδειγμα, προφανές ότι εάν η ποιότητα της εξυπηρέτησης είναι χαμηλής τότε το ύψος του φιλοδωρήματος θα την ακολουθήσει. Βλέπουμε δηλαδή ότι άμεσα μπορούμε να γράψουμε κάποιους αιτιακούς κανόνες για το συγκεκριμένο πρόβλημα. Πριν το κάνουμε όμως αυτό, θα πρέπει να «ορίσουμε τη γλώσσα» μας, δηλαδή να ορίσουμε τους γλωσσικούς όρους.

Επιλέγοντας αυθαίρετα μορφή και πυκνότητα για τα ασαφή σύνολα, καταλήγουμε σε μια διαμέριση σαν αυτή του παρακάτω σχήματος για τις δύο μεταβλητές εισόδου.



Σχήμα 12. Ορισμός γλωσσικών όρων παραδείγματος

Για την ποιότητα εξυπηρέτησης επιλέξαμε σιγμοειδή ασαφή σύνολα και για την ποιότητα του φαγητού τραπεζοειδή. Για την έξοδο (ύψος φιλοδωρήματος) επιλέγουμε μια αντίστοιχη διαμέριση με τριγωνοειδή ασαφή σύνολα και ορίζουμε τους αντίστοιχους όρους.

Με βάση τους αυτούς τους γλωσσικούς όρους, μπορούμε τώρα να ενσωματώσουμε «γνώση» στο σύστημά μας με τη μορφή αιτιακών κανόνων. Καταλήγουμε στους τρεις παρακάτω κανόνες:

ΕΑΝ η Ποιότητα Εξυπηρέτησης είναι **ΧΑΜΗΛΗ** ή η Ποιότητα Φαγητού είναι **ΧΑΜΗΛΗ** ΤΟΤΕ το Ύψος Φιλοδωρήματος είναι **ΧΑΜΗΛΟ**

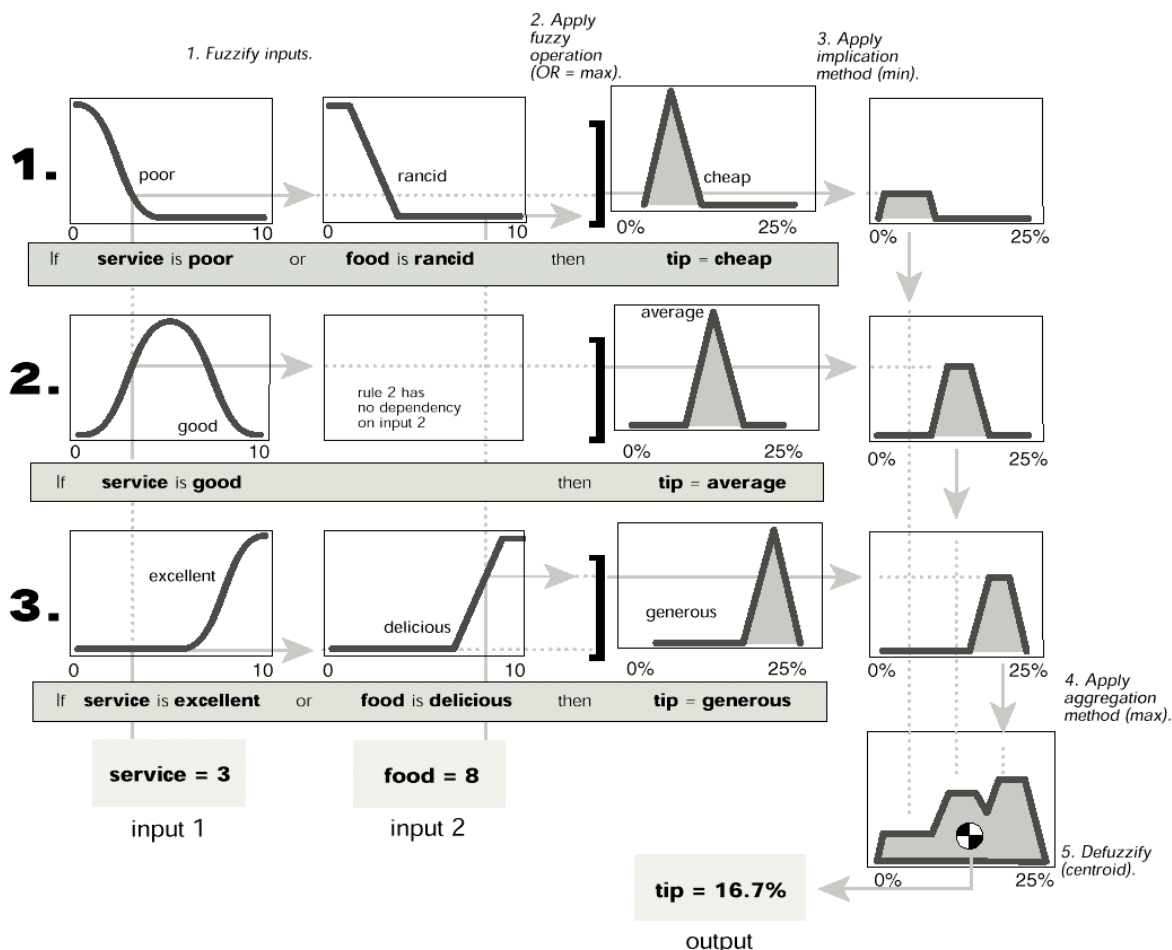
ΕΑΝ η Ποιότητα Εξυπηρέτησης είναι **ΜΕΤΡΙΑ** ΤΟΤΕ το Ύψος Φιλοδωρήματος είναι **ΜΕΤΡΙΟ**

ΕΑΝ η Ποιότητα Εξυπηρέτησης είναι **ΥΨΗΛΗ** ή η Ποιότητα Φαγητού είναι **ΥΨΗΛΗ** ΤΟΤΕ το Ύψος Φιλοδωρήματος είναι **ΥΨΗΛΟ**

Παρατηρούμε ότι στον δεύτερο κανόνα δεν λαμβάνουμε υπόψη την ποιότητα του φαγητού: οποιαδήποτε και αν είναι η ποιότητα του φαγητού εάν η ποιότητα εξυπηρέτησης είναι μέτρια επιθυμούμε να αφήσουμε μέτριο φιλοδώρημα. Επίσης σημειώνουμε ότι οι δύο προτάσεις στο πρώτο μέλος των κανόνων 1 και 3 συνδέονται με λογικό-ή.

Οι τρεις αυτοί κανόνες παρουσιάζονται πιο περιγραφικά στο Σχήμα 13.

¹ Το παράδειγμα αυτό είναι προσαρμοσμένο από το εγχειρίδιο Fuzzy Logic Toolbox που συνοδεύει το λογισμικό MATLAB της εταιρίας MATHWORKS.



Σχήμα 13. Οι κανόνες που χρησιμοποιούνται από το σύστημα του παραδείγματος

Στο σχήμα αυτό επιδεικνύεται και η λειτουργία του συστήματος. Ας θεωρήσουμε ότι έχουμε πλέον σχεδιάσει το σύστημα και το χρησιμοποιούμε για να προσδιορίσουμε το ύψος του φιλοδώρηματος σε μια περίπτωση όπου έχουμε βαθμολογήσει την ποιότητα της εξυπηρέτησης με 3 και την ποιότητα του φαγητού με 8.

Η τιμή 3 της εξυπηρέτησης, τέμνει τα ασαφή σύνολα της πρώτης στήλης (που σχετίζονται με την εξυπηρέτηση) σε κάποια σημεία, π.χ. χονδρικά στα 0.2, 0.6 και 0 από τον πρώτο κανόνα προς τον τρίτο.

Ομοίως, η τιμή 8 του φαγητού, τέμνει τα ασαφή σύνολα της δεύτερης στήλης (που σχετίζονται με το φαγητό) σε κάποια σημεία, π.χ. χονδρικά στα 0, - και 0.7 από τον πρώτο κανόνα προς τον τρίτο.

Λόγω του λογικού-ή μεταξύ των υποθέσεων, για κάθε κανόνα παίρνουμε τα μέγιστα των σημείων τομής. (Για τον δεύτερο κανόνα έχουμε μόνο ένα σημείο τομής.) Έτσι, έχουμε τα σημεία 0.2, 0.6 και 0.7.

Οι τιμές αυτές, καθορίζουν τα σημεία στα οποία θα τέμνουμε τα ασαφή σύνολα της μεταβλητής εξόδου (φιλοδώρημα) κάθε κανόνα.

Το αποτέλεσμα της τομής είναι τρεις ασαφή σύνολα (ένα για κάθε κανόνα). Ουσιαστικά κάθε κανόνας αποφάνθηκε ξεχωριστά για την περίπτωση και το συμπέρασμά του εμφανίζεται στο ασαφές σύνολο εξόδου του (τελευταία στήλη). Παρατηρούμε ότι οι απαντήσεις των κανόνων δεν έχουν όλες την ίδια «βαρύτητα». Ο πρώτος κανόνας έδωσε «μικρό» ασαφές σύνολο εξόδου. Γενικά, όσο πιο σχετικός είναι ένας κανόνας με τη συγκεκριμένη είσοδο τόσο πιο πολύ τείνει να επηρεάσει την έξοδο εξάγοντας πιο «ψηλά» ασαφή σύνολα.

Στο συγκεκριμένο πρόβλημα, ο τρίτος κανόνας ήταν ο πλέον σχετικός και έτσι πυροδοτήθηκε σε επίπεδο 0.7 έναντι του 0.2 και 0.6 των άλλων.

Στο τελευταίο βήμα, οι μερικές απαντήσεις κάθε κανόνα (ασαφή σύνολα) πρέπει να συνδυαστούν σε ένα μοναδικό ασαφές σύνολο. Η υπέρθεση των μερικών απαντήσεων οδηγεί στο ασαφές σύνολο στην κάτω δεξιά θέση του σχήματος 13.

Η τελική απόφαση για το ύψος του φιλοδωρήματος λαμβάνεται μετά από την αποασαφοποίηση του αυτού του ασαφούς συνόλου, χρησιμοποιώντας, για παράδειγμα, τη μέθοδο του κέντρου βάρους. Έτσι, για το ασαφές σύνολο εξόδου του παραδείγματος, η αποασαφοποίηση δίνει την τιμή 16.7%.

Νευρωνικά Δίκτυα

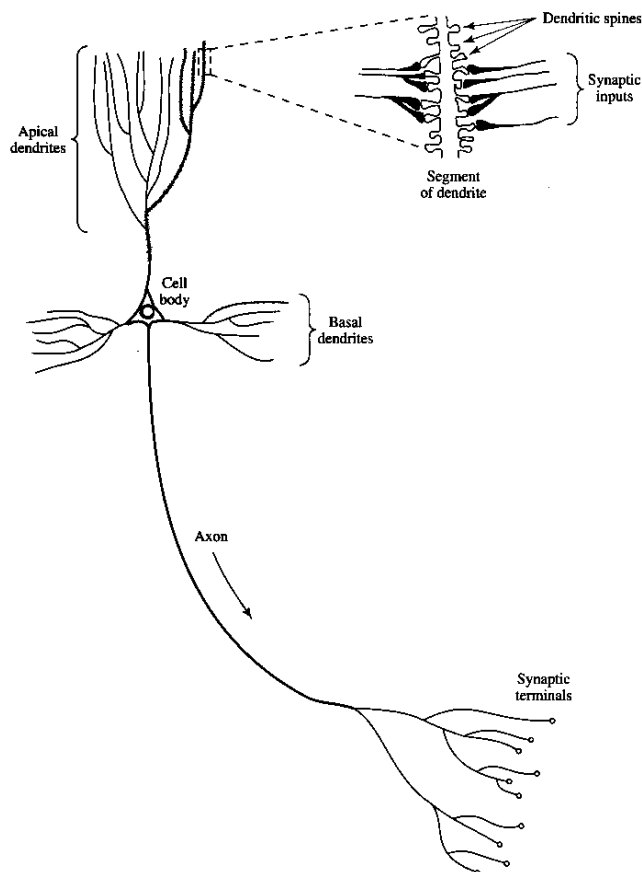
Εισαγωγή

Τα τεχνητά νευρωνικά δίκτυα υιοθετούν το δομικό πρότυπο του ανθρώπινου εγκεφάλου. Η ανάπτυξη της θεωρίας των νευρωνικών δικτύων επηρεάστηκε σημαντικά από την χρονική εξέλιξη των συμβάντων στον τομέα της Βιολογίας και της Φυσιολογίας.

Η πρώτη συστηματική κατανόηση του ανθρώπινου εγκεφάλου αποδίδεται στον Ramón y Cajal ο οποίος αναγνώρισε και περιέγραψε τους νευρώνες ως βασικά δομικά στοιχεία του εγκεφάλου το 1911. Από τότε, η έρευνα στα βιολογικά νευρωνικά δίκτυα έχει πραγματοποιήσει σημαντικές ανακαλύψεις αν και αποτελεί ακόμη ένα ανοικτό ερευνητικό πεδίο.

Με τη γέννηση του ανθρώπου, ο εγκέφαλος παρουσιάζεται **υψηλά δομημένος με έμφυτη ικανότητα διαμόρφωσης κανόνων** μέσω της «εμπειρίας». Την πρώτη διετία της ζωής, η ανάπτυξη του εγκεφάλου πραγματοποιείται με δραματικά υψηλούς ρυθμούς: σχηματίζονται περίπου 1 εκατομμύριο συνάψεις κάθε δευτερόλεπτο! Στην συνέχεια αν και η ανάπτυξη αυτή συνεχίζεται, εμφανίζει κάμψη στους ρυθμούς της.

Η Δομή του Εγκεφάλου



Σχήμα 14. Η δομή του νευρώνα

Νευρώνας

Ο **νευρώνας** (ή νευρικό κύτταρο) αποτελεί το βασικό δομικό συστατικό του εγκεφάλου.

Χονδρικά, ένας νευρώνας απαρτίζεται (i) από ένα αισθητήριο τμήμα όπου συλλέγονται οι είσοδοι από προηγούμενους νευρώνες, (ii) ένα τμήμα (ηλεκτροχημικής) «επεξεργασίας» και (iii) ένα τμήμα μέσω του οποίου διαδίδεται η έξοδος.

Σύναψη

Η **σύναψη** (ή σύνδεσμος) αποτελεί στοιχειώδη δομική και λειτουργική μονάδα που μεσολαβεί στην αλληλεπίδραση μεταξύ νευρώνων. Με τη βοήθεια μιας σύναψης, ένας νευρώνας μπορεί να δρα προτρεπτικά ή αποτρεπτικά στο νευρώνα-αποδέκτη του.

Δενδρίτης

Οι **δενδρίτες** (dendrites) απαρτίζουν την αισθητήρια ζώνη του νευρώνα και συλλέγουν τις εισόδους του από άλλους νευρώνες.

Άξονας

Ο **άξονας** (axon) αποτελεί τη γραμμή μετάδοσης ενός νευρώνα και διανέμει την έξοδό του σε άλλους νευρώνες.

Βιολογικά Νευρωνικά Δίκτυα Έναντι Υπολογιστικών Συστημάτων

Συγκρινόμενα με τα ψηφιακά επιτεύγματα της σύγχρονης τεχνολογίας, τα βιολογικά νευρωνικά δίκτυα εμφανίζονται σχετικά αργά. Το γεγονός αυτό, όμως, αντισταθμίζεται από τη μαζικότητα που

εμφανίζουν: υπολογίζεται ότι ο ανθρώπινος εγκέφαλος αποτελείται από 10 δισεκατομμύρια νευρώνες (neurons) και 60 τρισεκατομμύρια συνάψεις (synapses).

Ο παρακάτω πίνακας περιέχει συγκριτικές τιμές για κάποια μεγέθη:

	Βιολογικά ΝΔ	Η/Υ
Χρονική τάξη συμβάντων	10^{-3}sec	10^{-9}sec
Ενεργητική ικανότητα	10^{-16}J/op/sec	10^{-6}J/op/sec
Λειτουργίες αντιληπτικής αναγνώρισης (π.χ. αναγνώριση γνώριμων προσώπων σε άγνωστο περιβάλλον)	100-200 msec	~μέρες (ενδεικτικά)

Πίνακας. Σύγκριση βιολογικών νευρωνικών δικτύων με σύγχρονους υπολογιστές

Τεχνητά Νευρωνικά Δίκτυα

Τεχνητά νευρωνικά
δίκτυα

Ένα **Τεχνητό Νευρωνικό Δίκτυο** (ΤΝΔ — Artificial Neural Network, ANN) είναι ένα **κατανεμημένο** σύστημα επεξεργασίας, **μαζικής παραλληλίας**, με ικανότητα αποθήκευσης και ανάκλησης γνώσης προερχόμενης από εμπειρία. Το σύστημα αυτό μοιάζει στον ανθρώπινο εγκέφαλο στα εξής:

- Αποκτά γνώση μέσω μιας διαδικασίας **μάθησης**
- Αποθηκεύει τη γνώση στα **βάρη** των συνδέσεων μεταξύ νευρώνων.

Ιστορικές αναφορές

Στην ιστορία των τεχνητών νευρωνικών δικτύων μπορούμε να αναγνωρίσουμε κάποιες σημαντικές ημερομηνίες που καθόρισαν τις εξελίξεις στο χώρο αυτό:

1943

Η εργασία των McCulloch και Pitts όπου παρουσιάζεται και μελετάται το μοντέλο ενός τεχνητού νευρώνα. Η περιοχή ελκύει το πρώτο έντονο ερευνητικό ενδιαφέρον και πραγματοποιούνται τα πρώτα βασικά βήματα.

1969

Η δημοσίευση από τους Minsky και Papert είναι ένα ισχυρό πλήγμα για τα τεχνητά νευρωνικά δίκτυα αφού «αναδεικνύει» καίριες αδυναμίες τους. Η μαζική έρευνα περιορίζεται και τα νευρωνικά δίκτυα περνούν για αρκετό καιρό στο παρασκήνιο των εξελίξεων.

1986

Αναζωογονείται το ενδιαφέρον στην περιοχή λόγω των ιδιαίτερα αισιόδοξων αποτελεσμάτων της εργασίας των Rumelhart, Hinton και Williams που παρουσίασαν τον περίφημο αλγόριθμο μάθησης που τιτλοφορείται Ανάστροφη Διάδοση Σφάλματος (back-propagation of error).

Ιδιότητες Τεχνητών Νευρωνικών Δικτύων

Τα τεχνητά νευρωνικά δίκτυα παρουσιάζουν ένα σύνολο ελκυστικών (από σχεδιαστική άποψη) ιδιοτήτων που τα καθιστούν ιδιαίτερα χρήσιμα για ορισμένες κατηγορίες προβλημάτων:

• Μη-γραμμικότητα

Όπως θα φανεί και στις επόμενες παραγράφους, τα ΤΝΔ είναι ισχυρά μη γραμμικά συστήματα. Το χαρακτηριστικό αυτό, μπορεί μεν να δυσχεραίνει συχνά την μελέτη και τον χειρισμό τους, αλλά τα καθιστά ικανά για την αντιμετώπιση ιδιαίτερα πολύπλοκων προβλημάτων, δεδομένου ότι τα περισσότερα ρεαλιστικά προβλήματα είναι, άλλωστε, μη γραμμικά.

• Απεικόνιση Εισόδων σε Εξόδους

Τα ΤΝΔ ουσιαστικά απεικονίζουν εισόδους σε εξόδους, αποτελούν δηλαδή τυπική περίπτωση συστημάτων όπως αυτά παρουσιάστηκαν στις προηγούμενες παραγράφους. Το σημαντικό με τα ΤΝΔ όμως, είναι ότι αποτελούν αποδεδειγμένα **καθολικούς εκτιμητές άνευ μοντέλου**.

Ο όρος «εκτιμητές άνευ μοντέλου» αναφέρεται σε συστήματα που δεν απαιτούν (ούτε κατά τη σχεδίαση ούτε κατά τη λειτουργία τους) την διάρθρωση μοντέλων για το εκάστοτε πρόβλημα που καλούνται να αντιμετωπίσουν. Ο όρος «καθολικός» αναφέρεται στο γεγονός ότι τα ΤΝΔ είναι ικανά να επιλύσουν **οποιοδήποτε** πρόβλημα, δηλαδή να πραγματοποιήσουν οποιαδήποτε επιθυμητή απεικόνιση από το χώρο των εισόδων στο χώρο των εξόδων. Ένα σημείο που πρέπει να διευκρινίσουμε σχετικά με την «καθολικότητα», είναι ότι αυτή εξασφαλίζει ουσιαστικά ότι **για κάθε επιθυμητή απεικόνιση υπάρχει κάποιο τεχνητό νευρωνικό δίκτυο που να την πραγματοποιεί με αυθαίρετη ακρίβεια** αλλά δεν μας υποδεικνύει **ποιο είναι αυτό ή πώς θα το σχεδιάσουμε**.

- **Προσαρμοστικότητα**

Τα ΤΝΔ έχουν την δυνατότητα να προσαρμόζονται στο περιβάλλον τους. Για παράδειγμα, αφού σχεδιαστεί ένα ΤΝΔ για την αντιμετώπιση ενός συγκεκριμένου προβλήματος σε ένα καθορισμένο περιβάλλον, το δίκτυο αυτό μπορεί (με κατάλληλες μεθόδους) να μεταβάλλει κατάλληλα τις παραμέτρους του ώστε να αντεπεξέλθει σε πιθανές μεταβολές του περιβάλλοντος προσαρμόζοντας τη συμπεριφορά του σε αυτό. Η προσαρμογή στο περιβάλλον αποτελεί και αυτή μια μορφή μάθησης με τη διαφορά ότι αποτελεί μια διαρκή διεργασία που λαμβάνει χώρα κατά τη φάση λειτουργίας του συστήματος σε αντίθεση με την μάθηση που πραγματοποιείται μία μοναδική φορά κατά τη φάση σχεδίασης.

Η προσαρμοστικότητα συνεπάγεται αυτο-μεταβλητότητα. Η μεταβλητότητα, όμως, αυτή μπορεί να συντελέσει αρνητικά στην ευστάθεια του συστήματος. Ένα σύστημα που προβαίνει σε συνεχείς μεταβολές της συμπεριφοράς του παρασυρόμενο από τις όποιες μεταβολές του περιβάλλοντός του (όπως το σύστημα τις αντιλαμβάνεται) χωρίς τις κατάλληλες δικλίδες ασφαλείας μπορεί να οδηγηθεί σε μια ασταθή (και σίγουρα μη επιθυμητή) συμπεριφορά.

Συνεπώς, η προσαρμογή είναι ένα χαρακτηριστικό που θα πρέπει να καθορίζεται με προσοχή. Μικρή προσαρμοστικότητα οδηγεί σε ευσταθή συστήματα ενώ μεγάλη σε εύπλαστα. Το **δίλημμα ευστάθειας / ευπλαστίας** βρίσκει, συνήθως, τη βέλτιστη απάντησή του κάπου στο ενδιάμεσο.

- **Ανοχή σε σφάλματα**

Η γνώση αποθηκεύεται στα ΤΝΔ **κατανεμημένα** στα βάρη των συνδέσεων μεταξύ των νευρώνων. Έτσι, το βάρος μιας σύνδεσης δεν αντιστοιχεί σε ένα αυτοτελές ανεξάρτητο «κομμάτι» γνώσης αλλά καθορίζεται συλλογικά από όλη τη γνώση που το ΤΝΔ καλείται να αποθηκεύσει και να ανακαλεί. Ως αποτέλεσμα, δεν είναι καθόλου προφανής η «αποκρυπτογράφηση» της γνώσης που βρίσκεται αποθηκευμένη σε ένα ΤΝΔ απλά κοιτώντας τα βάρη των συνδέσεών του.

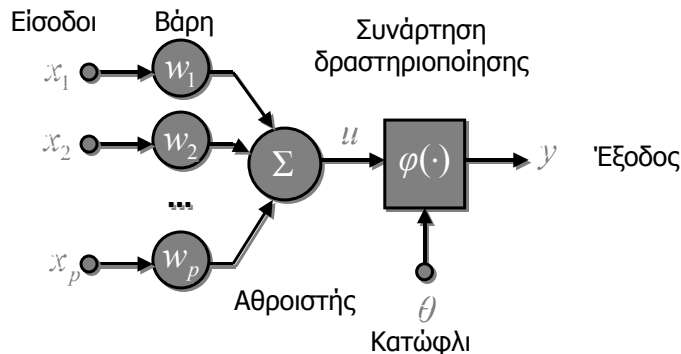
Το γεγονός όμως αυτό, προικίζει τα ΤΝΔ με μια ιδιαίτερη ικανότητα ανοχής σε σφάλματα. Αν, για παράδειγμα, υποθέσουμε ότι σε ένα ΤΝΔ ένα μικρό σύνολο από συνδέσεις για κάποιο λόγο καταστρέφονται ή δυσλειτουργούν, το γεγονός αυτό ελάχιστα θα επηρεάσει τη συνολική συμπεριφορά του δικτύου. Καθώς ο αριθμός των «ελαττωματικών» συνδέσεων μεγαλώνει σταδιακά θα αποκλιμακώνεται και η συνολική απόδοση του συστήματος. Το χαρακτηριστικό αυτό δίνει στα ΤΝΔ μια σημαντική ευρωστία συμπεριφοράς.

- **Υλοποίηση σε VLSI**

Ένα ΤΝΔ αποτελείται από έναν αριθμό αλληλοσυνδεδεμένων νευρώνων, καθένας από τους οποίους επιτελεί μια ιδιαίτερα απλή λειτουργία. Παρόλα αυτά το δίκτυο συνολικά μπορεί να αντιμετωπίσει ιδιαίτερα πολύπλοκα προβλήματα. Τα ΤΝΔ αποτελούν, δηλαδή, ένα παράδειγμα συστήματος που τα ενώ τα δομικά υλικά του είναι α-νόητα (δεν παρουσιάζουν νοημοσύνη) συνολικά το σύστημα εμφανίζεται «νοήμον». Συνεπώς, η ανακάλυψή τους συλλογική ικανότητα του δικτύου είναι σημαντικά ανώτερη από μια απλή άθροιση των ικανοτήτων των επιμέρους στοιχείων του. Δηλαδή, παρατηρούμε «νοημοσύνη» να προκύπτει από την κατάλληλη δόμηση ενός αριθμού α-νόητων στοιχείων [MINSKY, 1986]. Η δομή αυτή των νευρωνικών δικτύων προσφέρεται άμεσα για υλοποίηση σε κυκλώματα υψηλής ολοκλήρωσης (very large scale integration).

Το Μοντέλο του Νευρώνα

Στη βάση του βιολογικού νευρώνα, όπως αυτός παρουσιάστηκε στην αντίστοιχη παράγραφο, στηρίχθηκε το μοντέλο του τεχνητού νευρώνα που χρησιμοποιείται για τη σχεδίαση ΤΝΔ. Το μοντέλο αυτό παρουσιάζεται στο παρακάτω σχήμα (Σχήμα 15).



Σχήμα 15. Το μοντέλο του τεχνητού νευρώνα

Η είσοδος στο μοντέλο του τεχνητού νευρώνα αποτελείται από ένα σύνολο συνάψεων με βάρη. Κάθε είσοδος πολλαπλασιάζεται με το αντίστοιχο βάρος και αφού αθροιστούν τροφοδοτούν την συνάρτηση δραστηριοποίησης του νευρώνα. Η έξοδος της συνάρτησης αυτής, που είναι και η συνολική έξοδος του νευρώνα, θα αποτελέσει είσοδο για τους νευρώνες που έπονται.

Ένα σημείο που πρέπει να τονιστεί είναι η παράμετρος «κατωφλίου». Το κατώφλι, θ , δεν αποτελεί πραγματική είσοδο του νευρώνα αλλά μια εσωτερική παράμετρό του που επηρεάζει σημαντικά την έξοδο του αφού καθορίζει τη συμπεριφορά της συνάρτησης δραστηριοποίησης. Ο ακριβής του ρόλος θα αναδειχθεί στις ακόλουθες παραγράφους.

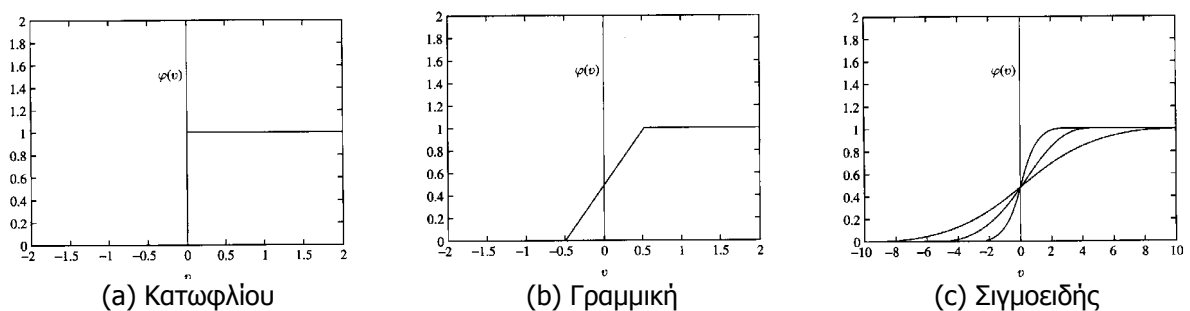
Μαθηματικό μοντέλο νευρώνα

Καταstrώνοντας το μαθηματικό μοντέλο του τεχνητού νευρώνα, καταλήγουμε στην εξίσωση:

$$y = \varphi \left(\sum_{j=1}^p w_j x_j - \theta \right)$$

Συνάρτηση δραστηριοποίησης

Ο ρόλος της συνάρτησης δραστηριοποίησης είναι να εισάγει μη γραμμικότητα στη συμπεριφορά του νευρώνα. Έτσι, οι πιο συνηθισμένες μορφές που χρησιμοποιούνται είναι αυτές που απεικονίζονται στο Σχήμα 16.



Σχήμα 16. Συνήθεις συναρτήσεις δραστηριοποίησης

Αξίζει να σημειώσουμε ότι ενώ η συνάρτηση δραστηριοποίησης είναι, στη γενική περίπτωση, μια μη γραμμική συνάρτηση δεν αποκλείονται και γραμμικές συναρτήσεις. Νευρωνικά δίκτυα που χρησιμοποιούν τέτοιες συναρτήσεις είναι μεν γραμμικά αλλά χρησιμοποιούνται είτε για την επίλυση γραμμικών προβλημάτων είτε για την προσεγγιστική γραμμική επίλυση μη γραμμικών προβλημάτων.

Ακόμη και με την χρήση γραμμικών στοιχείων, η νευρωνική προσέγγιση στην επίλυση προβλημάτων επωφελείται από τις αρχιτεκτονικές νευρωνικών δικτύων και τους αλγόριθμους νευρωνικής μάθησης.

Αρχιτεκτονικές Νευρωνικών Δικτύων

Με βάση το μοντέλο του τεχνητού νευρώνα, μπορούμε να οργανώσουμε δίκτυα αποτελούμενα από έναν αριθμό τέτοιων νευρώνων. Τα τεχνητά νευρωνικά δίκτυα μπορούν να διαχωριστούν σε δύο ευρείες κατηγορίες: τα δίκτυα προσοτροφοδότησης και τα δίκτυα ανατροφοδότησης.

Δίκτυα
προσοτροφοδότησης

Τα δίκτυα **προσοτροφοδότησης** (feedforward) αποτελούνται από έναν αριθμό νευρώνων οργανωμένων κατά τέτοιο τρόπο ώστε η ροή των δεδομένων να έχει μία μοναδική κατεύθυνση: από τις εισόδους προς τις εξόδους του δικτύου. Σε καμία περίπτωση, δηλαδή, ένας νευρώνας, έστω β , που έλαβε είσοδο από κάποιον άλλον νευρώνα, έστω α , δεν θα επηρεάσει με την έξοδό του (άμεσα ή έμμεσα) τον νευρώνα α .

Δίκτυα
ανατροφοδότησης

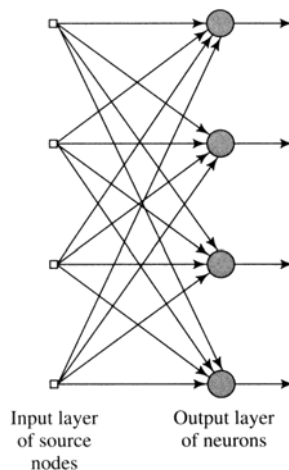
Στα δίκτυα **ανατροφοδότησης** (feedback) ο προηγούμενος περιορισμός δεν ισχύει. Έτσι, οποιοδήποτε νευρώνας του δικτύου μπορεί να συνδέονται με αυθαίρετο τρόπο. Λόγω αυτού, τα δίκτυα ανατροφοδότησης είναι αρκετά πιο πολύπλοκα σε σχέση με τα δίκτυα προσοτροφοδότησης.

Στα περισσότερα νευρωνικά δίκτυα, οι νευρώνες είναι οργανωμένοι σε ευδιάκριτες ομάδες, τα λεγόμενα στρώματα. Νευρώνες του ίδιου στρώματος συνήθως δεν συνδέονται μεταξύ τους και συνεπώς δεν αλληλοεπηρεάζονται. Ένα δίκτυο μπορεί να έχει αυθαίρετο αριθμό στρωμάτων, αλλά τουλάχιστον ένα.

Με βάση τη ροή των δεδομένων (προσοτροφοδότηση / ανατροφοδότηση) και το αριθμό των στρωμάτων, μπορούμε να διακρίνουμε τις παρακάτω συνήθεις αρχιτεκτονικές νευρωνικών δικτύων.

Δίκτυα Προσοτροφοδότησης Ενός Στρώματος

Αποτελούν την απλούστερη μορφή νευρωνικών δικτύων. Οι εξωτερικές εισοδοί τροφοδοτούνται σε όλους τους νευρώνες και η έξοδος κάθε νευρώνα συνεισφέρει στη συνολική έξοδο του δικτύου. Η συμπεριφορά των δικτύων αυτών είναι αρκετά απλή και γι αυτό αποτελούν ένα από τα πλέον χρησιμοποιούμενα είδη δικτύων. Φυσικά, η «ισχύς» των δικτύων αυτών είναι σχετικά περιορισμένη αφού δεν έχουν την ικανότητα να πραγματοποιήσουν αυθαίρετα πολύπλοκες απεικονίσεις από τις εισόδους στις εξόδους.

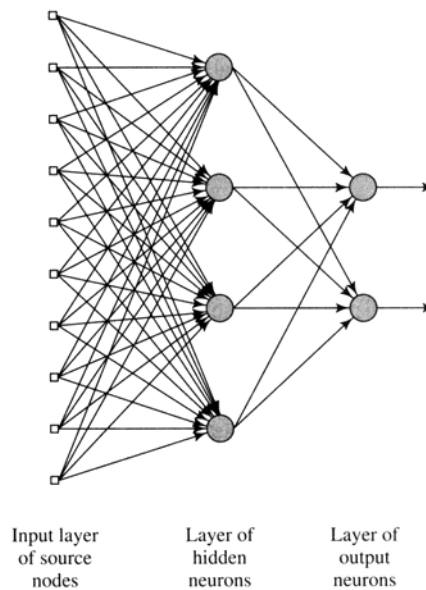


Σχήμα 17. Δίκτυο προσοτροφοδότησης ενός στρώματος

Μια ειδική κατηγορία νευρωνικών δικτύων προσοτροφοδότησης ενός ή περισσότερων στρωμάτων που αποκαλούνται δίκτυα Perceptron θα αναλυθούν σε ξεχωριστή παράγραφο λόγω της ιδιαίτερης σημασίας και της ευρείας χρήσης τους.

Δίκτυα Προσοτροφοδότησης Πολλών Στρωμάτων

Αποτελούν την προφανή επέκταση των δικτύων ενός στρώματος. Τα πολλά στρώματα δίνουν αυξημένες ικανότητες διαμόρφωσης πολύπλοκων απεικονίσεων των εισόδων στις εξόδους.

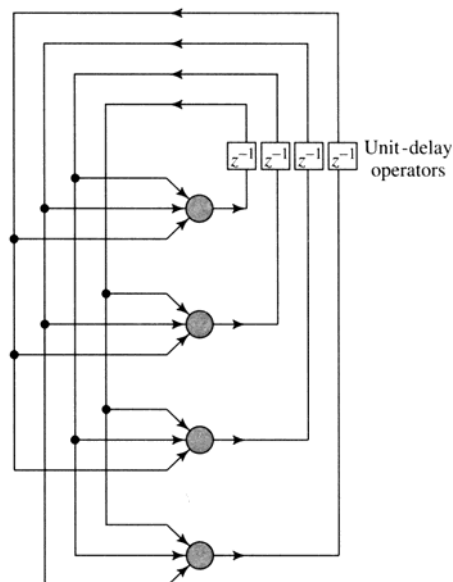


Σχήμα 18. Δίκτυο προστροφοδότησης δύο στρωμάτων

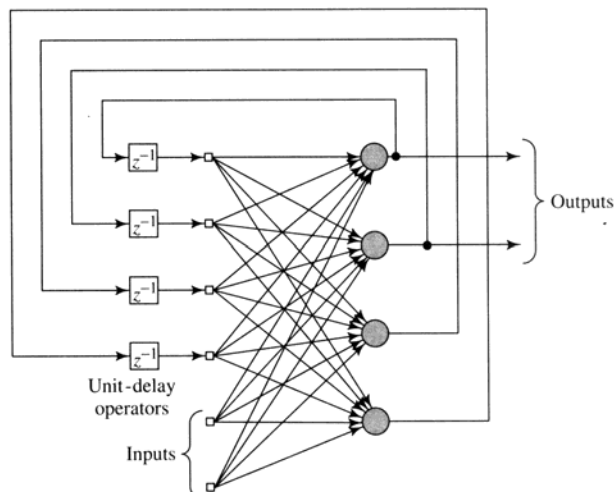
Δίκτυα Ανατροφοδότησης

Τα δίκτυα αυτά έχουν αυθαίρετη δομή και παρουσιάζουν δυναμική (χρονικά μεταβαλλόμενη) συμπεριφορά.

Διακρίνονται σε δίκτυα με ή χωρίς κρυμμένους νευρώνες. Κρυμμένοι θεωρούνται οι νευρώνες των οποίων η έξοδος δεν αποτελεί μέρος της συνολικής εξόδου του δικτύου αλλά χρησιμοποιείται μόνο ως είσοδος για άλλους νευρώνες.



Σχήμα 19. Δίκτυο με ανατροφοδότηση χωρίς κρυμμένους νευρώνες



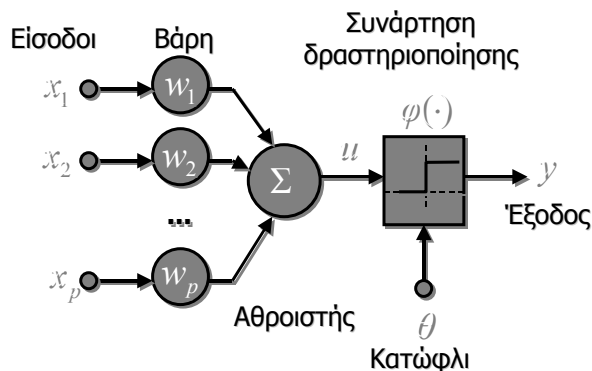
Σχήμα 20. Δίκτυο ανατροφοδότησης με κρυμμένους νευρώνες

Συνήθεις Αρχιτεκτονικές Νευρωνικών Δικτύων

Δίκτυα Perceptron

Τα δίκτυα Perceptron είναι τυπικά δίκτυα προστροφοδότησης αποτελούμενα από ένα ή περισσότερα στρώματα. Το ειδοποιό χαρακτηριστικό τους είναι ότι όλοι οι νευρώνες τους χρησιμοποιούν την συνάρτηση δραστηριοποίησης κατωφλίου (Σχήμα 16α). Συνεπώς, η έξοδος κάθε νευρώνα μπορεί να λάβει δύο και μόνο τιμές: 0 ή 1. Το γεγονός αυτό καθιστά τα δίκτυα Perceptron ιδανικά για συγκεκριμένες κατηγορίες εφαρμογών όπως η κατηγοριοποίηση (classification).

Στο παρακάτω σχήμα απεικονίζεται ένας τυπικός νευρώνας που χρησιμοποιείται στα δίκτυα τύπου Perceptron.



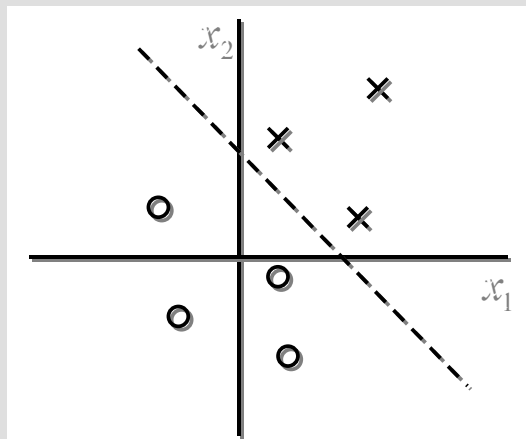
Σχήμα 21. Ένας τυπικός νευρώνας τύπου Perceptron

Παράδειγμα: Ταξινόμηση με Έναν Νευρώνα

Ας θεωρήσουμε το εξής πρόβλημα:

*Επιθυμούμε να σχεδιάσουμε ένα σύστημα που να ταξινομεί τα σημεία ενός διδιάστατου χώρου (επιπέδου) x_1 - x_2 σε δύο κατηγορίες: την κατηγορία **A** και την κατηγορία **B**. Στο παρακάτω σχήμα, φαίνονται τα δείγματα του επιπέδου που επιθυμούμε να διαχωρίσουμε. Με «**x**» σημειώνονται τα σημεία της κατηγορίας **A** και με «**ο**» τα σημεία της κατηγορίας **B**.*

Για το πρόβλημα αυτό επιλέγουμε ένα νευρωνικό δίκτυο τύπου Perceptron με έναν μοναδικό νευρώνα. Ο νευρώνας αυτός θα δέχεται δύο εισόδους (τις τιμές x_1 και x_2 ενός σημείου) και θα έχει μία έξοδο που θα λαμβάνει την τιμή 0 όταν το σημείο ανήκει στην κατηγορία **A** ή 1 όταν το σημείο ανήκει στην κατηγορία **B**.



Σχήμα 22. Τα προς ταξινόμηση σημεία του επιπέδου

Ανακαλώντας την γενική εξίσωση για τη συνάρτηση μεταφοράς ενός νευρώνα, έχουμε:

$$y = \varphi(u), \text{ όπου } u = \sum_{j=1}^p (w_j x_j - \theta)$$

Όμως, η συνάρτηση δραστηριοποίησης του νευρώνα είναι τύπου κατώφλιου και λαμβάνει μόνο την τιμή 0 (όταν το άθροισμα των εισόδων πολλαπλασιασμένων με τα αντίστοιχα βάρη είναι μικρότερο του θ) ή 1 (όταν είναι μεγαλύτερο). Δηλαδή:

$$\varphi(u) = \begin{cases} 1, & \text{εάν } u \geq 0 \\ 0, & \text{εάν } u < 0 \end{cases}$$

Το όριο μεταξύ των δύο περιοχών είναι το σύνολο σημείων όπου $u=0$, δηλαδή ο τόπος:

$$\sum_{j=1}^2 (w_j x_j - \theta) = w_1 x_1 + w_2 x_2 - \theta = 0$$

Η τελευταία εξίσωση είναι η κλασσική εξίσωση μιας ευθείας στο χώρο, γεγονός που υποδηλώνει ότι το όριο μεταξύ των δύο διαφορετικών περιοχών απόφασης που μπορεί να παράγει ο νευρώνας είναι μια ευθεία.

Μπορούμε άμεσα να υπολογίσουμε τις απαραίτητες τιμές για τα βάρη και το κατώφλι του νευρώνα εάν απλά σχεδιάσουμε μια ευθεία που να διαχωρίζει πλήρως τα σημεία «x» από τα σημεία «o» και υπολογίσουμε την κλίση και τα σημεία τομής της με τους άξονες x_1 και x_2 . Παρατηρούμε ότι η κλίση της ευθείας καθορίζεται από τα βάρη w_1 και w_2 ενώ η θέση της σε σχέση με την αρχή των αξόνων από το κατώφλι θ .

Οι πιθανές επιλογές που έχουμε για μια ευθεία που να διαχωρίζει πλήρως τα σημεία των δύο διαφορετικών κατηγοριών είναι άπειρες. Μια τέτοια ευθεία παρουσιάζεται διακεκομμένη στο Σχήμα 22.

Είναι φανερό ότι ένα δίκτυο Perceptron ενός μοναδικού νευρώνα μπορεί να επιλύσει μόνο προβλήματα ταξινόμησης συνόλων που είναι **γραμμικά διαχωρίσιμα**. Εάν δεν υπάρχει ευθεία που να διαχωρίζει τις δύο κατηγορίες σημείων τότε δεν υπάρχουν τιμές για τα βάρη και το κατώφλι που να μπορούν να οδηγήσουν το νευρωνικό στη διαμόρφωση μιας σωστής ευθείας απόφασης.

Στην περίπτωση που οι επιθυμητές κατηγορίες ήταν περισσότερες από δύο, θα απαιτούνταν περισσότεροι νευρώνες (ένα στρώμα νευρώνων). Αντίστοιχα, στην περίπτωση που ο χώρος του οποίου τα σημεία επιθυμούσαμε να διαχωρίσουμε ήταν μεγαλύτερης διάστασης από δύο (δεν ήταν, δηλαδή, επίπεδο), τότε θα είχαμε περισσότερες εισόδους σε κάθε νευρώνα του δικτύου. Τέλος, εάν οι κατηγορίες σημείων δεν ήταν γραμμικά διαχωρίσιμες, τότε περισσότερα του ενός στρώματα θα ήταν απαραίτητα.

Γραμμικά Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα αυτού του τύπου είναι καθ'όλα όμοια με τα νευρωνικά δίκτυα τύπου Perceptron με μόνη διαφορά την συνάρτηση δραστηριοποίησης που χρησιμοποιούν που είναι η γραμμική σε όλο το πεδίο ορισμού της. Δηλαδή, σε κάθε σημείο του πεδίου ορισμού:

$$\varphi(u) = u$$

Αν και η συμπεριφορά των δικτύων αυτού του τύπου δεν είναι μη-γραμμική, τα δίκτυα αυτά βρίσκουν εκτεταμένη εφαρμογή σε μεγάλο εύρος προβλημάτων. Δεν είναι σπάνιο δε, να χρησιμοποιηθούν σε περιπτώσεις γραμμικής προσεγγιστικής λύσης ενός μη γραμμικού προβλήματος.

Αλγόριθμοι Εκπαίδευσης Νευρωνικών Δικτύων

Μάθηση	Κεντρική θέση, τόσο στη θεωρία των νευρωνικών δικτύων όσο και στο ευρύτερο πλαίσιο της Υπολογιστικής Νοημοσύνης, κατέχει η έννοια της μάθησης. Η μάθηση συνίσταται στον υπολογισμό κατάλληλων τιμών για τις ελεύθερες παραμέτρους ενός μοντέλου ώστε το προκύπτων σύστημα να παρουσιάζει την επιθυμητή συμπεριφορά.
Νευρωνική μάθηση	Στο πλαίσιο των νευρωνικών δικτύων, ελεύθερες παραμέτρους αποτελούν τα βάρη των συνδέσεων και οι τιμές των κατώφλιών των νευρώνων. Συνεπώς ο σκοπός κάθε αλγορίθμου μάθησης είναι ο προσδιορισμός κατάλληλων τιμών για αυτά τα μεγέθη. Μπορούμε να διακρίνουμε δύο βασικές κατηγορίες αλγορίθμων νευρωνικής μάθησης: την μάθηση υπό επίβλεψη και την μάθηση χωρίς επίβλεψη.
Μάθηση υπό επίβλεψη	Η μάθηση υπό επίβλεψη (supervised learning) προϋποθέτει την ύπαρξη κάποιας μορφής «επιβλέποντος» ο οποίος γνωρίζει την «σωστή» για κάθε περίπτωση συμπεριφορά και διορθώνει το νευρωνικό δίκτυο όταν πραγματοποιεί λάθη. Στην μάθηση υπό επίβλεψη, η «σωστή» συμπεριφορά προσδιορίζεται, συνήθως, με τη βοήθεια ενός συνόλου από υποδειγματικά ζευγάρια εισόδων-εξόδων (παραδειγματικά πρότυπα). Τα ζευγάρια αυτό προδιαγράφουν την έξοδο που επιθυμούμε να εμφανίζει το σύστημα όταν διεγείρεται από την συγκεκριμένη είσοδο. Πριν το νευρωνικό ολοκληρώσει την φάση εκπαίδευσης, η πραγματική έξοδος που θα παράγει όταν διεγείρεται από τη συγκεκριμένη είσοδο διαφέρει από την επιθυμητή. Η διαφορά αυτή είναι το σφάλμα του νευρωνικού δικτύου και χρησιμοποιείται από τον αλγόριθμο εκπαίδευσης για να προσδιορίσει νέες τιμές για τα βάρη και τα κατώφλια. Οι αλγόριθμοι εκπαίδευσης μπορεί να είναι επαναληπτικοί ή μη. Στην πρώτη περίπτωση τα παραδειγματικά πρότυπα παρουσιάζονται πολλές φορές διαδοχικά στο νευρωνικό δίκτυο και κάθε φορά υπολογίζεται το σφάλμα που χρησιμοποιείται για την κατάλληλη μεταβολή των ελεύθερων παραμέτρων του δικτύου. Η διαδικασία αυτή τερματίζεται είτε όταν πραγματοποιηθεί συγκεκριμένος αριθμός επαναλήψεων είτε όταν το σφάλμα πέσει κάτω από κάποιο επιθυμητό επίπεδο. Πριν από την εκκίνηση του βρόχου εκπαίδευσης, οι ελεύθερες παράμετροι του δικτύου αρχικοποιούνται τυχαία λαμβάνοντας μικρές (συνήθως θετικές) τιμές.
Μάθηση χωρίς επίβλεψη	Η μάθηση χωρίς επίβλεψη (unsupervised learning) δεν απαιτεί την ύπαρξη οποιασδήποτε ρητής περιγραφής της επιθυμητής συμπεριφοράς όπως η μάθηση υπό επίβλεψη. Οι αντίστοιχοι αλγόριθμοι μάθησης είτε βασίζονται σε πληροφορία που αντλείται κατά τη διάρκεια της λειτουργίας του νευρωνικού, ή οδηγούν τους νευρώνες του δικτύου σε κάποια μορφή ανταγωνισμού / συνεργασίας ώστε καθένας να ωθείται να «ειδικευτεί» σε συγκεκριμένες περιοχές του προβλήματος που προσδιορίζονται δυναμικά. Γενικά, οι αλγόριθμοι εκπαίδευσης νευρωνικών δικτύων δεν είναι ανεξάρτητοι από την επιλεχθείσα αρχιτεκτονική του δικτύου. Δηλαδή δεν μπορούμε να συνδυάσουμε οποιονδήποτε αλγόριθμο μάθησης με νευρωνικά δίκτυα οποιασδήποτε αρχιτεκτονικής. Η επιλογή των δύο αυτών συντελεστών είναι εξαρτημένη. Άλλωστε, οι περισσότεροι αλγόριθμοι εκπαίδευσης αναπτύχθηκαν στο πλαίσιο κάποιας συγκεκριμένης αρχιτεκτονικής.

Αλγόριθμος Εκπαίδευσης για Δίκτυα Perceptron Ενός Στρώματος

Στο παράδειγμα της προηγούμενης παραγράφου, εστιάσαμε το ενδιαφέρον μας σε ένα απλούστατο δίκτυο τύπου Perceptron και υπολογίσαμε «με το χέρι» κατάλληλες τιμές για τα βάρη και το κατώφλι του μοναδικού νευρώνα. Είναι σαφές ότι για προβλήματα πιο αυξημένης δυσκολίας η μέθοδος αυτή έχει περιορισμένη χρησιμότητα.

Παρακάτω περιγράφεται ένας αλγόριθμος εκπαίδευσης για τον αυτόματο υπολογισμό των βαρών και κατωφλίων ενός νευρωνικού δικτύου τύπου Perceptron ενός στρώματος με αυθαίρετο αριθμό νευρώνων.

Λόγω της παρουσίας περισσότερων του ενός νευρώνων, αναπόφευκτα καταφεύγουμε στην υιοθέτηση ενός ακόμη δείκτη που υποδηλώνει τον νευρώνα στον οποίο αναφερόμαστε. Έτσι, ο δείκτης j αναφέρεται στον τρέχοντα νευρώνα και ο δείκτης i στην τρέχουσα είσοδο. Συνεπώς, η μεταβλητή w_{ij} αναφέρεται στη βάρος της σύνδεσης από την είσοδο i στον νευρώνα j και η μεταβλητή θ_j στο κατώφλι του νευρώνα j .

Θα θεωρήσουμε ότι διαθέτουμε ένα σύνολο από ζεύγη εισόδων / επιθυμητών εξόδων που θα συμβολίσουμε με $\mathbf{x}$ και t . Εξετάζοντας κάθε νευρώνα ξεχωριστά, το $\mathbf{x}$ θα είναι ένα διάνυσμα εισόδων (αφού έχουμε παραπάνω από μία εισόδους) και το t βαθμωτό αφού κάθε νευρώνας έχει μία μόνο έξοδο. Με y θα συμβολίσουμε την πραγματική έξοδο του νευρώνα.

Υπενθυμίζουμε ότι στα δίκτυα Perceptron οι εξοδοί των νευρώνων μπορούν να λάβουν μόνο δύο τιμές: 0 ή 1. Έτσι, τα y και t μπορεί να είναι μόνο 0 ή 1.

Βήμα 1.

Αρχικοποιούμε τα βάρη w_{ij} και τα κατώφλια θ_j με τυχαίες μικρές θετικές τιμές.

Βήμα 2.

Για κάθε ένα από τα παραδειγματικά πρότυπα εισόδων (έστω $\mathbf{x}$) / επιθυμητών εξόδων (έστω t) υπολογίζουμε την πραγματική έξοδο του νευρώνα, έστω y . Συγκρίνοντας το y με το t , έχουμε τις παρακάτω περιπτώσεις:

- **Περίπτωση 1.** $t = y$.

Τότε το σφάλμα για αυτόν τον νευρώνα θα είναι $e = t - y = 0$. Συνεπώς καμία αλλαγή δεν απαιτείται για αυτόν τον νευρώνα.

- **Περίπτωση 2.** $t = 1$ και $y = 0$.

Τότε το σφάλμα για αυτόν τον νευρώνα θα είναι $e = t - y = 1$. Μεταβάλλουμε τα βάρη των εισόδων του νευρώνα σύμφωνα με τον κανόνα:

$$\mathbf{w} = \mathbf{w} + \mathbf{x}$$

- **Περίπτωση 3.** $t = 0$ και $y = 1$.

Τότε το σφάλμα για αυτόν τον νευρώνα θα είναι $e = t - y = -1$. Μεταβάλλουμε τα βάρη των εισόδων του νευρώνα σύμφωνα με τον κανόνα:

$$\mathbf{w} = \mathbf{w} - \mathbf{x}$$

Οι δύο κανόνες των δύο τελευταίων περιπτώσεων μπορούν να γραφούν με μία ενοποιημένη μορφή ως:

$$\mathbf{w} = \mathbf{w} + e\mathbf{x}$$

Τελείως απλοϊκά, το σκεπτικό του κανόνα αυτού είναι ότι το όριο της περιοχής απόφασης του νευρώνα είναι τέτοιο ώστε το δείγμα $\mathbf{x}$ του χώρου βρέθηκε σε λάθος πλευρά της. Συνεπώς, η περιοχή απόφασης θα πρέπει να μετακινηθεί προς το μέρος του σημείου $\mathbf{x}$ ώστε να το φέρει στην άλλη πλευρά της.

Βήμα 3.

Επαναλαμβάνουμε το Βήμα 2 για κάθε νευρώνα.

Βήμα 4.

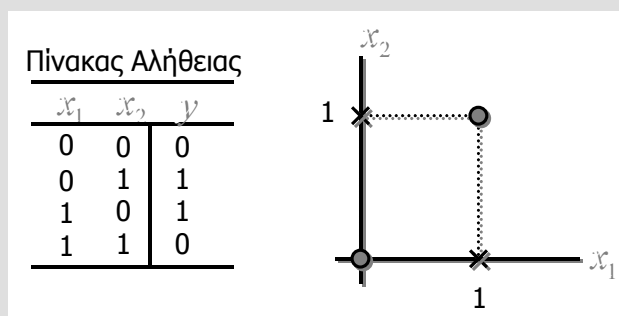
Επαναλαμβάνουμε από το Βήμα 2 για κάθε ζεύγος του συνόλου των παραδειγματικών προτύπων.

Ο αλγόριθμος για την εκπαίδευση νευρωνικών δικτύων τύπου Perceptron ενός στρώματος **συγκλίνει** αποδεδειγμένα εάν τα σημεία των διαφορετικών κατηγοριών είναι γραμμικά διαχωρίσιμα. Αυτό σημαίνει ότι εάν υπάρχει λύση στο πρόβλημα ο αλγόριθμος θα καταφέρει να υπολογίσει σε πεπερασμένο αριθμό επαναλήψεων τις παραμέτρους ενός νευρωνικού που θα το λύνει.

Παράδειγμα: Το Πρόβλημα της Αποκλειστικής Διάζευξης (XOR)

Ένα κλασσικό πρόβλημα στον χώρο των νευρωνικών δικτύων είναι το πρόβλημα της αποκλειστικής διάζευξης (exclusive or - XOR). Το δυαδικό αυτό πρόβλημα περιγράφεται από τον πίνακα αληθείας του Σχήματος 23 οι τιμές του οποίου απεικονίζονται στο διπλανό διάγραμμα.

Παρά την φαινομενική απλότητά του, το πρόβλημα αυτό δεν μπορεί να επιλυθεί από νευρωνικά δίκτυα τύπου Perceptron ενός στρώματος ακόμη και με περισσότερους του ενός νευρώνες. Ο απλούστατος λόγος για αυτό είναι ότι τα σημεία των δύο κατηγοριών δεν είναι γραμμικά διαχωρίσιμα, δηλαδή δεν υπάρχει ευθεία που να «αφήνει» όλα τα «**x**» σημεία (και μόνο αυτά) από τη μια μεριά της και όλα τα «**ο**» σημεία από την άλλη.



Σχήμα 23. Το πρόβλημα της αποκλειστικής διάζευξης (XOR)

Για την επίλυση αυτού του προβλήματος, μπορεί να χρησιμοποιηθεί νευρωνικό δίκτυο τύπου Perceptron αρκεί να περιέχει τουλάχιστον δύο στρώματα νευρώνων.

Αλγόριθμος Εκπαίδευσης για Γραμμικά Νευρωνικά Δίκτυα

Για την εκπαίδευση των γραμμικών νευρωνικών δικτύων χρησιμοποιείται ο **αλγόριθμος ελαχίστων μέσων τετραγώνων** (least mean squares, LMS). Ο αλγόριθμος αυτός επιδιώκει τον υπολογισμό βέλτιστων τιμών για τις ελεύθερες παραμέτρους ενός γραμμικού νευρωνικού δικτύου μέσω της ελαχιστοποίησης κάποιου κατάλληλα επιλεγμένου κριτηρίου κόστους (σφάλματος). Και ο αλγόριθμος αυτός απαιτεί την ύπαρξη ενός συνόλου παραδειγματικών προτύπων.

Ας υποθέσουμε ότι διαθέτουμε ένα σύνολο από Q παραδειγματικά πρότυπα. Για κάθε ένα από αυτά, έστω το k , υπολογίζουμε το σφάλμα με βάση την σχέση:

$$e_k = t_k - y_k$$

Σκοπός του αλγορίθμου είναι η ελαχιστοποίηση του μέσου τετραγώνου των σφαλμάτων σε όλα τα δεδομένα εκπαίδευσης συνολικά, την ελαχιστοποίηση δηλαδή της ποσότητας:

$$mse = \frac{1}{Q} \sum_{k=1}^Q e_k = \frac{1}{Q} \sum_{k=1}^Q (t_k - y_k)^2$$

Χωρίς να υπεισέλθουμε σε λεπτομέρειες, αναφέρουμε ότι από την επίλυση της παραπάνω εξίσωσης προκύπτουν βέλτιστες τιμές για τις ελεύθερες παραμέτρους του νευρωνικού δικτύου.

Παρατηρούμε μια σημαντική διαφορά από τον αλγόριθμο εκπαίδευσης που χρησιμοποιείται για τα νευρωνικά δίκτυα τύπου Perceptron. Εκεί, ο αλγόριθμος εκπαίδευσης ήταν επαναληπτικός και σε κάθε βήμα μεταβάλλαμε κατά μια ποσότητα τις ελεύθερες παραμέτρους του δικτύου. Εδώ, η τελευταία εξίσωση δίνει άμεσα τιμές για τις παραμέτρους αυτές χωρίς την ανάγκη κάποιας επαναληπτικής διαδικασίας. Πρόκειται λοιπόν περισσότερο για αλγόριθμο **σχεδίασης** παρά για αλγόριθμο **μάθησης**.

Μία επιπρόσθετη διαφορά των δύο αλγορίθμων είναι η δυνατότητα του δεύτερου να υπολογίσει **βέλτιστες τιμές** για τις παραμέτρους. Στο πρόβλημα της κατηγοριοποίησης δεδομένων από γραμμικά διαχωρίσιμες περιοχές, ήταν ορατό ότι οι πιθανές λύσεις ήταν άπειρες. Χονδρικά, μπορούμε να πούμε ότι η βέλτιστη ευθεία απόφασης (και συνεπώς οι βέλτιστες τιμές για τις παραμέτρους του νευρωνικού) είναι εκείνη που όχι μόνο κατηγοριοποιεί σωστά όλα τα δεδομένα αλλά και βρίσκεται όσο το δυνατόν πιο μακριά από αυτά. Αυτή η απόσταση «ασφαλείας» προφέρει σχετική ευρωστία στο νευρωνικό δίκτυο προφυλάσσοντάς το από θορυβώδη δεδομένα.

Κανόνας Μάθησης Widrow-Hoff

Ο κανόνας μάθησης των Widrow-Hoff προσφέρεται ως μια εναλλακτική επιλογή στον αλγόριθμο ελάχιστων μέσων τετραγώνων της προηγούμενης παραγράφου. Όντας επαναληπτικός, είναι κατάλληλος για περιπτώσεις όπου τα παραδειγματικά πρότυπα δεν είναι διαθέσιμα εκ των προτέρων και συνεπώς το μέσο τετραγωνικό σφάλμα δεν μπορεί να υπολογιστεί.

Βαθμωτή κάθοδος

Ο κανόνας μάθησης των Widrow-Hoff είναι τύπου **βαθμωτής καθόδου** (gradient descent). Αυτό σημαίνει ότι ο αλγόριθμος βασιζόμενος σε τοπική πληροφορία του σφάλματος προσπαθεί να οδηγήσει το σύστημα προς περιοχές της επιφάνειας σφάλματος όπου αυτό μειώνεται. Προς τοπικά, δηλαδή, ελάχιστα.

Επιγραμματικά παραθέτουμε την εξίσωση του κανόνα μάθησης των Widrow-Hoff:

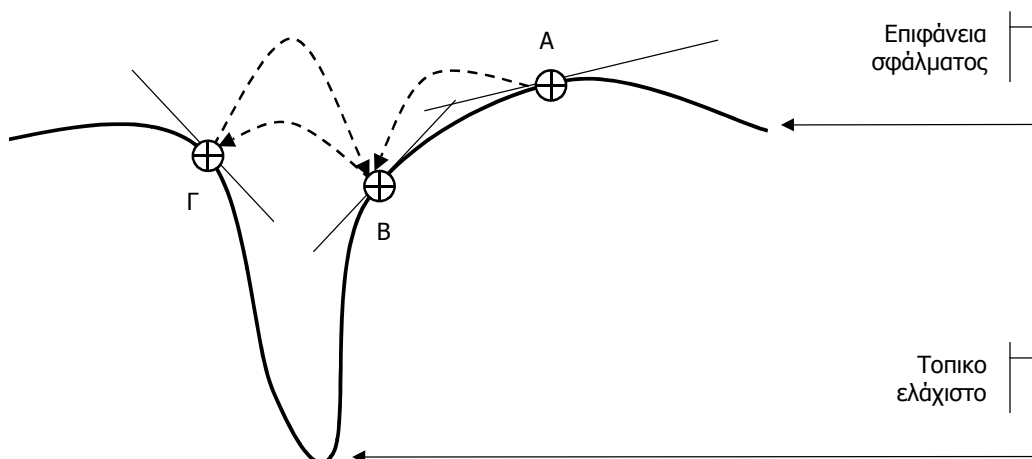
$$\mathbf{w} = \mathbf{w} + 2 \cdot a \cdot e \cdot \mathbf{x}$$

όπου a είναι ένας συντελεστής ρυθμού μάθησης.

Ρυθμός μάθησης

Ο **ρυθμός μάθησης** (learning rate) καθορίζει την ταχύτητα με την οποία ο αλγόριθμος οδηγεί το σύστημα προς την κατεύθυνση όπου (φαινομενικά) μειώνεται το σφάλμα. Το γνωστό **δίλημμα σύγκλιση / ταχύτητας** είναι συνδεδεμένο άμεσα με την επιλογή της τιμής του συντελεστή αυτού. Μεγάλες τιμές οδηγούν σε γρήγορη κίνηση προς την κατεύθυνση μείωσης του σφάλματος με μεγάλα βήματα. Τα μεγάλα όμως βήματα έχουν ως αποτέλεσμα την απώλεια τη διακριτικής ικανότητας του αλγορίθμου: στενά «πηγάδια» ελαχιστοποίησης στην επιφάνεια του σφάλματος είναι πιθανόν να παραβλεφθούν και ο αλγόριθμος να παγιδευτεί σε αναπηδήσεις μεταξύ των άκρων του πηγαδιού χωρίς την δυνατότητα να εισέλθει.

Στο σχήμα 24 απεικονίζεται η εξέλιξη ενός αλγορίθμου βαθμωτής καθόδου όπου ο ρυθμός μάθησης έχει λάβει υψηλές τιμές. Το σύστημα ξεκινώντας από το σημείο A και με βάση την κλίση του σφάλματος στο σημείο αυτό, οδηγείται στο σημείο B πραγματοποιώντας ένα βήμα μεγέθους ανάλογου με το ρυθμό μάθησης. Ομοίως από το σημείο B οδηγείται στο σημείο Γ. Η κλίση της επιφάνειας σφάλματος αλλάζει πρόσημο. Ως αποτέλεσμα, ο αλγόριθμος οδηγεί το σύστημα προς την αντίθετη κατεύθυνση, δηλαδή πίσω προς το σημείο B, εκκινώντας έτσι μια διαδοχή αναπηδήσεων μεταξύ των σημείων B και Γ.

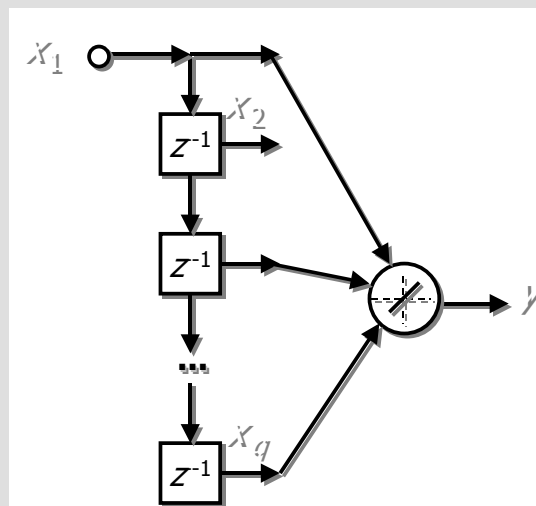


Σχήμα 24. Βαθμωτή κάθοδος με υψηλό ρυθμό μάθησης

Παράδειγμα: Γραμμική Νευρωνική Πρόβλεψη με Μονάδες Καθυστέρησης

Ένα τυπικό πεδίο εφαρμογής των γραμμικών νευρωνικών δικτύων είναι αυτό της πρόβλεψης. Με βάση ένα αριθμό από παρελθόντα δείγματα μιας ακολουθίας, επιδιώκουμε να προβλέψουμε το δείγμα που θα εμφανιστεί κάποια επόμενη χρονική στιγμή.

Για την επίλυση αυτού του προβλήματος, μπορεί να χρησιμοποιηθεί ένα γραμμικό νευρωνικό δίκτυο ενός νευρώνα όπως αυτό του σχήματος 25 που ακολουθεί.



Σχήμα 25. Νευρωνική πρόβλεψη

Τα στοιχεία καθυστέρηση που στο σχήμα συμβολίζονται με τα μπλοκ z^{-1} , έχουν το ρόλο μνήμης. Έτσι, εάν η πρόβλεψή μας βασίζεται στα προηγούμενα q δείγματα, αντί να παρέχουμε στο σύστημα κάθε φορά q εισόδους εξωτερικά, το σύστημα αναλαμβάνει να τις αποθηκεύει μόνο του με τη βοήθεια των στοιχείων καθυστέρησης και να τις τροφοδοτεί στο νευρωνικό δίκτυο (που στη συγκεκριμένη περίπτωση αποτελείται από ένα μοναδικό νευρώνα).

Εάν όλα τα δεδομένα της χρονοσειράς είναι διαθέσιμα εκ των προτέρων, το δίκτυο μπορεί να εκπαιδευτεί με τη βοήθεια του αλγορίθμου ελαχίστων μέσων τετραγώνων. Στην αντίθετη περίπτωση, ο κανόνας μάθησης Widrow-Hoff θα πρέπει να χρησιμοποιηθεί και το δίκτυο να περάσει μια φάση εκπαίδευσης όπου κάθε νέο δείγμα θα χρησιμοποιείται για να βελτιώσει τη συμπεριφορά του, μέχρι αυτή να φτάσει στα επιθυμητά επίπεδα.

Αλγόριθμος Ανάστροφης Διάδοσης Σφάλματος

Ο **αλγόριθμος ανάστροφης διάδοσης σφάλματος** (backpropagation of error) αποτελεί έναν από τους βασικότερους αλγόριθμους εκπαίδευσης στο χώρο των νευρωνικών δικτύων. Έχει τη δυνατότητα να εκπαιδεύσει νευρωνικά δίκτυα προστροφοδότησης αυθαίρετου αριθμού στρωμάτων.

Αποτελεί και αυτός ένα παράδειγμα επαναληπτικού αλγορίθμου βαθμωτής καθόδου και στηρίζεται στην τοπική πληροφορία της κλίσης της επιφάνειας του σφάλματος για να υπολογίσει (τοπικά) βέλτιστες τιμές για τις παραμέτρους του δικτύου.

Συνοπτικά αναφέρουμε τον βασικό κανόνα που χρησιμοποιεί ο οποίος είναι της γνώριμης μορφής:

$$\mathbf{w} = \mathbf{w} - a \mathbf{g}$$

όπου $\mathbf{g}$ είναι η κλίση τη επιφάνειας σφάλματος στο τρέχον σημείο και a ο επιθυμητός ρυθμός μάθησης

Το βασικό πρόβλημα το οποίο επιλύει ο αλγόριθμος ανάστροφης διάδοσης σφάλματος, είναι εκείνο της απόδοσης σφάλματος.

Η **πρόβλημα της απόδοσης σφάλματος** (credit assignment) σχετίζεται με τον υπολογισμό του σφάλματος για νευρώνες ενδιάμεσων στρωμάτων που δεν συμμετέχουν άμεσα στην έξοδο. Κάθε αλγόριθμος εκπαίδευσης υπό επίβλεψη, στηρίζεται στο σφάλμα ενός νευρώνα για να υπολογίσει νέες

τιμές για τις παραμέτρους του νευρώνα αυτού (όπως τα βάρη των συνδέσεων εισόδου και το κατώφλι). Για του κόμβους εξόδου, η πληροφορία αυτή είναι άμεσα διαθέσιμη αφού υπολογίζεται άμεσα ως η διαφορά μεταξύ επιθυμητής και πραγματικής εξόδου στον νευρώνα αυτό. Για τους «κρυμμένους» νευρώνες, η επιθυμητή έξοδος δεν είναι προφανής. Ο αλγόριθμος ανάστροφης διάδοσης σφάλματος καταφέρνει να υπολογίσει απευθείας το σφάλμα στους κρυμμένους νευρώνες «διαδίδοντάς» κατάλληλα το σφάλμα των νευρώνων εξόδου προς τα πίσω (προς τους νευρώνες εισόδου).

Βελτιώσεις στον Αλγόριθμο Ανάστροφης Διάδοσης Σφάλματος

Αν και ο αλγόριθμος ανάστροφης διάδοσης σφάλματος αποτελεί βασικό εργαλείο, παρουσιάζει ένα αριθμό προβλημάτων πολλά από τα οποία έχουν αντιμετωπιστεί είτε με προσθήκες στη βασική λογική του αλγορίθμου, είτε με τροποποιημένες εκδόσεις του.

Μερικές πρακτικές που συχνά ακολουθούνται σε συνδυασμό με τον αλγόριθμο ανάστροφης διάδοσης σφάλματος ώστε να βελτιωθεί η απόδοσή του, αναφέρονται επιγραμματικά παρακάτω:

- **Κανονικοποίηση των δεδομένων εκπαίδευσης**

Μεγάλες διακυμάνσεις στις απόλυτες τιμές των παραδειγματικών προτύπων δυσχεραίνουν τη διαδικασία εκπαίδευσης. Τα δεδομένα αυτά, σκόπιμο είναι να κανονικοποιούνται.

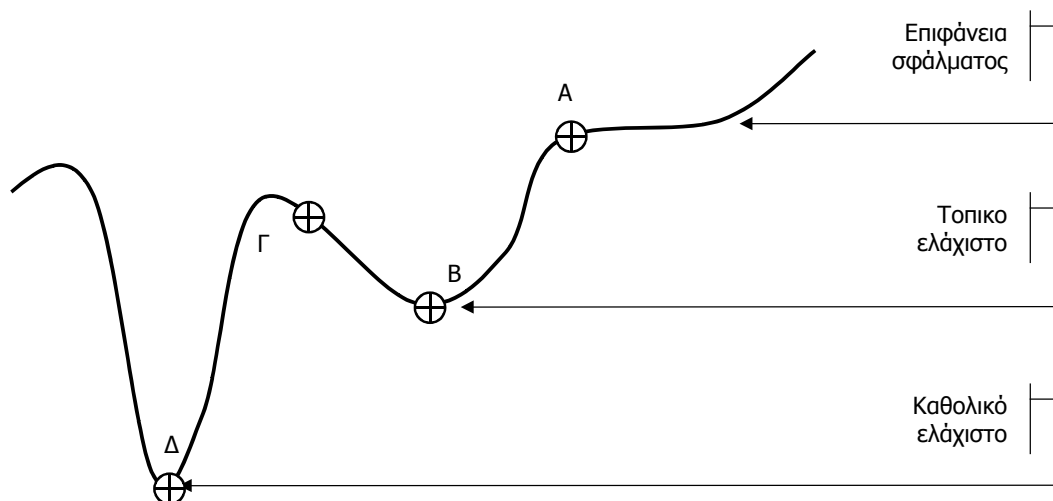
- **Μεταβλητός ρυθμός μάθησης**

Το δίλημμα σύγκλισης / ταχύτητας που αφορά στην επιλογή της τιμής του ρυθμού μάθησης, συχνά αντιμετωπίζεται με τη χρήση μεταβλητού ρυθμού. Κατά το στάδιο εκπαίδευσης, ο ρυθμός μάθησης λαμβάνει αρχικά μεγάλες τιμές ώστε το σφάλμα να οδηγηθεί γρήγορα σε περιοχές χαμηλού σφάλματος. Στην συνέχεια, ο ρυθμός μειώνεται ώστε να βελτιωθεί η διακριτική ικανότητα του αλγορίθμου.

- **Παράμετρος ορμής**

Η προσθήκη στην λογική του αλγορίθμου μιας **παραμέτρου ορμής** (momentum) μπορεί να αποτελέσει αρκετά σημαντικό βοήθημα στην αντιμετώπιση καταστάσεων «παγίδευσης» σε τοπικά ελάχιστα.

Παγίδευση σε τοπικά
ελάχιστα



Σχήμα 26. Τοπικά ελάχιστα της επιφάνειας σφάλματος

Αν υποθέσουμε ότι ο αλγόριθμος εκπαίδευσης ξεκινάει από το σημείο Α της επιφάνειας σφάλματος, κινούμενος σύμφωνα με την κλίση της θα κατευθυνθεί προς το σημείο Γ. Φτάνοντας όμως εκεί, θα υποχρεωθεί να κινηθεί αντίθετα αφού φαινομενικά το σφάλμα αυξάνεται προς τα αριστερά. Θα καταλήξει έτσι στο σημείο Β το οποίο είναι μεν ελάχιστο αλλά τοπικό και όχι καθολικό. Η παράμετρος ορμής θα συνηγορήσει στη διατήρηση της πορείας του αλγορίθμου προς την κατεύθυνση την οποία είχε αρχικά επιλέξει, επιτρέποντάς του να αποφύγει το σημείο Β και να εντοπίσει το καθολικό ελάχιστο Δ.

• Περικοπή συνδέσεων και νευρώνων

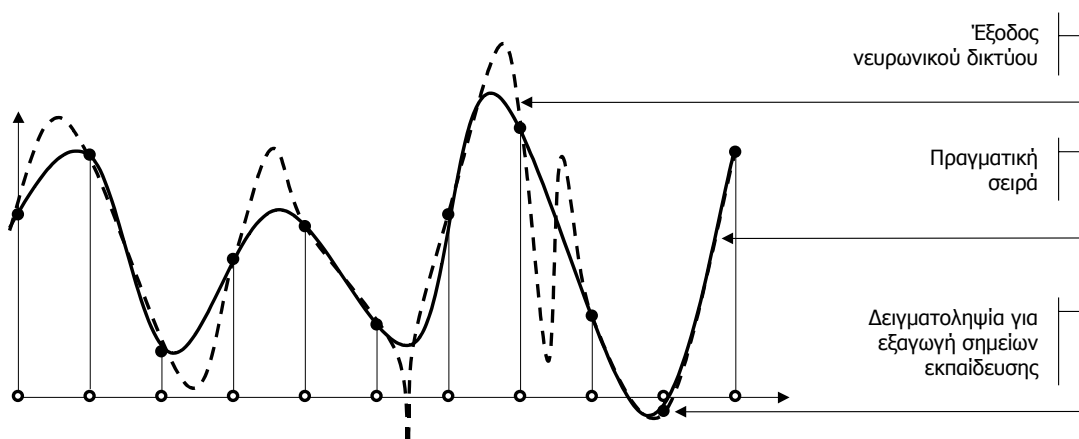
Κατά την σχεδίαση νευρωνικών δικτύων, η ακριβής δομή του δικτύου δεν υποδεικνύεται (γενικά) από συγκεκριμένους κανόνες. Η επιλογή του αριθμού των στρώματων, του αριθμού των νευρώνων ανά στρώμα, κ.λπ. παραμένει μια σημαντική απόφαση που λαμβάνεται με βάση εμπειρικούς κανόνες. Οι επιλογές, όμως, αυτές είναι ιδιαίτερα σημαντικές αφού επηρεάζουν άμεσα τόσο για την διαδικασία εκπαίδευσης του δικτύου όσο και την τελική απόδοσή του κατά τη φάση λειτουργίας του.

Η ισχύς ενός δικτύου, η ικανότητά του δηλαδή να διαμορφώνει απεικονίσεις εισόδων σε εξόδους, βρίσκεται σε άμεση σχέση με τον αριθμό των νευρώνων και συνδέσεων του αφού ουσιαστικά, οι ελεύθερες παράμετροι του μοντέλου είναι τα βάρη των συνδέσεων και τα κατώφλια των νευρώνων.

Επιλέγοντας ένα δίκτυο με λίγα δομικά στοιχεία, ισοδυναμεί με την επιλογή ενός μοντέλου μικρού αριθμού παραμέτρων για την επίλυση ενός προβλήματος. Όταν οι παράμετροι αυτές δεν επαρκούν (υπο-παραμετροποίηση — under-parameterization) το νευρωνικό δεν θα καταφέρει να επιτύχει την ζητούμενη απεικόνιση εισόδων / εξόδων στο επιθυμητό επίπεδο σφάλματος. Αντίθετα, όταν ο αριθμός των παραμέτρων είναι αρκετά μεγαλύτερος από αυτόν που πραγματικά χρειάζονται (υπερ-παραμετροποίηση — over-parameterization), το τελικό δίκτυο θα καταφέρει σχετικά εύκολα να μειώσει το σφάλμα κάτω από το επιθυμητό επίπεδο, αλλά θα παρουσιάζει πολύ περιορισμένες ικανότητες γενίκευσης.

Η **γενίκευση** (generalization) είναι μια έννοια που σχετίζεται με τη συμπεριφορά του δικτύου σε «καινοφανή» δεδομένα, δηλαδή δεδομένα που δεν περιλαμβάνονταν στο σύνολο των δεδομένων εκπαίδευσης. Η έννοια της γενίκευσης είναι ιδιαίτερα σημαντική αφού καθορίζει την συμπεριφορά του δικτύου κατά τη φάση λειτουργίας του, δηλαδή σε «πραγματικές» συνθήκες.

Υπερ-παραμετροποιημένα δίκτυα τείνουν να παρουσιάζουν μειωμένες ικανότητες γενίκευσης για τον απλό λόγο ότι η υπερβολική ισχύς που διαθέτουν ως μοντέλα (λόγω του μεγάλου αριθμού ελεύθερων παραμέτρων) δεν έχει μορφοποιηθεί από τα δεδομένα εκπαίδευσης. Έτσι, το δίκτυο «αυτοσχεδιάζει» μεταξύ των προτύπων εκπαίδευσης παρουσιάζοντας μια απρόβλεπτη συμπεριφορά.



Σχήμα 27. Παράδειγμα υπερ-παραμετροποίησης

Στο σχήμα 27, φαίνεται μια ενδεικτική περίπτωση συμπεριφοράς υπερ-παραμετροποιημένου συστήματος. Με τη συνεχή γραμμή απεικονίζεται η σειρά για την οποία επιθυμούμε να σχεδιάσουμε έναν προβλέπτη. Για το σκοπό αυτό δειγματοληψούμε τη σειρά στα σημεία που φαίνονται ώστε να δημιουργήσουμε ένα σύνολο παραδειγματικών προτύπων για την εκπαίδευση του νευρωνικού προβλέπτη. Αν θεωρήσουμε ότι το νευρωνικό που επιλέγουμε είναι υπερβολικά μεγάλης διάστασης τότε, σχετικά γρήγορα, το επίπεδο σφάλματος κατά την φάση εκπαίδευσης θα πέσει στα επιθυμητά όρια. Όμως, κατά την φάση λειτουργίας, όταν το νευρωνικό θα διεγείρεται με εισόδους που δεν περιλαμβάνονται στο σύνολο των παραδειγματικών προτύπων εκπαίδευσης, μπορεί να παρουσιάζει συμπεριφορά όπως εκείνη που απεικονίζει η διακεκομμένη γραμμή.

Υπο- και υπερ-
παραμετροποίηση

Γενίκευση

Μια λύση για το πρόβλημα της υπερ-παραμετροποίησης δίνεται από την τεχνική της περικοπής.

Η **περικοπή** (pruning) ενός δικτύου συνίσταται στη αφαίρεση συνδέσεων (ή και νευρώνων) που κατά την φάση εκπαίδευσης συμμετείχαν ελάχιστα στη διαμόρφωση των εξόδων και έλαβαν αμελητέες τιμές σε σχέση με τον μέσο όρο.

Διάφοροι αλγόριθμοι έχουν κατά καιρούς προταθεί ως εναλλακτικές επιλογές του αλγόριθμου ανάστροφης διάδοσης σφάλματος. Οι αλγόριθμοι αυτοί παρουσιάζουν σημαντικές βελτιώσεις έναντι αυτού σε θέματα όπως η ταχύτητα σύγκλισης (π.χ. αλγόριθμοι Levenberg-Marquardt, quasi-Newton, κ.λπ.), η αντιμετώπιση καταστάσεων παγίδευσης σε τοπικά ελάχιστα, κ.λπ.

Γενετικοί Αλγόριθμοι

Εισαγωγή

Οι **γενετικοί αλγόριθμοι** (ΓΑ — Genetic algorithms, GA) αποτελούν, ουσιαστικά, ικανούς αλγόριθμους βελτιστοποίησης. Δανείζονται πρότυπα και μηχανισμούς εξέλιξης από τη φύση και τη γενετική για να δημιουργήσουν, να συντηρήσουν και να εξελίξουν ένα πληθυσμό από κωδικοποιημένες ομάδες παραμέτρων, κάθε μία από τις οποίες αποτελεί και δυνητική λύση του εκάστοτε προβλήματος.

Οι γενετικοί αλγόριθμοι είναι ιδιαίτερα ελκυστικοί αφού απαιτούν ελάχιστη πληροφορία σχετικά με το συγκεκριμένο πρόβλημα που καλούνται να επιλύσουν και συγκεκριμένα μόνο την αντικειμενική συνάρτηση που αποτιμά την ποιότητα μιας συγκεκριμένης ομάδας παραμέτρων.

Λόγω της σημαντικής αντιστοιχίας μεταξύ γενετικών αλγορίθμων και βιολογικών διεργασιών, μεγάλο ποσοστό των χρησιμοποιούμενων όρων παραπέμπουν σ'αυτές. Η τελευταία παράγραφος περιέχει έναν πίνακα με συνήθεις όρους συνοδευόμενος με τους αντίστοιχους αγγλικούς.

Μετα-εργασία

Ως αλγόριθμοι αναζήτησης, οι γενετικοί δεν ενδείκνυνται για προβλήματα που απαιτούν λύσεις σε πραγματικό χρόνο. Αντίθετα, λόγω της μεγάλης ευελιξίας που προσφέρουν, οι γενετικοί αλγόριθμοι είναι ιδιαίτερα ικανοί ως «μετα-εργαλεία», δηλαδή ως συστήματα «υψηλότερου» επιπέδου ικανά να παράγουν (σχεδιάσουν) άλλα συστήματα.

Οι γενετικοί αλγόριθμοι προτάθηκαν από τον John Holland το 1975 μέσα από την (διάσημη πλέον) μονογραφία του *Adaptation in Natural and Artificial Systems* [HOLLAND, 1975]. Η έρευνα αυτή ήταν προσανατολισμένη στην μελέτη των μηχανισμών της φύσης και την αναπαραγωγή του στο πλαίσιο των υπολογιστικών μηχανών. Η ερευνητική ομάδα του Holland για πολύ καιρό αποτέλεσε και συνεχίζει να αποτελεί ένα σημαντικό κέντρο ανάπτυξης και εξέλιξης της θεωρίας των γενετικών αλγορίθμων.

Βελτιστοποίηση

Η ιδέα της **βελτιστοποίησης** κατέχει κεντρική θέση σε διάφορες περιοχές της επιστήμης καθώς ένα μεγάλο σύνολο από προβλήματα μπορούν να ενταχθούν στο πλαίσιο αυτό. Η μάθηση, για παράδειγμα, μπορούμε σαφώς να θεωρήσουμε ότι ουσιαστικά συνιστά μια διαδικασία βελτιστοποίησης αφού το ζητούμενο είναι ένα σύνολο τιμών για τις ελεύθερες παραμέτρους του μοντέλου οι οποίες θα του δώσουν την επιθυμητή συμπεριφορά.

Η βελτιστοποίηση, με τη σειρά της, έχει αδιάρρηκτους δεσμούς με την ιδέα της **αναζήτησης**. Από ένα (συνήθως) ευρύ πεδίο δυνατών τιμών για ένα σύνολο παραμέτρων ενός συστήματος, αναζητούμε εκείνες που θα μεγιστοποιήσουν ή ελαχιστοποιήσουν ένα προκαθορισμένο κριτήριο ποιότητας για το σύστημα αυτό.

Αυτές είναι οι δύο βασικές έννοιες μέσα από τις οποίες θα πρέπει να εξεταστούν οι γενετικοί αλγόριθμοι.

Θεωρία
βελτιστοποίησης

Εάν προσπαθούσαμε να περιγράψουμε την θεωρία βελτιστοποίησης θα μπορούσαμε να πούμε ότι:

*Η θεωρία βελτιστοποίησης ασχολείται με τον εντοπισμό ενός «**βέλτιστου**» δεδομένης μιας μεθόδου αποτίμησης του «**καλού**» και του «**κακού**» και ενός «**μηχανισμού μεταβολής**» τους.*

Αν και αφαιρετική, η παραπάνω περιγραφή εμπλέκει όλα τα βασικά συστατικά της βελτιστοποίησης. Για να βελτιστοποιήσουμε οτιδήποτε, θα πρέπει να έχουμε ένα μέτρο για το τί είναι καλό, τί είναι καλύτερο και τί όχι. Επιπρόσθετα, αφού διαγνώσουμε την παρούσα κατάσταση, θα πρέπει να έχουμε κάποιον τρόπο να ωθήσουμε το σύστημα προς καλύτερες ή, τουλάχιστον, προς διαφορετικές.

Πλεονεκτήματα Γενετικών Αλγορίθμων

Οι γενετικοί αλγόριθμοι παρουσιάζουν ισχυρά πλεονεκτήματα έναντι αντίστοιχων μεθόδων αναζήτησης και βελτιστοποίησης:

- **Υπολογιστική απλότητα**

Οι γενετικοί αλγόριθμοι είναι ιδιαίτερα απλοί στην υλοποίησή τους. Το γεγονός αυτό δημιούργησε μεγάλη εντύπωση σε αντιδιαστολή με την ευελιξία και τη γενική ικανότητα που

επιδεικνύουν. Όμως, η απλότητα αυτή υποστηρίζεται από μια μαθηματικά θεμελιωμένη θεωρία, τη θεωρία των σχημάτων (schema theory).

- **Αποτελεσματικότητα**

Η αποτελεσματικότητα των γενετικών αλγορίθμων στην επίλυση προβλημάτων αναζήτησης και βελτιστοποίησης έχει πιστοποιηθεί τόσο θεωρητικά μέσω της θεωρίας των σχημάτων όσο και πρακτικά μέσω των πολυάριθμων εφαρμογών που έχουν βρει σε ένα μεγάλο εύρος επιστημονικών περιοχών.

- **Γενικότητα**

Χαρακτηριστικό στοιχείο των γενετικών αλγορίθμων αποτελεί η αποσύζευξή τους από τις λεπτομέρειες του εκάστοτε προβλήματος το οποίο καλούνται να επιλύσουν. Αυτή η «απόσταση» που κρατούν από τα ειδικά χαρακτηριστικά του κάθε προβλήματος τους απαλλάσσει από την ανάγκη δεσμευτικών υποθέσεων σχετικά με τον χώρο όπου πραγματοποιείται η αναζήτηση. Έτσι, σε αντίθεση με πιο μαθηματικές μεθόδους, οι γενετικοί αλγόριθμοι δεν απαιτούν καμία υπόθεση συνέχειας, παραγωγισιμότητας, μονοτονίας, κ.λπ.

Παραδοσιακές Μέθοδοι Αναζήτησης

Κατηγορίες
παραδοσιακών
μεθόδων
αναζήτησης

Οι τυπικές μέθοδοι αναζήτησης μπορούν να διαιρεθούν στις παρακάτω κατηγορίες:

- **Μέθοδοι βασισμένες σε μαθηματική ανάλυση**

Τις μεθόδους που βασίζονται σε μαθηματική ανάλυση (calculus-based), μπορούμε να τις διαχωρίσουμε σε δύο υποκατηγορίες:

- **Έμμεσες μέθοδοι**

Οι μέθοδοι αυτές θεωρούν ως πιθανά σημεία ακρότατων (μεγίστων ή ελαχίστων) τα σημεία μηδενισμού της παραγώγου της επιφάνειας σφάλματος στον χώρο των παραμέτρων.

- **Άμεσες μέθοδοι**

Οι μέθοδοι αυτές, ξεκινώντας από κάποιο σημείο, πραγματοποιούν μια βηματική αναζήτηση προς την κατεύθυνση που υποδεικνύει η τοπική τιμή της παραγώγου. Τυπικό παράδειγμα τέτοιας μεθόδου αποτελούν οι μέθοδοι βαθμωτής καθόδου που εξετάστηκαν σε προηγούμενη παράγραφο.

Βασικά προβλήματα των μεθόδων αυτής της κατηγορίας είναι ότι είναι τοπικές (περιορίζονται, δηλαδή, σε τοπικά ακρότατα) και ότι βασίζονται σε παραγώγους πράγμα συχνά μη ρεαλιστικό αφού οι παράγωγοι όχι μόνο μπορεί να μην είναι διαθέσιμες αλλά ούτε καν να υπάρχουν ή να μπορούν να οριστούν.

- **Μέθοδοι απαρίθμησης**

Οι μέθοδοι απαρίθμησης (enumeration) επιχειρούν μια εξαντλητική αποτίμηση κάθε δυνατού σημείου του χώρου αναζήτησης (κάθε δυνατού συνδυασμού των παραμέτρων).

Η χρησιμότητα των αλγορίθμων αυτών είναι, προφανώς, περιορισμένη σε περιπτώσεις χώρων μικρής διάστασης. Κάθε προσπάθεια εφαρμογής των μεθόδων αυτών σε ευρύτερους χώρους έρχεται αντιμέτωπη με την περίφημη «κατάρτα της πολυδιάστασης» (curse of dimensionality), η οποία συνίσταται στο ότι προσθέτοντας μία ακόμη διάσταση στο πρόβλημα, η πολυπλοκότητα του προβλήματος αυξάνεται πολλαπλασιαστικά.

- **Τυχαίες μέθοδοι**

Οι μέθοδοι τυχαίας αναζήτησης (random search) βασίζονται στην τυχαία «περιπλάνηση» στον χώρο ευελπιστώντας ότι έτσι θα καταφέρουν τελικά να εντοπίσουν «καλά» σημεία.

Οι τυχαίες μέθοδοι έχουν το πρόβλημα που δηλώνει ο τίτλος τους: είναι τυχαίες και αδυνατούν να παρέχουν εύρωστες λύσεις σε μη τετριμμένα προβλήματα.

Curse of
dimensionality

• Υβριδικές μέθοδοι

Οι μέθοδοι αυτές συνδυάζουν κάποιες από τις παραπάνω μεθόδους. Για παράδειγμα, μια υβριδική μέθοδος θα μπορούσε να χρησιμοποιήσει τυχαία αναζήτηση μέχρι να εντοπίσει κάποιο υποσχόμενο σημείο οπότε και θα εκκινήσει μια μέθοδο βασισμένη σε παραγώγους για τον ακριβή εντοπισμό κάποιου ελαχίστου στην περιοχή αυτή.

Γενετικοί Αλγόριθμοι Έναντι Παραδοσιακών Μεθόδων

Με βάση την προηγούμενη συνοπτική παρουσίαση των παραδοσιακών μεθόδων αναζήτησης, μπορούμε να προβούμε σε μια πρώτη σύγκρισή τους με τους γενετικούς αλγορίθμους. Οι διαφορές που μπορούμε άμεσα να αναγνωρίσουμε είναι οι ακόλουθες:

- Οι ΓΑ δεν επιδρούν στις ίδιες τις παραμέτρους του υπό επίλυση προβλήματος αλλά σε μία **κωδικοποίηση** των παραμέτρων αυτών. Συνεπώς οι ΓΑ παραμένουν ανεπηρέαστοι από τις ακριβείς τιμές των εμπλεκόμενων μεγεθών καθώς εργάζονται στη μορφή τους όπως αυτή καθορίζεται από την χρησιμοποιούμενη κωδικοποίηση.
- Οι ΓΑ δεν εργάζονται με ένα μοναδικό σημείο του χώρου αλλά διατηρούν ένα **πληθυσμό σημείων** για την διεξοδικότερη εξερεύνηση του χώρου. Έτσι, προβλήματα όπως η παγίδευση σε τοπικά ελάχιστα δεν αφορούν τους ΓΑ.
- Οι ΓΑ χρησιμοποιούν **συναρτήσεις αποτίμησης** και όχι παραγώγους ή άλλη παρόμοια πληροφορία. Η κατάστρωση μιας συνάρτησης αποτίμησης είναι σαφώς απλούστερη από τον υπολογισμό παραγώγων και απαιτεί από ελάχιστες ως και καθόλου παραδοχές.
- Οι ΓΑ χρησιμοποιούν **πιθανοτικούς κανόνες μετάβασης** και όχι ντετερμινιστικούς.

Τα παραπάνω θέματα αυτά θα γίνουν πιο κατανοητά μετά την αναλυτικότερη παρουσίαση των γενετικών αλγορίθμων που δίνεται στην επόμενη παράγραφο.

Η Λειτουργία των Γενετικών Αλγορίθμων

Ας θεωρήσουμε ένα αφαιρετικό πρόβλημα για την επίλυση του οποίου θα χρησιμοποιήσουμε γενετικούς αλγορίθμους.

Έστω, λοιπόν, ότι επιθυμούμε να βρούμε την κατάλληλη τιμή x η οποία όταν διεγείρει ένα άγνωστο σύστημα (μαύρο κουτί) θα δώσει την μέγιστη δυνατή έξοδο.



Σχήμα 28. Το πρόβλημα του «μαύρου κουτιού»

Τα απαιτούμενα βήματα για την γενετική επίλυση του δεδομένου προβλήματος, είναι τα ακόλουθα:

1. Επιλογή σχήματος κωδικοποίησης των παραμέτρων του προβλήματος
2. Επιλογή μεγέθους πληθυσμού και αρχικοποίηση
3. Επιλογή τιμών για τις παραμέτρους των γενετικών τελεστών
4. Βρόχος επαναληπτικής εκτέλεσης αλγορίθμου
 - 4.1. Αποτίμηση μελών πληθυσμού
 - 4.2. Επιλογή ζευγών μελών με τη μέθοδο ρουλέτας
 - 4.3. Εφαρμογή τελεστή διασταύρωσης
 - 4.4. Εφαρμογή τελεστή μετάλλαξης
 - 4.5. Επανάληψη από το βήμα 4.1

Τα βήματα αυτά θα εξεταστούν ξεχωριστά στις ακόλουθες παραγράφους.

Κωδικοποίηση Παραμέτρων

Οι ΓΑ εργάζονται πάνω σε κωδικοποίηση των παραμέτρων και όχι στις παραμέτρους καθαυτές. Ένα κατάλληλο **σχήμα κωδικοποίησης** μετατρέπει τις παραμέτρους σε συμβολοσειρές από ένα πεπερασμένο αλφάβητο. Το πιο σύννηθες αλφάβητο είναι το δυαδικό αλφάβητο που παράγει δυαδικές συμβολοσειρές, δηλαδή συμβολοσειρές αποτελούμενες αποκλειστικά από δύο ψηφία: το 0 και το 1. Για το σκοπό αυτό, μπορούν να χρησιμοποιηθούν οι κλασσικές μέθοδοι δυαδικής αναπαράστασης αριθμών.

Η αντίστροφη διαδικασία, αυτή της **αποκωδικοποίησης**, συνίσταται στην μετατροπή μιας κωδικοποιημένης συμβολοσειράς στην αντίστοιχη τιμή της παραμέτρου.

Επιλογή Μεγέθους Πληθυσμού και Αρχικοποίηση

Το μέγεθος του πληθυσμού ενός ΓΑ αποτελεί βασική επιλογή κατά το στάδιο της σχεδίασής του. Κάθε αλγόριθμος αναζήτησης αναμένουμε να παρουσιάζει μια χρονική εξέλιξη. Για να είναι εφικτή αυτή η εξέλιξη, ο αλγόριθμος θα πρέπει να παρουσιάζει κάποιο άμεσο ή έμμεσο είδος μνήμης ώστε η απόδοσή του να βελτιώνεται με το χρόνο. Ο ρόλος της μνήμης σε έναν γενετικό αλγόριθμο έχει ανατεθεί στον πληθυσμό. Η μόνη πληροφορία που μπορεί να ανακαλέσει ένας γενετικός είναι η πληροφορία που βρίσκεται κατανεμημένη στο μέλη του πληθυσμού του την οποία ο ίδιος παρήγαγε. Συνεπώς, το άμεσο συμπέρασμα που προκύπτει είναι ότι όσο μεγαλύτερος είναι ο πληθυσμός τόσο πιο ισχυρός είναι ο αλγόριθμος. Ένα αντίπαλο επιχείρημα, όμως, έχει να κάνει με την παρατήρηση ότι το μέγεθος του πληθυσμού επηρεάζει αρνητικά την «κινητικότητα» του. Μεγαλύτεροι πληθυσμοί παρουσιάζουν μεγαλύτερη αδράνεια και συνεπώς βραδύτερη εξέλιξη.

Γενετικοί Τελεστές

Ως αλγόριθμος αναζήτησης, ο βασικός στόχος του γενετικού αλγορίθμου είναι η κινητοποίηση των μελών του πληθυσμού ώστε να ανακαλυφθούν νέα και (αισίως) καλύτερα σημεία του χώρου. Όμως, η πορεία αυτή στον υπό αναζήτηση χώρο χαρακτηρίζεται από δύο αντιφατικές επιδιώξεις:

- την επιδίωξη για την ευρύτερη δυνατή εξερεύνηση του χώρου ώστε να ανακαλυφθούν όσο το δυνατόν περισσότερες «ενδιαφέρουσες» περιοχές, πράγμα που απαιτεί υψηλή κινητικότητα, και
- την επιδίωξη για την διεξοδική διερεύνηση των περιοχών αυτών ώστε να επιτευχθεί μεγαλύτερη διακριτική ικανότητα και να εντοπιστούν με ακρίβεια τα «βέλτιστα» σημεία, πράγμα που απαιτεί χαμηλή κινητικότητα.

Το δίλημμα αυτό είναι γνωστό ως **δίλημμα εξερεύνησης / διερεύνησης** (exploration / exploitation dilemma).

Η κινητικότητα ενός γενετικού αλγορίθμου καθορίζεται έμμεσα μέσω των δύο θεμελιωδών γενετικών τελεστών του: του τελεστή της διασταύρωσης και του τελεστή της μετάλλαξης. Οι τελεστές αυτοί επιδρούν στα μέλη του πληθυσμού ενός γενετικού αλγορίθμου με σκοπό την εξέλιξή τους και την παραγωγή ενός νέου βελτιωμένου πληθυσμού σε κάθε βήμα του αλγορίθμου.

Ο **τελεστής της διασταύρωσης** (crossover) ανταλλάσσει μέρος της κωδικοποιημένης πληροφορίας δύο χρωμοσωμάτων (συμβολοσειρών). Επιλέγοντας ένα σημείο για την πραγματοποίηση της διασταύρωσης και ανταλλάσσοντας τα δύο τμήματα μεταξύ των χρωμοσωμάτων-γονέων προκύπτουν δύο νέα χρωμοσώματα-απόγονοι.

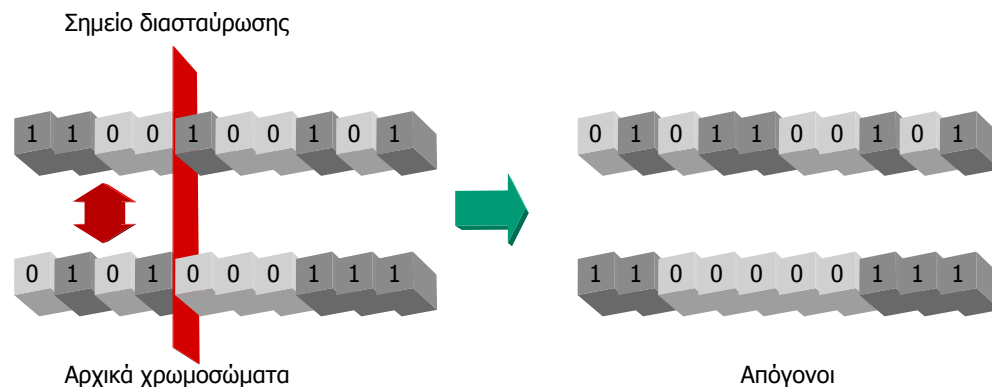
Καθώς η ανταλλαγή της πληροφορίας των δύο γονέων λαμβάνει χώρα στην κωδικοποιημένη τους αναπαράσταση, οι δύο απόγονοι περιέχουν καινοφανείς τιμές των παραμέτρων αφού η αποκωδικοποίησή τους μπορεί να δώσει εντελώς διαφορετικές τιμές για τις παραμέτρους του προβλήματος ενδιαφέροντος.

Μια σχηματοποίηση του τελεστή δίνεται στο σχήμα 29.

Η διασταύρωση λαμβάνει χώρα κατά το στάδιο παραγωγής ενός νέου πληθυσμού από έναν υπάρχων. Η διασταύρωση δεν πραγματοποιείται πάντα αλλά με μια πιθανότητα που συνήθως είναι αρκετά μεγάλη (τυπικά της τάξης του 0.7).

Δίλημμα
εξερεύνησης /
διερεύνησης

Τελεστής
διασταύρωσης



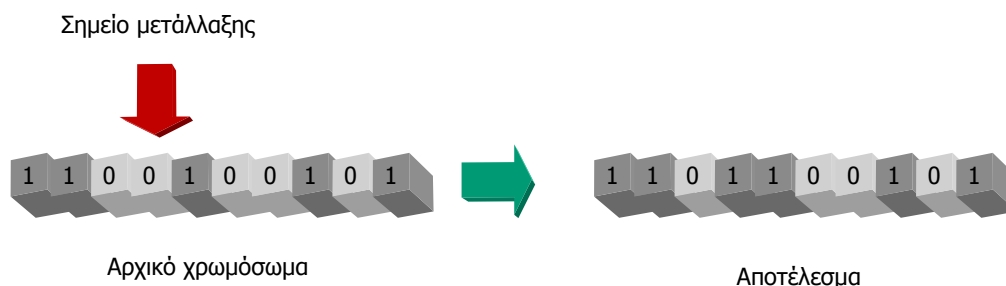
Σχήμα 29. Ο γενετικός τελεστής διασταύρωσης

Τελεστής
μετάλλαξης

Ο **τελεστής της μετάλλαξης** (mutation) είναι υπεύθυνος για την μεταβολή του γονιδιακού υλικού του χρωμοσώματος στο οποίο επιδρά, αντίθετα με τον τελεστή της διασταύρωσης ο οποίος απλά ανακατανέμει το υπάρχων υλικό δημιουργώντας νέους συνδυασμούς.

Η μετάλλαξη εφαρμόζεται με μια σχετικά μικρή πιθανότητα (τυπικά της τάξης του 0.03) κατά την φάση δημιουργίας ενός νέου πληθυσμού από τον πληθυσμό της προηγούμενης γενιάς.

Στο σχήμα 30 δίνεται μια σχηματοποίηση της δράσης του τελεστή αυτού.



Σχήμα 30. Ο γενετικός τελεστής μετάλλαξης

Οι δύο αυτοί τελεστές είναι οι πλέον χρησιμοποιούμενοι. Παρόλα αυτά δεν είναι και οι μοναδικοί δυνατοί. Πλήθος τελεστών έχουν προταθεί στη βιβλιογραφία, άλλοι γενικού σκοπού και άλλοι ειδικά σχεδιασμένοι για τη επίτευξη ειδικής συμπεριφοράς που απαιτείται από εξειδικευμένα προβλήματα.

Συνάρτηση Αποτίμησης

Μια ομάδα παραμέτρων υπό βελτιστοποίηση μετατρέπεται σε συμβολοσειρά μέσω της κωδικοποίησης. Η κωδικοποιημένη συμβολοσειρά υφίσταται την επίδραση των γενετικών τελεστών και νέες συμβολοσειρές παράγονται.

Για την αποτίμηση των νέων αυτών συμβολοσειρών, την εκτίμηση δηλαδή, του επιπέδου ποιότητάς τους ως υποψήφιες λύσεις στο εκάστοτε πρόβλημα, απαιτείται η ύπαρξη μιας **συνάρτησης αποτίμησης** (αντικειμενικής συνάρτησης — objective function). Οι νέες συμβολοσειρές αποκωδικοποιούνται για να δώσουν και πάλι ομάδες παραμέτρων οι οποίες τροφοδοτούνται στην αντικειμενική συνάρτηση για αποτίμηση. Το μόνο που απαιτείται από την συνάρτηση αυτή είναι η επιστροφή μίας και μόνο τιμής που να χαρακτηρίζει το πόσο καλής ποιότητας λύση δίνει η κάθε ομάδα παραμέτρων στο πρόβλημα (βαθμός καταλληλότητας παραμέτρων).

Η πληροφορία αυτή είναι η μοναδική εξαρτώμενη από το εκάστοτε πρόβλημα που ρητά απαιτείται από τον γενετικό αλγόριθμο και καθορίζει την «τύχη» της ομάδας αυτής στον αγώνα επιβίωσης που πραγματοποιείται κατά την εξέλιξη από γενιά σε γενιά κάτω από την πίεση του «κανόνα επιβίωσης του ισχυρότερου».

Βρόχος Εκτέλεσης Γενετικού Αλγορίθμου

Έχοντας επιλέξει κατάλληλο σχήμα κωδικοποίησης για το πρόβλημα, το μέγεθος του πληθυσμού και κατάλληλες τιμές για τους γενετικούς τελεστές και με δεδομένη μια κατάλληλη συνάρτηση αποτίμησης, ο γενετικός αλγόριθμος είναι πλέον έτοιμος για να εκκινήσει.

Το πρώτο βήμα είναι η **αρχικοποίηση** του πληθυσμού που λαμβάνει χώρα στην κωδικοποιημένη μορφή των μελών του. Εάν έχει επιλεγθεί η κλασσική δυαδική κωδικοποίηση, η αρχικοποίηση απλά συνίσταται στην διαδοχική ρίψη ενός κέρματος τόσες φορές όσο είναι και το συνολικό πλήθος των δυαδικών ψηφίων του πληθυσμού.

Στην συνέχεια, ο αλγόριθμος εισέρχεται σε ένα βρόχο επαναλήψεων που τερματίζεται είτε όταν ένας προδιαγεγραμμένος αριθμός βημάτων έχει ολοκληρωθεί ή όταν ο αλγόριθμος καταφέρει να εντοπίσει ένα σύνολο παραμέτρων για τις οποίες η συνάρτηση αποτίμησης θα επιστρέψει μια τιμή που να βρίσκεται στο επιθυμητό επίπεδο.

Σε κάθε επανάληψη, ο γενετικός αλγόριθμος πραγματοποιεί διαδοχικά τα παρακάτω βήματα:

- **Αποτίμηση κάθε μέλος του παρόντος πληθυσμού**

Κάθε μέλος του πληθυσμού αποκωδικοποιείται δίνοντας την αντίστοιχη ομάδα παραμέτρων οι οποίες τροφοδοτούνται στην συνάρτηση αποτίμησης. Η συνάρτηση αυτή θα αναθέσει ένα «βαθμό» σε κάθε μέλος που να απεικονίζει την «καταλληλότητα» της ομάδας αυτής και χαρακτηρίζει το αντίστοιχο μέλος.

- **Επιλογή ζευγαριών από μέλη για αναπαραγωγή**

Η επιλογή γίνεται με τον κανόνα της ρουλέτας. Μια ρουλέτα μοιράζεται σε έναν αριθμό τμημάτων ίσο με τον αριθμό των μελών του πληθυσμού. Το μέγεθος κάθε τμήματος της ρουλέτας είναι ανάλογο με βαθμό που έλαβε το μέλος του πληθυσμού στο οποίο αντιστοιχεί. Έτσι, όσο πιο υψηλά αποτιμήθηκε ένα μέλος τόσο μεγαλύτερο μερίδιο της ρουλέτας θα καταλάβει και το αντίστροφο.

Αν υποθέσουμε ότι έχουμε N μέλη στον πληθυσμό, τότε γυρνάμε τη ρουλέτα N φορές επιλέγοντας κάθε φορά από ένα μέλος. (Το πείραμα πραγματοποιείται με επαναποθέτηση, δηλαδή υπάρχει η πιθανότητα το ίδιο μέλος να επιλεγθεί δύο ή και περισσότερες φορές.)

- **Εφαρμογή γενετικών τελεστών στα επιλεγθέντα μέλη**

Σε κάθε ζεύγος μελών που επιλέγουμε με την μέθοδο της ρουλέτας:

- εφαρμόζουμε τον τελεστή της διασταύρωσης με κάποια πιθανότητα, έστω p_c καταλήγοντας σε δύο νέες συμβολοσειρές.
- εφαρμόζουμε τον τελεστή της μετάλλαξης σε κάθε μία από τις συμβολοσειρές που δημιουργήθηκαν στο προηγούμενο βήμα με κάποια πιθανότητα, έστω p_m

Οι δύο νέες συμβολοσειρές προστίθενται στον νέο πληθυσμό (νέα γενιά).

Παρατηρούμε ότι τα μέλη που έλαβα υψηλό βαθμό κατά την αποτίμηση έχουν μεγαλύτερη πιθανότητα να επιλεγθούν για διασταύρωση και άρα μεγαλύτερη πιθανότητα να πολλαπλασιαστούν και να επιβιώσουν (είτε αμιγώς είτε σε μερικά τροποποιημένη μορφή) στην επόμενη γενιά.

Παρόλα αυτά, μέλη που έλαβαν χαμηλό βαθμό κατά την αποτίμηση δεν αποκλείονται από την διαδικασία εξέλιξης αφού και αυτά έχουν κάποια (έστω και μικρή) πιθανότητα να επιλεγθούν για διασταύρωση και να περάσουν κάποιο τμήμα του γενετικού υλικού τους στην επόμενη γενιά.

Παράδειγμα Γενετικού Αλγορίθμου

Παράδειγμα: Μεγιστοποίηση Συνάρτησης²

Ας θεωρήσουμε το απλό παράδειγμα της μεγιστοποίησης μιας άγνωστης συνάρτησης. Διατυπωμένο πιο αυστηρά:

Επιθυμούμε να εντοπίσουμε το σημείο $x_{max} \in \mathbb{N}$ όπου μια άγνωστη συνάρτηση $f(x)$ λαμβάνει τη μέγιστη τιμή της. Η μεταβλητή x λαμβάνει τιμές στο διάστημα $\mathbb{N} = [0, 31] \subset \mathbb{N}$ (λαμβάνει δηλαδή μόνο ακέραιες τιμές μεταξύ 0 και 31).

Στο παράδειγμα αυτό θα προδιαγράψουμε έναν γενετικό αλγόριθμο και θα τον εκτελέσουμε «με το χέρι» ώστε να διευκρινιστούν οι διεργασίες που επιτελούνται. Επιπρόσθετα, θα αποκαλύψουμε την άγνωστη συνάρτηση ώστε να έχουμε πλήρη επίγνωση του προβλήματος αλλά προφανώς δεν θα χρησιμοποιήσουμε καθόλου την πληροφορία αυτή σε κανένα σημείο της διαδικασίας επίλυσης παρά μόνο στην συνάρτηση αποτίμησης. Η συνάρτηση αυτή είναι η:

$$f(x) = x^2$$

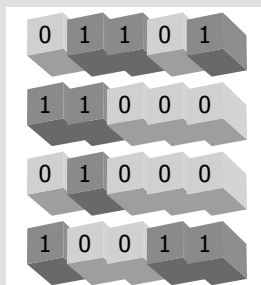
Βήμα 1: Σχήμα κωδικοποίησης

Για την κωδικοποίηση της μοναδικής παραμέτρου της συνάρτησης θα χρησιμοποιήσουμε την δυαδική της μεταγραφή. Ένας ακέραιος στο διάστημα $[0, 31]$ απαιτεί 5 δυαδικά ψηφία για την δυαδική αναπαράστασή του. Συνεπώς, κάθε μέλος του πληθυσμού (συμβολοσειρά) θα αποτελείται από 5 ψηφία.

Βήμα 2: Μέγεθος πληθυσμού, αρχικοποίηση, πιθανότητες τελεστών και συνάρτηση αποτίμησης

Επιλέγουμε για μέγεθος πληθυσμού το 4. Έτσι, σε κάθε γενιά, ο πληθυσμός μας θα έχει 4 μέλη.

Για κάθε δυαδικό ψηφίο κάθε μέλους του πληθυσμού, ρίχνουμε ένα νόμισμα για να το αρχικοποιήσουμε τυχαία στην τιμή 0 ή 1. Έστω ότι καταλήγουμε έτσι στον πληθυσμό του σχήματος 31.



Σχήμα 31. Ο τυχαία αρχικοποιημένος πληθυσμός

Υιοθετούμε τυπικές τιμές για τις πιθανότητες των γενετικών τελεστών διασταύρωσης και μετάλλαξης, για παράδειγμα:

$$p_c = 0.7 \text{ και } p_m = 0.001$$

Η συνάρτηση αποτίμησης αντιστοιχεί στο μαύρο κουτί το οποίο θα τροφοδοτούμε με τις αποκωδικοποιημένες παραμέτρους και θα επιστρέφει ένα μέτρο (βαθμό) καταλληλότητας για αυτές.

Στην περίπτωση μας, η συνάρτηση αποτίμησης είναι η $f(x) = x^2$.

Βήμα 3: Βρόχος εκτέλεσης αλγορίθμου

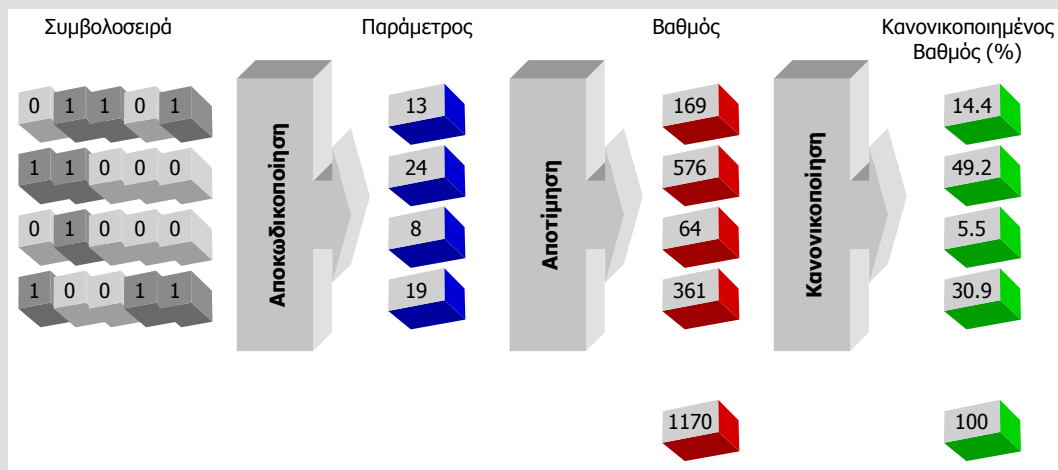
Σε κάθε επανάληψη του βρόχου πραγματοποιούνται τα παρακάτω βήματα:

- Αποκωδικοποίηση κάθε μέλους του πληθυσμού (μετατροπή από δυαδικό σε δεκαδικό)

² Το απλό αυτό παράδειγμα, είναι προσαρμοσμένο από το βιβλίο του Goldberg [GOLDBERG, 1989].

- Αποτίμηση κάθε μέλους του πληθυσμού (χρήση της συνάρτησης f)
- Κανονικοποίηση των βαθμών

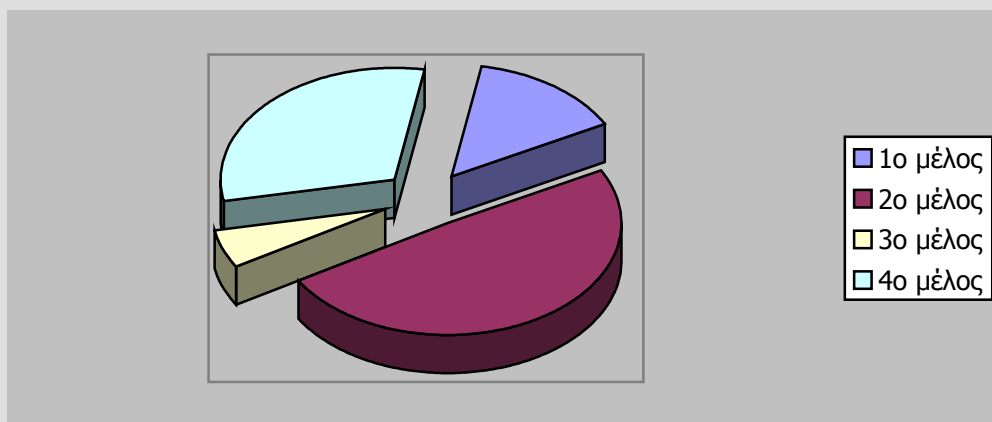
Τα βήματα αυτά απεικονίζονται γραφικά στο σχήμα που ακολουθεί.



Σχήμα 32. Αποκωδικοποίηση και αποτίμηση των μελών του πληθυσμού

Έχοντας υπολογίσει τα μεγέθη αυτά, μπορούμε να προχωρήσουμε στην εφαρμογή των γενετικών τελεστών.

Σχηματίζουμε μια ρουλέτα με τόσες περιοχές όσος είναι ο αριθμός των μελών του πληθυσμού και μεγέθη ανάλογα με το βαθμό του κάθε μέλους.

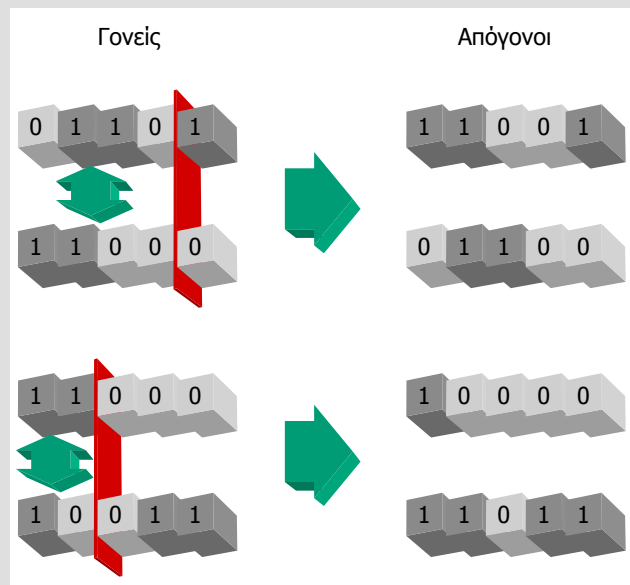


Σχήμα 33. Επιλογή με τον κανόνα της ρουλέτας

Για την επιλογή μελών του πληθυσμού που θα υποβληθούν στην επίδραση των γενετικών τελεστών και θα αποτελέσουν την νέα γενιά, χρησιμοποιούμε την ρουλέτα του παραπάνω σχήματος και πραγματοποιούμε 4 επαναλήψεις. Έστω ότι κατά τις επαναλήψεις επιλέχθηκαν τα ακόλουθα μέλη με την συγκεκριμένη σειρά: 1ο, 2ο, 2ο, 4ο. Έτσι τα δύο ζεύγη που θα διασταυρωθούν είναι: 1ο με 2ο και 2ο με 4ο.

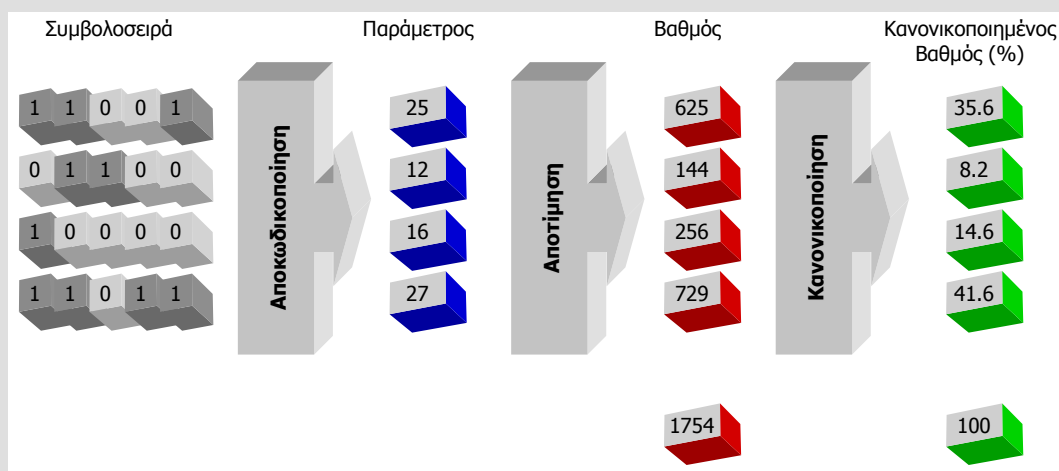
Για το ακριβές σημείο διασταύρωσης, χρησιμοποιούμε και πάλι τυχαία επιλογή (ρίψη νομίσματος) καταλήγοντας για το πρώτο ζευγάρι στην θέση 4 ενώ για το δεύτερο στην θέση 2.

Το αποτέλεσμα της διασταύρωσης μεταξύ των ζευγαριών δίνεται σχηματικά στο σχήμα 34.



Σχήμα 34. Εφαρμογή του τελεστή διασταύρωσης

Στους απογόνους, πρέπει τώρα να εφαρμόσουμε τον τελεστή της μετάλλαξης ψηφίο-ψηφίο. Λόγω της μικρής πιθανότητας που έχει, η μετάλλαξη είναι σπάνιο να συμβεί. Έτσι, στο συγκεκριμένο πείραμα δεν προέκυψε καμία μετάλλαξη. Συνεπώς οι συμβολοσειρές του προηγούμενου σχήματος αποτελούν και τα μέλη της νέας γενιάς του πληθυσμού.



Σχήμα 35. Η νέα γενιά του πληθυσμού

Συγκρίνοντας τον νέο πληθυσμό του σχήματος 35 με τον προηγούμενο του σχήματος 32, εύκολα παρατηρούμε ότι η νέα γενιά είναι βελτιωμένη και ως προς τον μέσο όρο των βαθμών των μελών (436 έναντι 293) αλλά και ως προς το βέλτιστο μέλος του πληθυσμού (βαθμός βέλτιστου μέλους 729 έναντι 576) πράγμα που σημαίνει ότι ο πληθυσμός εξελίχθηκε προς την σωστή κατεύθυνση.

Στο σημείο αυτό κλείνει το βήμα του επαναληπτικού βρόχου και η εκτέλεση συνεχίζεται με τον νέο πληθυσμό από την αρχή του βήματος 3.

Σχόλιο

Αν και το παράδειγμα αυτό είναι αρκετά απλοϊκό, δίνει μια ακριβή ιδέα για τους μηχανισμούς των γενετικών αλγορίθμων και τις μεθόδους εξέλιξης που υιοθετούν. Βασικό χαρακτηριστικό των γενετικών αλγορίθμων είναι ότι με την ίδια απλότητα θα μπορούσαν να αντιμετωπίσουν και προβλήματα πολύ μεγαλύτερης πολυπλοκότητας.

Πίνακας Ορολογίας Γενετικών Αλγορίθμων

Όρος	Μετάφραση	Ερμηνεία
Καταλληλότητα (Βαθμός)	Fitness	Ποσοτικό κριτήριο αξιολόγησης μιας ομάδας παραμέτρων σχετικά με την ποιότητα της λύσης που δίνουν στο πρόβλημα
Παράμετρος	Value (Allele)	
Πληθυσμός	Population	Ένα σύνολο χρωμοσωμάτων
Συμβολοσειρά (Χρωμόσωμα)	String (Chromosome)	Αποτελεί το βασικό δομικό στοιχείο ενός γενετικού αλγορίθμου. Κωδικοποιεί ένα σύνολο παραμέτρων.
Σχήμα	Schema	Μια φόρμα ομοιότητας που περιγράφει ένα υποσύνολο χρωμοσωμάτων με κοινά μέρη κώδικα σε συγκεκριμένα σημεία.
Φαινότυπος	Phenotype	Το αποτέλεσμα της αποκωδικοποίησης μιας ομάδας παραμέτρων

Υβριδικά Μοντέλα

Εισαγωγή

Κάθε μία από τις ξεχωριστές μεθοδολογίες της Υπολογιστικής Νοημοσύνης που παρουσιάστηκαν στα προηγούμενα (ασαφής συστήματα, νευρωνικά δίκτυα και γενετικοί αλγόριθμοι) παρουσιάζουν ιδιαίτερα χαρακτηριστικά και διαφορετικά ισχυρά και ασθενή σημεία.

Σκοπός των **υβριδικών μοντέλων** (hybrid models) είναι να συνδυάσουν δύο (ή, λιγότερο συχνά, περισσότερες) από αυτές τις μεθοδολογίες για την ανάπτυξη συστημάτων με μικτά χαρακτηριστικά τα οποία:

- θα παρουσιάζουν **συνδυαστικά πλεονεκτήματα**
- θα **εξαλείφουν ελαττώματα** των επιμέρους μεθοδολογιών, ή
- θα προσφέρουν ένα πιο **κατάλληλο πλαίσιο** διατύπωσης και αποτελεσματικότερης επίλυσης συγκεκριμένων κατηγοριών προβλημάτων

Αν και η βιβλιογραφία είναι πλούσια από παραδείγματα υβριδικών συστημάτων που συνδυάζουν κάποια από τις μεθοδολογίες της Υπολογιστικής Νοημοσύνης με παραδοσιακούς αλγορίθμους ή τεχνικές, η παρακάτω παρουσίαση θα περιοριστεί σε υβριδικά συστήματα που προκύπτουν από το συνδυασμό αποκλειστικά μεθοδολογιών της ΥΝ.

Οι ακόλουθες αράγραφοι έχουν περισσότερο την μορφή καταγραφής παρά αναλυτικής παρουσίασης.

Μοντέλα

Στην προσπάθεια σχεδίασης υβριδικών συστημάτων από συστατικά της ΥΝ, οι πιθανές επιλογές καθορίζονται από το παρακάτω διάγραμμα:



Σχήμα 36. Πλαίσιο υβριδικών συστημάτων

Το παραπάνω σχήμα, προσφέρει και κάποια σχετική πληροφορία αναφορικά με τον «όγκο» της έρευνας στις αντίστοιχες περιοχές. Έτσι, είναι βέβαιο ότι τα νευροασαφή συστήματα (τα συστήματα, δηλαδή, που συνδυάζουν μεθοδολογίες ασαφούς λογικής και νευρωνικών δικτύων) έχουν προσελκύσει περισσότερη έρευνα από εκείνα που συνδυάζουν ασαφή συστήματα και γενετικούς αλγορίθμους κ.λπ.

Ένα σημείο που θα πρέπει επίσης να σημειωθεί, είναι ότι ο συνδυασμός δύο μεθοδολογιών προσφέρει διάφορες εναλλακτικές. Έτσι, όπως θα φανεί στη συνέχεια, ο όρος «νευροασαφή» μπορεί να υπονοεί συστήματα που ενσωματώνουν νευρωνικές τεχνικές (όπως ένας κανόνας εκπαίδευσης) σε ένα ασαφές σύστημα ή ασαφή εργαλεία (π.χ. ένας ασαφής αλγόριθμος για την δυναμική μεταβολή του ρυθμού μάθησης) σε ένα νευρωνικό δίκτυο. Το ευρύτερο πλαίσιο, δηλαδή, μπορεί να είναι ασαφές όπου θα εισαχθούν νευρωνικά στοιχεία ή νευρωνικό όπου θα εισαχθούν ασαφή στοιχεία.

Νευροασαφή Μοντέλα

Ακολουθεί ένας μη εξαντλητικός κατάλογος διαφόρων υβριδικών μοντέλων που έχουν κατά καιρούς προταθεί στη διεθνή βιβλιογραφία για τον συνδυασμό ασαφών τεχνικών με νευρωνικές:



- **Ασαφή Νευρωνικά Δίκτυα**

(Fuzzy Neural Networks)

Αντιστοίχιση των σταδίων εξαγωγής ασαφούς συμπεράσματος στα στρώματα ενός νευρωνικού. Χρησιμοποιούνται διαφορετικοί ειδικά σχεδιασμένοι νευρώνες σε κάθε στρώμα.

- **ANFIS - Adaptive Neurofuzzy Inference Systems**

(Προσαρμοστικό νευροασαφές σύστημα εξαγωγής συμπεράσματος)

Εκπαίδευση ασαφούς συστήματος με τον αλγόριθμο ανάστροφης διάδοσης σφάλματος.

- **FALCON - Fuzzy Adaptive Learning Control Net**

(Ασαφές προσαρμοστικό δίκτυο ελέγχου)

Νευρωνική αναπαράσταση ασαφούς εξαγωγής συμπεράσματος.

- **Νευρωνικές συσχετιστικές μνήμες για την εκμάθηση ασαφών κανόνων**

(Neural associative memories for learning fuzzy rules)

Χρήση νευρωνικών δικτύων για την αντιγραφή της συμπεριφοράς ασαφών κανόνων.

- **Ασαφής έλεγχος του ρυθμού μάθησης σε νευρωνικά δίκτυα**

Ασαφή-Γενετικά Μοντέλα

Ακολουθεί ένας μη εξαντλητικός κατάλογος διαφόρων μεθόδων που έχουν κατά καιρούς προταθεί στη διεθνή βιβλιογραφία για τον συνδυασμό ασαφών τεχνικών με γενετικούς αλγόριθμους:



- **Αυτόματη βέλτιστη διαμέριση πεδίων εισόδων και εξόδων ασαφούς συστήματος με ΓΑ**

- **Αυτόματη εξαγωγή ασαφών κανόνων από δεδομένα με ΓΑ**

- **Γενετική βελτιστοποίηση ασαφών κανόνων**

- **Γενετική επιλογή ασαφών κανόνων**

- **Ασαφής έλεγχος παραμέτρων ΓΑ**

- **Ασαφείς Γενετικοί Αλγόριθμοι**

Νευρωνικά-Γενετικά Μοντέλα

Ακολουθεί ένας μη εξαντλητικός κατάλογος διαφόρων μεθόδων που έχουν κατά καιρούς προταθεί στη διεθνή βιβλιογραφία για τον συνδυασμό νευρωνικών τεχνικών με γενετικούς αλγόριθμους:

- **Γενετική εκπαίδευση νευρωνικού δικτύου**

Προσδιορισμός βαρών νευρωνικού δικτύου προκαθορισμένης (και σταθερής) δομής με ΓΑ.

- **Γενετική σχεδίαση νευρωνικού δικτύου**

Γενετικοί αλγόριθμοι για:

- Αυτόματο προσδιορισμό της δομής ενός νευρωνικού δικτύου για συγκεκριμένο πρόβλημα
- Προσδιορισμό «βέλτιστου» νευρωνικού δικτύου για συγκεκριμένο πρόβλημα.



Άλλα Μοντέλα

- **Γενετικός Μετα-Αλγόριθμος**

Γενετικός αλγόριθμος για τον προσδιορισμό «βέλτιστων» τιμών για τις παραμέτρους ενός «φωλιασμένου» γενετικού αλγορίθμου.



Βραχυγραφίες

Βραχυγραφία	Ανάπτυγμα
ΓΑ	Γενετικός Αλγόριθμος (Genetic Algorithm, GA)
ΤΝΔ	Τεχνητό Νευρωνικό Δίκτυο (Artificial Neural Network, ANN)
ΤΝ	Τεχνητή Νοημοσύνη (Artificial Intelligence, AI)
ΥΝ	Υπολογιστική Νοημοσύνη (Computational Intelligence, CI)
LMS	Least Mean Squares (Ελάχιστα Μέσα Τετράγωνα)

Προτεινόμενη Βιβλιογραφία

Ακολουθεί ένας περιορισμένος κατάλογος εισαγωγικών τίτλων οργανωμένος κατά θέμα. Βασικά βιβλία που καλύπτουν περισσότερα του ενός αντικείμενα επαναλαμβάνονται στις αντίστοιχες παραγράφους.

Ο αναγνώστης μπορεί να ανατρέξει στους προτεινόμενους τίτλους για (εκτεταμένους) καταλόγους πιο εξειδικευμένων άρθρων και βιβλίων.

Ασαφής Λογική

[KLIR & YUAN, 1995] G. J. KLIR and B. YUAN, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice-Hall, New Jersey (1995)

[LIN & LEE, 1996] C. -T. LIN and C. S. G. LEE, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice-Hall (1996)

[KOSKO, 1992A] B. KOSKO, *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ (1992)

[ROSS, 1995] T. J. ROSS, *Fuzzy Logic with Engineering Applications*, McGraw-Hill (1995)

Νευρωνικά Δίκτυα

[HAYKIN, 1994] S. HAYKIN, *Neural Networks: A Comprehensive Foundation*, Macmillan, NY (1994)

[LIN & LEE, 1996] C. -T. LIN and C. S. G. LEE, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice-Hall (1996)

[KOSKO, 1992A] B. KOSKO, *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ (1992)

[KOSKO, 1992B] B. KOSKO, *Neural Networks for Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ (1992)

Γενετικοί Αλγόριθμοι

[GOLDBERG, 1989] D. E. GOLDBERG, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Reading, MA: Addison-Wesley (1989)

[FOGEL, 1995] D. B. FOGEL, *Evolutionary Computation: Towards a New Philosophy of Machine Intelligence*, Piscataway, NJ: IEEE Press (1995)

[HOLLAND, 1975] J. H. HOLLAND, *Adaptation in Natural and Artificial Systems*, Ann Arbor: MI: Univ. of Michigan Press (1975)

Υβριδικές Μέθοδοι

[LIN & LEE, 1996] C. -T. LIN and C. S. G. LEE, *Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice-Hall (1996)

[BROWN & HARRIS, 1994] M. BROWN and C. HARRIS, *Neurofuzzy Adaptive Modelling and Control*, Prentice-Hall, UK (1994)

Άλλα Θέματα

[MINSKY, 1986] M. L. MINSKY, *The Society of Mind*, Simon & Schuster, New York (1986)

[MINSKY, 1995] M. L. MINSKY, *Η Κοινωνία της Νόησης*, Κάτοπτρο, Αθήνα (1995)
[Μετάφραση: Μ. Αντωνοπούλου και Σ. Μανουσέλης]

Κατάλογος Σχημάτων και Πινάκων

Σχήμα 1. Αφαιρετική θεώρηση ενός συστήματος	5
Σχήμα 2. Σύστημα βασισμένο σε μαθηματικά μοντέλα.....	5
Σχήμα 3. Ένα τυπικό σύστημα Τεχνητής Νοημοσύνης.....	6
Σχήμα 4. Αφαιρετική αναπαράσταση ενός τυπικού συστήματος ΥΝ.....	8
Σχήμα 5. Αφαιρετική αναπαράσταση ενός ασαφούς συστήματος	9
Σχήμα 6. Αναπαράσταση ενός νευρωνικού δικτύου	9
Σχήμα 7. Αφαιρετική αναπαράσταση γενετικών αλγορίθμων	10
Σχήμα 8. Παράδειγμα (α) κλασσικού και (β) ασαφούς συνόλου	13
Σχήμα 9. Πράξεις ασαφών συνόλων. (α) Συμπλήρωμα, (β) τομή και (γ) ένωση.....	15
Σχήμα 10. Ορισμός γλωσσικών όρων και γλωσσικών μεταβλητών	16
Σχήμα 11. Ένα τυπικό ασαφές σύστημα	17
Σχήμα 12. Ορισμός γλωσσικών όρων παραδείγματος	18
Σχήμα 13. Οι κανόνες που χρησιμοποιούνται από το σύστημα του παραδείγματος.....	19
Σχήμα 14. Η δομή του νευρώνα.....	21
Πίνακας. Σύγκριση βιολογικών νευρωνικών δικτύων με σύγχρονους υπολογιστές	22
Σχήμα 15. Το μοντέλο του τεχνητού νευρώνα.....	24
Σχήμα 16. Συνήθεις συναρτήσεις δραστηριοποίησης	24
Σχήμα 17. Δίκτυο προσοτροφοδότησης ενός στρώματος.....	25
Σχήμα 18. Δίκτυο προσοτροφοδότησης δύο στρωμάτων	26
Σχήμα 19. Δίκτυο με ανατροφοδότηση χωρίς κρυμμένους νευρώνες	26
Σχήμα 20. Δίκτυο ανατροφοδότησης με κρυμμένους νευρώνες.....	27
Σχήμα 21. Ένας τυπικός νευρώνας τύπου Perceptron.....	27
Σχήμα 22. Τα προς ταξινόμηση σημεία του επιπέδου	28
Σχήμα 23. Το πρόβλημα της αποκλειστικής διάζευξης (XOR).....	31
Σχήμα 24. Βαθμωτή κάθοδος με υψηλό ρυθμό μάθησης	32
Σχήμα 25. Νευρωνική πρόβλεψη.....	33
Σχήμα 26. Τοπικά ελάχιστα της επιφάνειας σφάλματος	34
Σχήμα 27. Παράδειγμα υπερ-παραμετροποίησης.....	35
Σχήμα 28. Το πρόβλημα του «μαύρου κουτιού»	39
Σχήμα 29. Ο γενετικός τελεστής διασταύρωσης.....	41
Σχήμα 30. Ο γενετικός τελεστής μετάλλαξης.....	41
Σχήμα 31. Ο τυχαία αρχικοποιημένος πληθυσμός.....	43
Σχήμα 32. Αποκωδικοποίηση και αποτίμηση των μελών του πληθυσμού	44
Σχήμα 33. Επιλογή με τον κανόνα της ρουλέτας.....	44
Σχήμα 34. Εφαρμογή του τελεστή διασταύρωσης.....	45
Σχήμα 35. Η νέα γενιά του πληθυσμού	45
Σχήμα 36. Πλαίσιο υβριδικών συστημάτων	47