

ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΠΕΙΡΑΙΑ
ΤΜΗΜΑ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

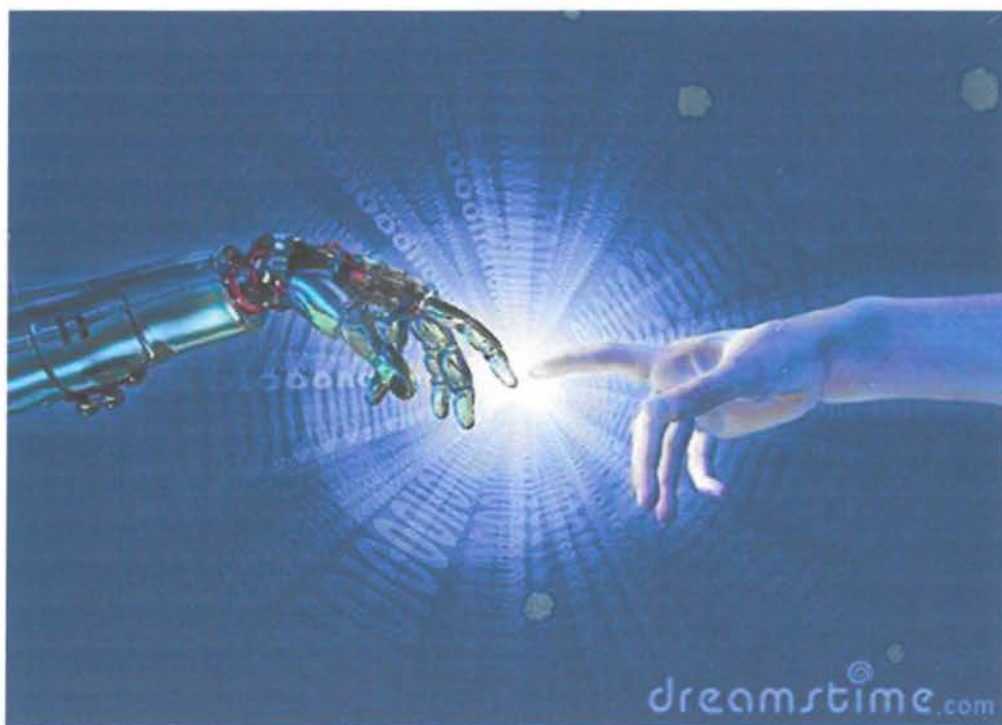


<<Τεχνητή Αίσθηση και Εφαρμογές>>

ΠΤΥΧΙΑΚΗ ΜΕΛΕΤΗ

ΧΑΡΑΛΑΜΠΟΠΟΥΛΟΣ ΑΛΕΞΑΝΔΡΟΣ 30442

2014



ΛΗΠΗ Ο ΚΑΘΗΓΗΤΗΣ
ΓΙΑΝΝΑΚΟΠΟΥΛΟΣ

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΠΕΙΡΑΙΑ
ΤΜΗΜΑ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ**



<<Τεχνητή Αίσθηση και Εφαρμογές>>

ΠΤΥΧΙΑΚΗ ΜΕΛΕΤΗ

ΧΑΡΑΛΑΜΠΟΠΟΥΛΟΣ ΑΛΕΞΑΝΔΡΟΣ 30442

2014

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

Περιεχόμενα

Εισαγωγή	5
1. ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ	7
1.1 ΕΙΣΑΓΩΓΗ	7
1.1.1 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΥΠΟΛΟΓΙΣΜΟΥ	7
1.1.2 ΟΡΙΣΜΟΣ ΤΟΥ ΤΕΧΝΗΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ	8
1.1.3 ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	9
1.1.4 Ο ΒΙΟΛΟΓΙΚΟΣ ΚΑΙ Ο ΤΕΧΝΗΤΟΣ ΝΕΥΡΩΝΑΣ	11
1.1.5 ΤΥΠΟΙ ΣΥΝΑΡΤΗΣΕΩΝ ΕΝΕΡΓΟΠΟΙΗΣΗΣ	15
1.1.6 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	19
1.1.7 ΕΦΑΡΜΟΓΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	24
1.1.8 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΚΑΙ ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	26
1.2 ΙΣΤΟΡΙΚΕΣ ΣΗΜΕΙΩΣΕΙΣ	28
1.3 ΑΝΑΠΑΡΑΣΤΑΣΗ ΓΝΩΣΗΣ	32
1.4 ΜΑΘΗΣΗ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	38
1.4.1 ΜΑΘΗΣΗ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ	39
1.4.2 ΠΑΡΑΔΕΙΓΜΑΤΑ ΜΑΘΗΣΗΣ	42
1.5. ΠΟΛΥΣΤΡΩΜΑΤΙΚΟ PERCEPTRON (MLP)	46
1.5.1 ΤΟ ΑΠΛΟ PERCEPTRON	46
1.5.2 ΤΟ ΠΟΛΥΣΤΡΩΜΑΤΙΚΟ PERCEPTRON	53
1.5.3 Ο ΑΛΓΟΡΙΘΜΟΣ ΑΝΑΣΤΡΟΦΗΣ ΔΙΑΔΟΣΗΣ	58
1.5.4 ΤΡΟΠΟΙ ΕΚΠΑΙΔΕΥΣΗΣ	61
1.5.5 ΕΦΑΡΜΟΓΕΣ – ΔΥΝΑΤΟΤΗΤΕΣ ΤΟΥ ΠΟΛΥΣΤΡΩΜΑΤΙΚΟΥ PERCEPTRON	62
1.6 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΚΑΙ ΜΑΤΛΑΒ	65
1.6.1 ΤΡΟΠΟΣ ΠΑΡΑΣΤΑΣΗΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	66
1.6.2 ΔΗΜΙΟΥΡΓΙΑ, ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΠΡΟΣΟΜΟΙΩΣΗ ΤΟΥ ΔΙΚΤΥΟΥ	70
1.6.3 ΑΛΓΟΡΙΘΜΟΙ ΕΚΠΑΙΔΕΥΣΗΣ	73
2. ΤΕΧΝΗΤΗ ΑΙΣΘΗΣΗ ΚΑΙ ΕΦΑΡΜΟΓΕΣ	77

2.1 ΕΙΣΑΓΩΓΗ	77
2.2 ΑΝΙΧΝΕΥΣΗ ΑΚΜΩΝ	77
2.4 ΑΝΑΛΥΣΗ ΣΧΗΜΑΤΟΣ-ΑΛΓΟΡΙΘΜΟΣ CHAIN CODE.....	97
2.5 ΑΝΑΛΥΣΗ ΣΧΗΜΑΤΟΣ-ΚΕΝΤΡΙΚΕΣ ΚΑΙ ΑΝΑΛΟΙΟΤΕΣ ΡΟΠΕΣ.....	105
2.6 ΧΑΡΤΗΣ ΒΑΘΟΥΣ-ΧΑΡΤΗΣ ΒΕΛΟΝΩΝ	116
2.7 ΕΥΡΕΣΗ ΑΝΤΙΣΤΟΙΧΩΝ ΣΗΜΕΙΩΝ ΣΕ ΖΕΥΓΗ ΣΤΕΡΕΟΣΚΟΠΙΚΩΝ ΕΙΚΟΝΩΝ (ΑΛΓΟΡΙΘΜΟΣ ΤΟΥ SHIRAI).....	121
2.8 ΛΟΓΙΣΜΙΚΟ ΑΝΑΓΝΩΡΙΣΗΣ ΦΩΝΗΣ.....	131
3.ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ	138
3.1 ΕΜΠΕΙΡΑ ΣΥΣΤΗΜΑΤΑ ΣΤΗΝ ΙΑΤΡΙΚΗ	141
3.1.1 ΓΕΝΙΚΑ	141
3.1.2 ΧΡΗΣΗ ΕΜΠΕΙΡΩΝ ΣΥΣΤΗΜΑΤΩΝ ΣΤΗΝ ΙΑΤΡΙΚΗ.....	142
3.1.3 ΑΝΑΣΚΟΠΙΣΗ ΙΑΤΡΙΚΩΝ ΕΜΠΕΙΡΩΝ ΣΥΣΤΗΜΑΤΩΝ	143
3.2 ΥΒΡΙΔΙΚΟ ΕΜΠΕΙΡΟ ΣΥΣΤΗΜΑ.....	150
3.2.1 ΓΕΝΙΚΑ	150
3.2.2 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ ΥΒΡΙΔΙΚΟΥ ΕΜΠΕΙΡΟΥ ΣΥΣΤΗΜΑΤΟΣ	151
3.2.3 ΠΑΡΑΔΕΙΓΜΑΤΑ ΚΑΝΟΝΩΝ.....	153
4.ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΣΤΗΝ ΧΕΙΡΟΥΡΓΙΚΗ ΡΟΜΠΟΤΙΚΗ.....	156
4.1 ΜΗΧΑΝΗΜΑΤΑ ΡΟΜΠΟΤΙΚΗΣ ΧΕΙΡΟΥΡΓΙΚΗΣ.....	157
4.1.1 ΤΟ ΧΕΙΡΟΥΡΓΙΚΟ ΣΥΣΤΗΜΑ DAVINCI	158
4.1.2 ΣΥΣΤΗΜΑΤΑ ΥΠΟΒΟΗΘΗΣΗΣ ΧΕΙΡΟΥΡΓΙΚΩΝ ΕΠΕΜΒΑΣΕΩΝ	159
4.1.3 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΓΙΑ ΤΟΝ ΑΣΘΕΝΗ ΚΑΙ ΤΟΝ ΧΕΙΡΟΥΡΓΟ	160
4.2 ΡΟΜΠΟΤΙΚΗ ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ	160
4.2.1 Η ΡΟΜΠΟΤΙΚΗ ΑΛΛΑΖΕΙ ΤΟΝ ΠΑΡΑΔΟΣΙΑΚΟ ΤΡΟΠΟ ΔΙΔΑΣΚΑΛΙΑΣ	161
4.2.2 ΜΕΙΟΝΕΚΤΗΜΑΤΑ.....	163
4.3 Η ΡΟΜΠΟΤΙΚΗ ΣΤΟΝ ΣΤΡΑΤΟ	164
4.4 ΡΟΜΠΟΤΙΚΗ ΣΤΟ ΔΙΑΣΤΗΜΑ.....	169
5 ΕΥΦΥΕΙΣ ΕΙΚΟΝΙΚΑ ΠΕΡΙΒΑΛΛΟΝΤΑ.....	170
5.1 ΕΙΣΑΓΩΓΗ	172
5.1.1 ΚΑΤΗΓΟΡΙΕΣ	173
5.1.2 ΤΑΞΙΝΟΜΗΣΗ.....	174
5.2 ΑΝΑΠΤΥΞΗ ΕΝΟΣ ΕΙΚΟΝΙΚΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ	175

5.2.1 ΜΟΝΤΕΛΟ ΠΕΡΙΒΑΛΛΟΝΤΟΣ	176
5.2.2 ΑΛΛΗΛΕΠΙΔΡΑΣΕΙΣ ΜΕ ΤΟΝ ΧΡΗΣΤΗ.....	179
4.3 ΤΕΧΝΟΛΟΓΙΕΣ	181
5.3.1 ΥΛΙΚΟ	181
5.3.2 ΛΟΓΙΣΜΙΚΟ.....	184
5.4 ΕΦΑΡΜΟΓΕΣ	187
5.5 ΕΙΚΟΝΙΚΟΙ ΠΡΑΚΤΟΡΕΣ	190
5.5.1 ΕΙΣΑΓΩΓΗ	190
5.5.1 ΤΑΞΙΝΟΜΗΣΗ ΕΙΚΟΝΙΚΩΝ ΠΡΑΚΤΟΡΩΝ	191
5.5.2 ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΟΥ ΣΩΜΑΤΟΣ.....	195
5.5.3 ΑΛΛΗΛΕΠΙΔΡΑΣΗ ΜΕ ΤΟ ΠΕΡΙΒΑΛΛΟΝ.....	203
5.5.4 ΟΛΟΚΛΗΡΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ.....	212
6 ΒΙΟΝΙΚΑ ΤΕΧΝΗΤΑ ΜΕΛΗ	218
Βιβλιογραφία	228

Εισαγωγή

Η προσπάθεια κατασκευής ευφυών μηχανών που διαθέτουν την ικανότητα να κατανοούν τον κόσμο αποτελεί πάντα μια ενδιαφέρουσα ενασχόληση των επιστημόνων πληροφορικής. Ενα μεγάλο μέρος του ανθρωπίνου εγκεφάλου είναι αφιερωμένο στη λειτουργία των αισθήσεων. Οι άνθρωποι λύνουν τα προβλήματα αίσθησης χωρίς προσπάθεια, παρ' όλο που λίγοι κατανοούν το μηχανισμό της. Ο Alan Turing, ένας από τους πατέρες των Ψηφιακών Υπολογιστών και του πεδίου της Τεχνητής Νοημοσύνης, πίστευε ότι είναι δυνατό ο ψηφιακός υπολογιστής να αποκτήσει ευφυία και την ικανότητα να κατανοεί το περιβάλλον. Αν και αυτός ο στόχος δεν έχει επιτευχθεί ακόμα σε βαθμό συγκρίσιμο με του ανθρώπου, έχει επιτευχθεί εκπληκτική πρόοδος.

Τα νευρωνικά δίκτυα είναι γνωστικό αντικείμενο της πληροφορικής δανεισμένο από τον πραγματικό εγκέφαλο. Πρόκειται για μαθηματικά μοντέλα τα οποία προσομοιώνουν κάποιες από τις αρχές λειτουργίας των πραγματικών νευρωνικών κυκλωμάτων. Τα νευρωνικά δίκτυα έφεραν στην πληροφορική μια «αναλογική αρχιτεκτονική» η οποία είναι εντελώς διαφορετική από την παραδοσιακή «ψηφιακή αρχιτεκτονική». Το σημαντικό με αυτή την διαφορετική προσέγγιση, είναι ότι λύθηκαν προβλήματα στα οποία οι κλασσικές μέθοδοι προγραμματισμού και υλοποίησης δεν μπορούσαν να ανταποκριθούν.

Ο κλάδος της Ιατρικής Τεχνολογίας αφορά οποιαδήποτε εφαρμογή της τεχνολογίας που σχετίζεται με την ανάπτυξη, διάθεση και χρήση των ιατροτεχνολογικών προϊόντων. Ο όρος «ιατροτεχνολογικό προϊόν» καλύπτει ένα ευρύ φάσμα προϊόντων, όπως τα υλικά μιας χρήσεως, τα εμφυτεύσιμα, τα αντιδραστήρια για διάγνωση in-vitro μέχρι και τα ιατρικά μηχανήματα υψηλού κόστους. Σήμερα, οι νέες τεχνολογίες προσφέρουν στους ασθενείς μια σειρά από οφέλη, όπως έγκαιρη διάγνωση, ταχύτερη ανάρρωση και αποθεραπεία, πολύ πιο σύντομη νοσηλεία και κατά συνέπεια γρήγορη επιστροφή στην καθημερινότητά τους. Εκτός όμως από την καλύτερη ποιότητα ζωής των ασθενών, τα οφέλη από τη χρήση των επιτευγμάτων της ιατρικής τεχνολογίας καθιστούν αποδοτικότερο το σύστημα υγείας, καθότι ο χρόνος παραμονής των ασθενών στο νοσοκομείο μειώνεται ενώ συγχρόνως επιτυγχάνεται και μικρότερη απασχόληση του ιατρονοσηλευτικού προσωπικού. Ευρύτερα, η υιοθέτηση της χρήσης της σύγχρονης ιατρικής τεχνολογίας έχει και κοινωνικοοικονομικές προεκτάσεις όπως η αύξηση της παραγωγικότητας και ιδιαίτερα του προσδόκιμου επιβίωσης.

Η τεχνολογία αναπτύσσεται σε γοργούς ρυθμούς και η ρομποτική τεχνολογία έχει μπει για τα καλά στον χώρο της ιατρικής. Πάρα πολλές εγχειρίσεις πραγματοποιούνται με την χρήση ρομπότ υψηλής τεχνολογίας, αλλά το πιο συναρπαστικό και εξελιγμένο στην ρομποτική είναι η αντικατάσταση ακρωτηριασμένων μελών από ρομποτικά μέλη. Αυτή η τεχνολογία διαθέτει άποιρες δυνατότητες ώστε να δημιουργεί τα λεγόμενα «βιονικά» μέλη. Ο κάτοχος του δεν καταλαβαίνει τις διαφορές από το να χρησιμοποιεί το δικό του μέλος. Ενώ το βιονικό

μέλος του εκτελεί όλες τις λειτουργίες που ο άνθρωπος ορίζει. Κατεστραμένα μέλη που αντικαθιστούνται με τεχνητά , αντικαθιστούν τις αισθήσεις , διαγείροντας με μικροτσιπ τον εγκέφαλο, τυφλοί που καταφέρνουν να δουν, κουφοί που ακούν, τεχνητό δέρμα κ.λ.π. Τα καινούργια μέλη είναι δυνατότερα, καλύτερα και άφθαρτα.

Εικονικά περιβάλλοντα ονομάζονται τα τρισδιάστατα περιβάλλοντα προσομοίωσης σε υπολογιστή των οποίων η απεικόνιση γίνεται σε πραγματικό χρόνο και εξαρτάται από τη συμπεριφορά του χρήστη. Όταν τα περιβάλλοντα αυτά εμπλουτίζονται με τεχνικές από το χώρο της Τεχνητής Νοημοσύνης και των Ευφυών Πρακτόρων με σκοπό να αυξηθεί η αυτονομία και αληθοφάνειά τους, καλούνται Ευφυή Εικονικά Περιβάλλοντα.

Ας προχωρήσουμε λοιπόν αναλύοντας όλα τα παραπάνω.....

1.ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

1.1 ΕΙΣΑΓΩΓΗ

1.1.1 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΥΠΟΛΟΓΙΣΜΟΥ

Η έννοια του υπολογισμού έχει πολλές μορφές. Ιστορικά, ο υπολογισμός κυριαρχήθηκε από την αντίληψη του προγραμματισμού, στον οποίο (κυρίως στον διαδικαστικό) σχεδιάζονταν και υλοποιούνταν ακολουθιακοί αλγόριθμοι. Όμως οι υπολογισμοί που είναι απαραίτητοι στα βιολογικά συστήματα απαιτούν έναν εναλλακτικό τρόπο αντιμετώπισης. Για παράδειγμα οι υπολογισμοί που εκτελεί ο ανθρώπινος εγκέφαλος διαφέρουν στο ότι:

- Οι υπολογισμοί είναι κατανεμημένοι και παράλληλοι.
- Η ανάπτυξη ενός προγράμματος αντικαθίσταται από τη διαδικασία μάθησης.

Χρησιμοποιώντας το ανάλογο της φύσης, αναπτύχθηκε το εμπνεόμενο από τη βιολογία υπολογιστικό παράδειγμα των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ). Χρησιμοποιείται ο όρος «τεχνητό» γιατί οι υπολογισμοί υλοποιούνται με τη χρήση ηλεκτρονικών και όχι βιολογικών στοιχείων. Η διαρκώς αναπτυσσόμενη τεχνολογία των ΤΝΔ αποτελεί ένα ευρύ σύνολο τεχνικών, το οποίο προσφέρει μια πρακτική, εναλλακτική των συμβατικών, προσέγγιση και σε πολλές περιπτώσεις παρέχει τη δυνατότητα αντιμετώπισης δύσκολων προβλημάτων. Η εξάπλωση των ΤΝΔ ευνοήθηκε σημαντικά από τις εξελίξεις στο πεδίο του υλικού των υπολογιστών (hardware) και ειδικά από την ανάπτυξη παράλληλων συστημάτων ανεκτού κόστους. Σήμερα έχουν αναμφίβολα φθάσει σε ένα θαυμάσιο επίπεδο ανάπτυξης και βεβαίως θα συνεχίσουν να αναπτύσσονται στο μέλλον ποικιλοτρόπως συνδυαζόμενα με τα ασαφή συστήματα,

τους γενετικούς αλγορίθμους και τα έμπειρα συστήματα, αλλά και με άλλες κλασικές τεχνικές ανάλυσης και σχεδίασης συστημάτων απόφασης και ελέγχου.

1.1.2 ΟΡΙΣΜΟΣ ΤΟΥ ΤΕΧΝΗΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ

Στη συνέχεια παρατίθεται ένας γενικός και πρακτικός ορισμός σχετικά με το τι είναι ένα ΤΝΔ..

ΤΕΧΝΗΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ: Μια αρχιτεκτονική δομή (δίκτυο) αποτελούμενη από ένα πλήθος διασυνδεδεμένων μονάδων (τεχνητοί νευρώνες). Κάθε μονάδα χαρακτηρίζεται από εισόδους και εξόδους και υλοποιεί τοπικά έναν απλό υπολογισμό. Κάθε σύνδεση μεταξύ δυο μονάδων χαρακτηρίζεται από μια τιμή βάρους. Οι τιμές των βαρών των συνδέσεων αποτελούν τη γνώση που είναι αποθηκευμένη στο δίκτυο και καθορίζουν τη λειτουργικότητά του. Η έξοδος κάθε μονάδας καθορίζεται από τον τύπο της μονάδας, τη διασύνδεση με τις υπόλοιπες μονάδες και πιθανώς κάποιες εξωτερικές εισόδους. Πέρα από μια πιθανή δεδομένη (εκ κατασκευής) λειτουργική ικανότητα ενός δικτύου, συνήθως ένα δίκτυο αναπτύσσει μια συνολική λειτουργικότητα μέσω μιας μορφής εκπαίδευσης.

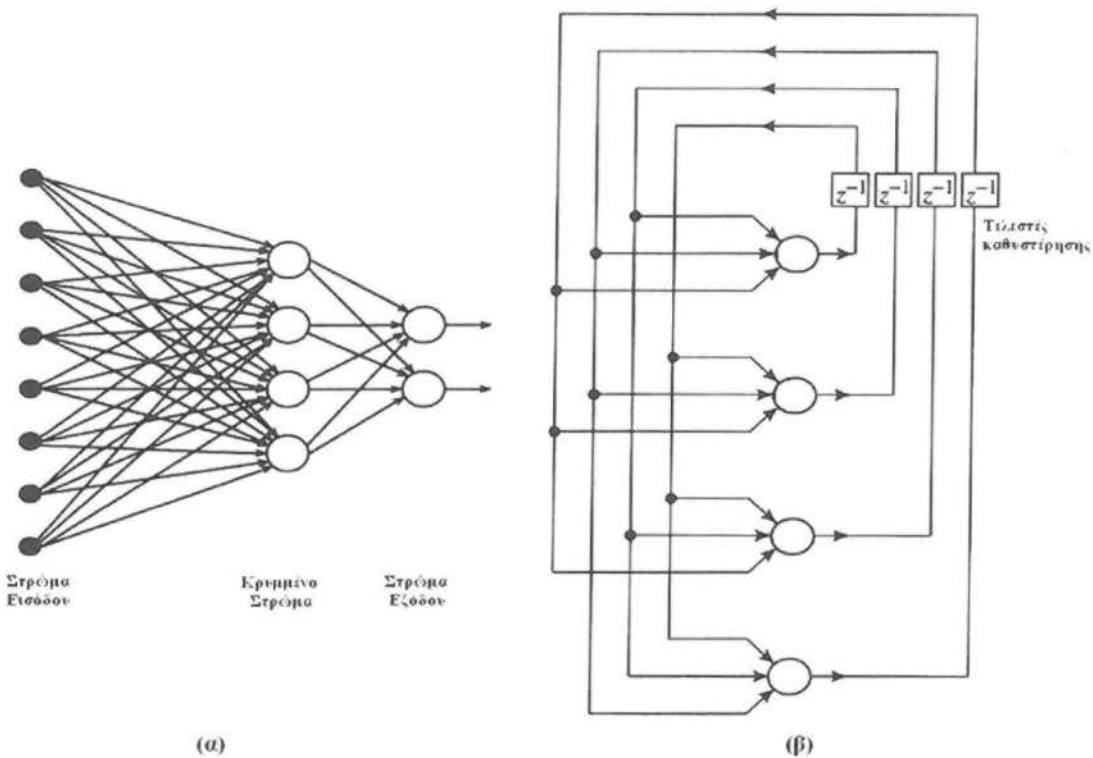
Η συνολική λειτουργικότητα ενός ΤΝΔ καθορίζεται από την τοπολογία του δικτύου, τα χαρακτηριστικά των νευρώνων, τη μέθοδο εκπαίδευσης και από τα δεδομένα με τα οποία γίνεται η εκπαίδευση. Παρότι δεν το απαιτεί ο προηγούμενος ορισμός, ο υπολογισμός που εκτελεί κάθε νευρώνας είναι απλός και κοινός για όλους τους νευρώνες. Επειδή οι νευρώνες λειτουργούν παράλληλα (ταυτόχρονα) και ο αριθμός τους μπορεί να είναι πολύ μεγάλος, τα ΤΝΔ αποτελούν χαρακτηριστικό παράδειγμα μαζικά παράλληλου υπολογισμού (massively parallel computing).

1.1.3 ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

• Όπως προκύπτει από τον ορισμό του ΤΝΔ, το υπολογιστικό μοντέλο αποτελείται από συνδέσεις απλών στοιχείων ή μονάδων (νευρώνων). Οι συνδέσεις αυτές χαρακτηρίζονται από βάρη. Στο σχήμα 1.1 απεικονίζονται δυο παραδείγματα δικτύων μικρής κλίμακας στα οποία οι νευρώνες παρουσιάζονται με κύκλους και οι συνδέσεις με βέλη. Στο σχήμα 1.1(α) απεικονίζεται ένα μη-ανατροφοδοτούμενο δίκτυο, το οποίο δηλαδή δεν περιέχει κλειστά μονοπάτια από συνδέσεις. Όπως φαίνεται οι νευρώνες ομαδοποιούνται σε επίπεδα (ή στρώματα) (layers). Αντίθετα στο σχήμα 1.1(β) παρουσιάζεται η αρχιτεκτονική της ανατροφοδότησης η οποία επιτρέπει την ύπαρξη κύκλων από συνδέσεις μεταξύ νευρώνων. Αυτή η δεύτερη αρχιτεκτονική προσδίδει στο δίκτυο πολύ περισσότερες δυνατότητες, αλλά είναι πιο δύσκολο να αντιμετωπιστεί μαθηματικά. Επίσης, πρέπει να σημειωθεί ότι η τοπολογία των δικτύων μπορεί να είναι είτε στατική είτε δυναμική (μεταβαλλόμενη). Τέλος, οι νευρώνες που δεν έχουν συνδέσεις με τον εξωτερικό κόσμο, λέμε ότι είναι «κρυμμένοι» (hidden) ή εσωτερικοί. Περισσότερα για τις αρχιτεκτονικές των δικτύων θα αναφερθούν σε επόμενη ενότητα.

• Κάθε νευρώνας υλοποιεί μια συνάρτηση τοπικά και ολόκληρο το δίκτυο υλοποιεί μια συγκεκριμένη λειτουργία. Στις περισσότερες περιπτώσεις, η εφαρμογή είναι αυτή που μέσω των προδιαγραφών της καθορίζει την απαιτούμενη λειτουργία. Ο καθορισμός των παραμέτρων του δικτύου (τιμών των βαρών) που θα ικανοποιούν αυτές τις προδιαγραφές επιτυγχάνεται μέσω της διαδικασίας μάθησης.

• Η γνώση, η εμπειρία και η εκπαίδευση του δικτύου αποθηκεύεται στις διασυνδέσεις των νευρώνων και στις τιμές των βαρών. Η εκπαίδευση επιτυγχάνεται με τη συνεχή τροποποίηση των τιμών των βαρών.



Σχήμα 1.1

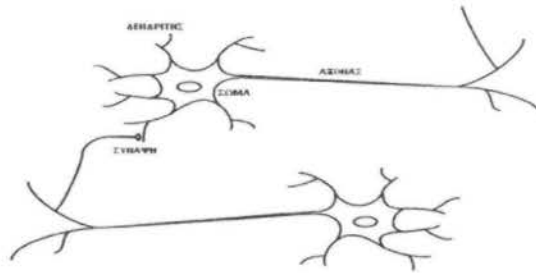
(α) Δίκτυο πρόσθιας τροφοδότησης (β) Επαναληπτικό δίκτυο.

• Στην πλειοψηφία τους τα ΤΝΔ εκπαιδεύονται με την ελπίδα ότι θα παρουσιάσουν καλή γενικευτική ικανότητα όταν θα τους ζητηθεί να αναγνωρίσουν ή να κατηγοριοποιήσουν καινούργια (άγνωστα) δεδομένα (πρότυπα). Αυτός είναι ο αντικειμενικός στόχος της διαδικασίας εκπαίδευσης, να αναπτύξει δηλαδή το δίκτυο κατάλληλη εσωτερική δομή ώστε να αναγνωρίζει πρότυπα που θα μοιάζουν με αυτά με τα οποία εκπαιδεύτηκε. Τα ΤΝΔ εκπαιδεύονται τόσο με τεχνικές μάθησης με επίβλεψη όσο και με τεχνικές μάθησης χωρίς επίβλεψη. Όλα αυτά θα αναλυθούν περισσότερο στη συνέχεια.

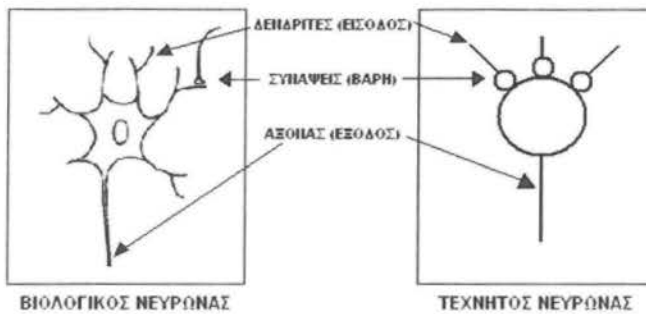
1.1.4 Ο ΒΙΟΛΟΓΙΚΟΣ ΚΑΙ Ο ΤΕΧΝΗΤΟΣ ΝΕΥΡΩΝΑΣ

Όπως προαναφέρθηκε, η λειτουργία του βιολογικού συστήματος βασίζεται στη διασύνδεση εξειδικευμένων φυσικών κυττάρων που ονομάζονται νευρώνες. Οι σημαντικές ιδιότητες των βιολογικών συστημάτων, όπως η προσαρμοστικότητα, η ικανότητα αναγνώρισης από τα συμφραζόμενα, η ανοχή στα λάθη, η μεγάλη χωρητικότητα μνήμης, η ικανότητα επεξεργασίας βιολογικών πληροφοριών σε πραγματικό χρόνο (κυρίως από τον εγκέφαλο), μας κατευθύνουν στην μελέτη και την προσπάθεια προσομοίωσης αυτών των εναλλακτικών βιολογικών αρχιτεκτονικών. Ωστόσο, δεν είναι ακόμα επαρκώς γνωστός ο τρόπος με τον οποίο λειτουργεί ο ανθρώπινος εγκέφαλος. Επιπλέον, παρά το γεγονός ότι το βασικό στοιχείο υπολογισμού του ανθρώπινου συστήματος επεξεργασίας είναι σχετικά αργό (σε σύγκριση με τα ηλεκτρονικά στοιχεία), η συνολική επεξεργασία επιτυγχάνεται σε μερικές εκατοντάδες msec. Αυτό μας οδηγεί στο συμπέρασμα ότι η βάση του βιολογικού υπολογισμού είναι ένας μικρός αριθμός ακολουθιακών βημάτων, καθένα από τα οποία εκτελείται με μεγάλο παραλληλισμό. Επιπλέον, στην έμφυτη αυτή παράλληλη αρχιτεκτονική, κάθε μονάδα επεξεργασίας είναι σχετικά απλή και τοπικά συνδεδεμένη.

Στο σχήμα 1.2 παρουσιάζεται η βασική δομή του βιολογικού νευρώνα, ο οποίος αποτελείται από το σώμα, τον άξονα, τους δενδρίτες και τις συνάψεις. Η βασική λειτουργία που επιτελεί ένας νευρώνας είναι η συσσώρευση των σημάτων που δέχεται από τους νευρώνες με τους οποίους συνδέεται η είσοδός του, το φιλτράρισμα και η ενίσχυση αυτών των σημάτων, και η παραγωγή ενός σήματος εξόδου το οποίο στη συνέχεια μεταδίδεται μέσω των συνάψεων προς τους νευρώνες με τους οποίους συνδέεται η έξοδός του. Ένα πολύ σημαντικό στοιχείο είναι ότι η επίδραση ενός νευρώνα στους γειτονικούς του μπορεί να είναι είτε διεγερτική (θετική σύνδεση) είτε ανασταλτική (αρνητική σύνδεση). Σε πλήρη αντιστοιχία με το απλοποιημένο αυτό μοντέλο του βιολογικού νευρώνα αναπτύχθηκε το μοντέλο του τεχνητού νευρώνα (Σχήμα 1.3).



Σχήμα 1.2
Δομή του βιολογικού νευρώνα.



Σχήμα 1.3

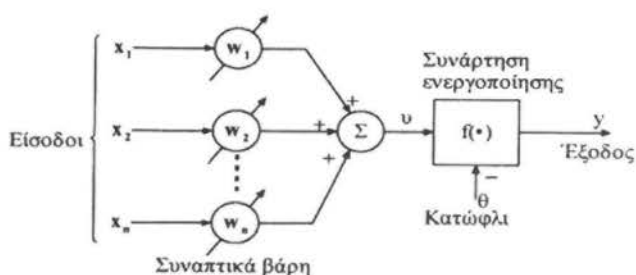
Αντιστοιχία βιολογικού-τεχνητού νευρώνα.

Έχοντας μόλις παρουσιάσει, εν συντομία, τη δομή και τη λειτουργία του βιολογικού νευρώνα, μπορούμε να προχωρήσουμε στην αναλυτική παρουσίαση του βασικού μοντέλου του τεχνητού νευρώνα. Παρατηρούμε, στο Σχήμα 1.4, ότι ο νευρώνας είναι μια θεμελιακή μονάδα επεξεργασίας πληροφορίας η οποία αποτελείται από τρία στοιχεία:

- Ένα σύνολο κλάδων διασύνδεσης (συνάψεων)
- Έναν κόμβο άθροισης (γραμμικής άθροισης)

- Μια συνάρτηση ενεργοποίησης (μη γραμμικότητα)

Κάθε κλάδος διασύνδεσης έχει ένα βάρος (weight) το οποίο είναι θετικό αν η σύναψη είναι διεγερτικού τύπου και αρνητικό αν η σύναψη είναι ανασταλτικού τύπου. Ο κόμβος άθροισης αθροίζει τα σήματα εισόδου πολλαπλασιαζόμενα με τα αντίστοιχα βάρη των συνάψεων. Συνεπώς ο κόμβος άθροισης είναι μια μονάδα γραμμικού συνδυασμού. Τέλος η συνάρτηση ενεργοποίησης (ή συνάρτηση σύνθλιψης / squashing function) περιορίζει (συνθλίβει) το επιτρεπόμενο πλάτος του σήματος εξόδου σε κάποια πεπερασμένη τιμή (συνήθως στο κανονικοποιημένο διάστημα $[0,1]$ ή εναλλακτικά στο $[-1,1]$). Το μοντέλο του νευρώνα περιέχει επίσης ένα κατώφλι θ που εφαρμόζεται εξωτερικά και πρακτικά υποβιβάζει την καθαρή είσοδο στη συνάρτηση ενεργοποίησης. Βεβαίως, η καθαρή είσοδος της συνάρτησης ενεργοποίησης μπορεί να αυξηθεί αν χρησιμοποιήσουμε έναν όρο πόλωσης b , ο οποίος είναι το αρνητικό του κατωφλίου ($b = -\theta$).



Σχήμα 1.4

Βασικό μοντέλο τεχνητού νευρώνα.

Από το Σχήμα 1.4 λοιπόν προκύπτει ότι ο νευρώνας περιγράφεται από τις παρακάτω εξισώσεις:

$$u = \sum_{j=1}^n w_j x_j$$

(4.1)

$$y = f(u - \theta), \theta > 0$$

(1.2)

όπου x_i ($i = 1, 2, \dots, n$) είναι τα σήματα εισόδου, w_i ($i = 1, 2, \dots, n$) είναι τα συναπτικά βάρη του νευρώνα, u είναι η είσοδος του γραμμικού συνδυαστή, θ είναι το κατώφλι, $f(\cdot)$ είναι η συνάρτηση ενεργοποίησης και y είναι το σήμα εξόδου του νευρώνα. Προφανώς, η συνάρτηση ενεργοποίησης εφαρμόζεται στο σήμα $u = u - \theta$, $\theta > 0$ και συνεπώς αν η είσοδος κατωφλίου θ συμπεριληφθεί στις κανονικές εισόδους θέτοντας $x_0 = -1$ και $w_0 = \theta$, τότε οι εξισώσεις (1.1) και (1.2) γράφονται ως:

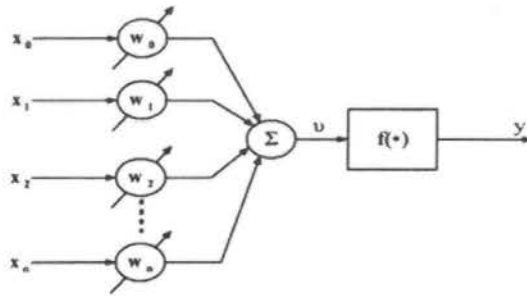
$$v = \sum_{j=0}^n w_j x_j$$

(1.3)

$$y = f(v)$$

(1.4)

και το μοντέλο του νευρώνα του Σχήματος 1.4 παίρνει τη μορφή του Σχήματος 1.5

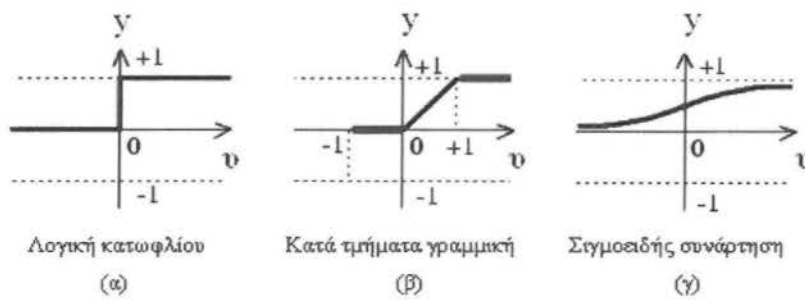


Σχήμα 1.5

Το βασικό μοντέλο του νευρώνα με το κατώφλι ($x_0=-1, w_0=\theta$) ή την πόλωση ($x_0=+1, w_0=b$) συμπεριλαμβανόμενο (η) στο σύνολο των κανονικών εισόδων.

1.1.5 ΤΥΠΟΙ ΣΥΝΑΡΤΗΣΕΩΝ ΕΝΕΡΓΟΠΟΙΗΣΗΣ

Ας μελετήσουμε τώρα τη συνάρτηση ενεργοποίησης, η οποία μπορεί να έχει μια από τις παρακάτω τρεις μορφές:



Σχήμα 1.6 - Συνήθεις μορφές συνάρτησης ενεργοποίησης νευρώνων.

Οι συναρτήσεις του Σχήματος 1.6 ορίζονται ως ακολούθως:

- Συνάρτηση κατωφλίου (Σχήμα 1.6 α)

$$f(v) = \begin{cases} 1, v \geq 0 \\ 0, v < 0 \end{cases}$$

(1.5)

Ο αντίστοιχος νευρώνας έχει τις εξισώσεις:

$$v = \sum_{j=1}^n w_j x_j - \theta$$

(1.6)

$$y = \begin{cases} 1, v \geq 0 \\ 0, v < 0 \end{cases}$$

(1.7)

• Κατά τμήματα γραμμική συνάρτηση (Σχήμα 1.6 β)

$$f(v) = \begin{cases} 1, v \geq 1 \\ v, 0 < v < 1 \\ 0, v \leq 0 \end{cases}$$

(1.8)

Η συνάρτηση αυτή παριστά ένα γραμμικό ενισχυτή με κορεσμό και μεταπίπτει στη συνάρτηση κατωφλίου αν η απολαβή (ενίσχυση) του ενισχυτή γίνει πάρα πολύ μεγάλη (θεωρητικά άπειρη).

• Σιγμοειδής συνάρτηση (Σχήμα 1.6 γ)

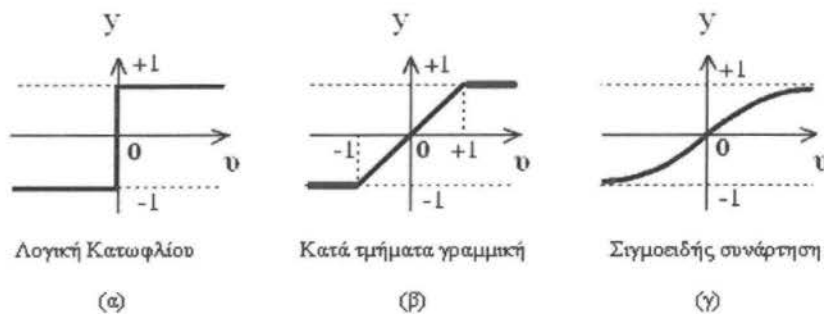
Η σιγμοειδής συνάρτηση είναι η ευρύτερα χρησιμοποιούμενη συνάρτηση ενεργοποίησης και μπορεί να οριστεί με πολλούς τρόπους. Ένας από αυτούς είναι η λογιστική (logistic) συνάρτηση:

$$f(v) = \frac{1}{1 + \exp(-\lambda v)}$$

(1.9)

όπου λ είναι η «παράμετρος κλίσης» της σιγμοειδούς συνάρτησης. Για διάφορα λ παίρνουμε σιγμοειδείς συναρτήσεις με διαφορετική κλίση. Η κλίση στην αρχή ($v=0$) είναι ίση με $\lambda/4$.

Οι συναρτήσεις ενεργοποίησης του Σχήματος 1.6 παίρνουν τιμές μεταξύ 0 και 1, δηλαδή $0 \leq f(v) \leq 1$. Εκτός από αυτές σε πολλές εφαρμογές χρειαζόμαστε τιμές στο διάστημα $[-1, +1]$, οπότε έχουμε τις συναρτήσεις του Σχήματος 1.7:



Σχήμα 1.7

Συναρτήσεις ενεργοποίησης με τιμές στο διάστημα $[-1,+1]$.

Οι συναρτήσεις αυτές περιγράφονται από τις ακόλουθες εξισώσεις.

- Συνάρτηση κατωφλίου {συνάρτηση προσήμου} (Σχήμα 1.7 α)

$$f(v) \begin{cases} +1, v > 0 \\ 0, v = 0 \\ -1, v < 0 \end{cases}$$

(1.10)

- Κατά τμήματα γραμμική συνάρτηση (Σχήμα 1.7 β)

$$f(v) = \begin{cases} 1, v \geq +1 \\ v, -1 < v < +1 \\ -1, v \leq -1 \end{cases}$$

(1.11)

- Σιγμοειδής συνάρτηση (υπερβολική εφαπτομένη) (Σχήμα 4.7 γ)

$$f(v) = \tanh\left(\frac{v}{2}\right) = \frac{1 - \exp(-v)}{1 + \exp(-v)}$$

(1.12)

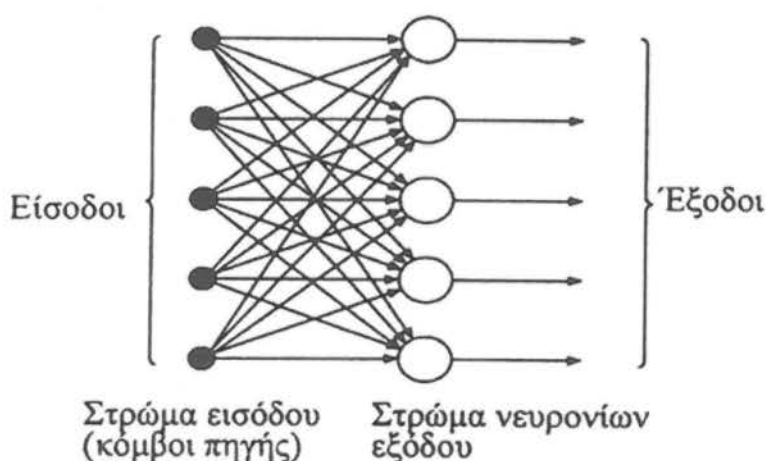
1.1.6 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Η τοπολογική δομή είναι το κύριο χαρακτηριστικό των ΤΝΔ και αναφέρεται στην αρχιτεκτονική στην οποία διευθετούνται και διασυνδέονται πολλαπλοί νευρώνες. Οι δυο βασικές ιδιότητες που καθορίζουν την αρχιτεκτονική ενός δικτύου είναι το πλήθος των στρωμάτων (layers) και οι συνδέσεις ανάμεσα στους νευρώνες. Το τρίτο χαρακτηριστικό, το οποίο σχετίζεται φυσιολογικά με τον τρόπο κατά τον οποίο είναι δομημένοι οι νευρώνες, είναι ο αλγόριθμος μάθησης που χρησιμοποιείται για την εκπαίδευση του δικτύου. Η μάθηση θα εξετασθεί σε επόμενη ενότητα.

ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΠΡΟΣΟΤΡΟΦΟΔΟΤΗΣΗΣ

Στην απλούστερη περίπτωση ένα διαστρωματωμένο νευρωνικό δίκτυο (δηλαδή ένα ΤΝΔ του οποίου οι νευρώνες είναι οργανωμένοι σε στρώματα) έχει ένα στρώμα εισόδου από «κόμβους πηγής» (source nodes) το οποίο προβάλλεται πάνω σε ένα στρώμα νευρώνων εξόδου (κόμβων υπολογισμού), αλλά όχι αντίστροφα. Ένα τέτοιο

ΤΝΔ είναι αυστηρά του τύπου «προσοτροφοδότησης» (feedforward) (Σχήμα 1.8) και καλείται «νευρωνικό δίκτυο προσοτροφοδότησης ενός μοναδικού στρώματος, όπου βέβαια το μοναδικό στρώμα είναι το στρώμα των νευρώνων εξόδου. Αυτό σημαίνει ότι το στρώμα των κόμβων πηγής δεν προσμετράται γιατί δεν λαμβάνει χώρα κανένας υπολογισμός σε αυτό.



Σχήμα 1.8

Μονοστρωματικό ΤΝΔ προσοτροφοδότησης με 5 νευρώνες εξόδου.

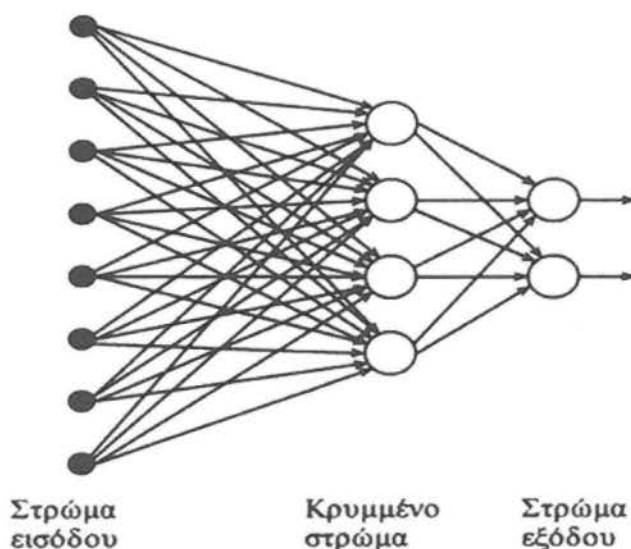
Ένα τυπικό παράδειγμα μονοστρωματικού ΤΝΔ προσοτροφοδότησης είναι η γραμμική συσχετιστική μνήμη η οποία αντιστοιχεί (συσχετίζει) ένα διάνυσμα (πρότυπο) εξόδου σε ένα διάνυσμα (πρότυπο) εισόδου και αποθηκεύει την πληροφορία ως αλλαγές στα συναπτικά βάρη.

Στη γενική περίπτωση, ένα ΤΝΔ προσοτροφοδότησης περιέχει ένα ή περισσότερα κρυμμένα στρώματα (hidden layers), των οποίων οι υπολογιστικοί κόμβοι είναι γνωστοί ως «κρυμμένοι νευρώνες» και παρεμβαίνουν μεταξύ των εξωτερικών εισόδων και των εξόδων του ΤΝΔ. Στα δίκτυα αυτά που ονομάζονται «πολυστρωματικά ΤΝΔ προσοτροφοδότησης», οι κόμβοι πηγής στο στρώμα εισόδου παρέχουν τα

στοιχεία του διανύσματος (προτύπου) δράσης τα οποία εισέρχονται ως είσοδοι στο πρώτο κρυμμένο στρώμα υπολογιστικών κόμβων. Ομοίως οι έξοδοι των κόμβων του πρώτου κρυμμένου στρώματος εισέρχονται ως είσοδοι στους κόμβους του δεύτερου κρυμμένου στρώματος, κ.ο.κ. Το τελικό στρώμα κόμβων (στρώμα εξόδου) δίνει τη συνολική απόκριση στα διανύσματα εισόδου (δηλαδή στα πρότυπα εξωτερικής δράσης).

Ένα παράδειγμα ΤΝΔ προσοτροφοδότησης με ένα κρυμμένο στρώμα τεσσάρων κόμβων, οκτώ κόμβων πηγής στο στρώμα εισόδου και δυο κόμβους στο στρώμα εξόδου εικονίζεται στο Σχήμα 1.9. Το δίκτυο αυτό αναφέρεται συνολικά ως ΤΝΔ προσοτροφοδότησης 8-4-2. Ένα άλλο παράδειγμα θα ήταν ένα ΤΝΔ 12-60-10-7 κ.ο.κ.

Το νευρωνικό δίκτυο του Σχήματος 1.9 της επόμενης σελίδας, ονομάζεται «πλήρως διασυνδεδεμένο» γιατί κάθε κόμβος (νευρώνας) οποιουδήποτε στρώματος συνδέεται με όλους τους κόμβους (νευρώνες) του γειτονικού του προς τα εμπρός (δηλαδή του επόμενου) στρώματος. Αν αυτό δεν ισχύει, δηλαδή αν λείπουν μερικοί σύνδεσμοι επικοινωνίας τότε το ΤΝΔ ονομάζεται «μερικώς διασυνδεδεμένο».



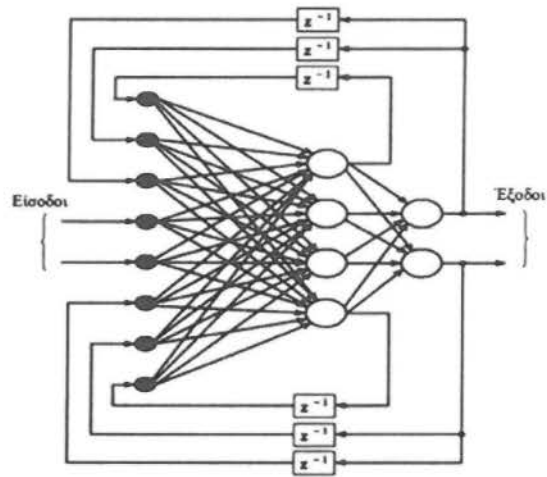
Σχήμα 1.9

Πολυστρωματικό ΤΝΔ προσοτροφοδότησης 8-4-2.

ΑΝΑΔΡΟΜΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ (ΑΝΑΤΡΟΦΟΔΟΤΗΣΗΣ)

Αν το ΤΝΔ περιέχει τουλάχιστον ένα βρόχο ανατροφοδότησης ο οποίος ανακυκλώνει πληροφορία μέσω του ίδιου ή προηγούμενων στρωμάτων, τότε ονομάζεται αναδρομικό νευρωνικό δίκτυο (recurrent NN) ή ΤΝΔ ανατροφοδότησης. Το αποτέλεσμα είναι ότι όταν ένα διάνυσμα (πρότυπο) εισόδου εισέρχεται στο αναδρομικό ΤΝΔ, δεν παράγει ένα πρότυπο εξόδου σε πεπερασμένο αριθμό χρονικών βημάτων, αλλά δρα με ένα κυκλικό τρόπο, με τα ίδια στρώματα να ενεργοποιούνται επαναληπτικά. Αν το δίκτυο είναι αφ' εαυτού ευσταθές πιθανά να ταλαντωθεί για κάποιο χρονικό διάστημα προτού φτάσει σε μια σταθερή κατάσταση, προτού παραχθεί δηλαδή μια σταθερή έξοδος. Διαφορετικά αν το δίκτυο δεν είναι ευσταθές οι ταλαντώσεις θα συνεχίσουν αδιάκοπα. Συνεπώς όταν εκπαιδεύουμε ένα αναδρομικό ΤΝΔ είναι σημαντικό να βρούμε το σύνολο των συναπτικών βαρών που του επιτρέπουν να σταθεροποιηθεί στις επιθυμητές τιμές εξόδου.

Ένα παράδειγμα αναδρομικού ΤΝΔ μπορεί να ληφθεί από το δίκτυο προσοτροφοδότησης 8-4-2 του Σχήματος 1.9 αν κλείσουμε διάφορους βρόχους όπως δείχνει το Σχήμα 1.10. Οι συνδέσεις ανατροφοδότησης μπορούν να ξεκινούν τόσο από τους κρυμμένους κόμβους όσο και από τους κόμβους εξόδου. Σημειώνεται ότι οι βρόχοι ανατροφοδότησης έχουν μια ευεργετική επίδραση πάνω στη μάθηση και τη συνολική συμπεριφορά του δικτύου.

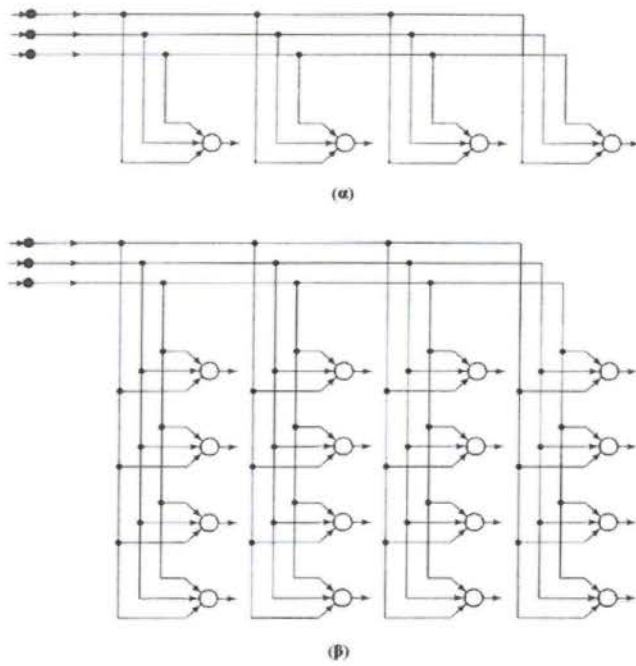


Σχήμα 1.10

Αναδρομικό ΤΝΔ ενός κρυμμένου στρώματος.

ΔΙΚΤΥΩΤΕΣ ΔΟΜΕΣ

Ένα πλέγμα, αποτελείται από έναν πίνακα μιας-δυο ή και μεγαλύτερης διάστασης από νευρώνες, με ένα αντίστοιχο σύνολο από πηγαίους κόμβους που παρέχουν τα σήματα εισόδου στον πίνακα. Δυο παραδείγματα τέτοιων δικτύων εικονίζονται στο Σχήμα 1.11.



Σχήμα 1.11

(α) Μονοδιάστατο πλέγμα με 4 νευρώνες (β) Δισδιάστατο πλέγμα με 4x4 νευρώνες.

1.1.7 ΕΦΑΡΜΟΓΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΠΡΟΒΛΗΜΑΤΩΝ ΚΑΤΑΛΛΗΛΩΝ ΓΙΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Η εξομίωση των βιολογικών υπολογιστικών παραδειγμάτων που πραγματοποιείται μέσω των ΤΝΔ έχει πολύ καλά αποτελέσματα για πολλές κατηγορίες προβλημάτων. Μεταξύ αυτών είναι:

- η κατηγορία προβλημάτων αναγνώρισης (π.χ. αναγνώριση φωνής, εικόνας κλπ.)
- η κατηγορία των προβλημάτων ελέγχου των οποίων τα δεδομένα είναι ελλιπή, ασαφή και στοχαστικά.
- η κατηγορία των NP-πλήρων προβλημάτων, τα οποία περιλαμβάνουν προβλήματα δρομολόγησης, αναζήτησης κλπ.

Όλα αυτά τα προβλήματα χαρακτηρίζονται από τον πολυδιάστατο χώρο τους, την πολυπλοκότητά τους και τις άγνωστες συσχετίσεις μεταξύ των μεταβλητών τους. Επιπλέον τα ΤΝΔ προσφέρουν λύσεις κυρίως σε προβλήματα που σχετίζονται με τον ανθρώπινο παράγοντα (αναγνώριση ομιλίας, εικόνας, χειρόγραφου κειμένου κλπ.)

ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΦΑΡΜΟΓΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

- Επεξεργασία εικόνας και μηχανική όραση (π.χ. ταίριασμα εικόνας, προεπεξεργασία, κατάτμηση, ανάλυση, συμπίεση εικόνας και επεξεργασία χρονικά μεταβαλλόμενων εικόνων).
- Επεξεργασία σήματος (π.χ. ανάλυση και μορφολογία σεισμικού σήματος).
- Αναγνώριση προτύπων (π.χ. εξαγωγή χαρακτηριστικών, ανάλυση και κατηγοριοποίηση σήματος radar, αναγνώριση φωνής, κειμένου, χειρονομιών και πιστοποίηση ταυτότητας).

- Ιατρική (π.χ. ανάλυση ηλεκτροκαρδιογραφήματος, ιατρική διάγνωση και επεξεργασία ιατρικής εικόνας).
- Αμυντικά συστήματα (π.χ. υποβρύχια ανίχνευση ναρκών).
- Οικονομία (π.χ. ανάλυση αγοράς μετοχών, ασφάλεια συναλλαγών, εκτίμηση φερεγγυότητας δανειζόμενου πελάτη, εκτίμηση ακίνητης περιουσίας).
- Σχεδίαση, έλεγχος και αναζήτηση (π.χ. παράλληλη υλοποίηση NP-πλήρων προβλημάτων, αυτόματος έλεγχος, ρομποτική).
- Τεχνητή νοημοσύνη (π.χ. υλοποίηση έμπειρων συστημάτων).
- Δυναμικά εξελισσόμενα συστήματα, πρόβλεψη χρονοσειρών (π.χ. εκτίμηση κατάστασης συστήματος, ανίχνευση βλαβών και ανάκαμψη).
- Επικοινωνία ανθρώπου – υπολογιστή.

1.1.8 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΚΑΙ ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Επειδή τα ΤΝΔ αποτελούν ένα σχετικά πρόσφατο μοντέλο υπολογισμού, θα πρέπει να πούμε ότι τα πλεονεκτήματα, μειονεκτήματα και η σχέση τους με τις κλασσικές μεθόδους υπολογισμού δεν είναι πλήρως ξεκαθαρισμένα. Οι προσδοκίες σε αυτόν τον τομέα είναι μεγάλες. Σε γενικές γραμμές τα ΤΝΔ είναι κατάλληλα να αντιμετωπίσουν συγκεκριμένες κατηγορίες προβλημάτων και ειδικότερα προβλήματα

εκπαίδευσης που αντιμετωπίζονται με συσχέτιση προτύπων. Η αντίληψη ότι τα ΤΝΔ μπορούν να λύσουν όλα τα προβλήματα αυτόματης απόκτησης γνώσης είναι μη ρεαλιστική.

Πλεονεκτήματα

- Έμφυτος παραλληλισμός.
- Ανοχή σε βλάβες (λόγω του παραλληλισμού).
- Σχεδιασμένα να είναι προσαρμοζόμενα.
- Δεν υπάρχει η ανάγκη για χαρακτηρισμό του προβλήματος πέρα από το σύνολο των προτύπων εκπαίδευσης.

Μειονεκτήματα

- Δεν υπάρχουν σαφείς κανόνες για την ανάπτυξη ΤΝΔ για οποιαδήποτε εφαρμογή.
- Δεν υπάρχει γενικός τρόπος για την ερμηνεία της εσωτερικής λειτουργίας του δικτύου.
- Η εκπαίδευση μπορεί να είναι δύσκολη ή αδύνατη.
- Η ικανότητα γενίκευσης είναι δύσκολα προβλέψιμη.

1.2 ΙΣΤΟΡΙΚΕΣ ΣΗΜΕΙΩΣΕΙΣ

Η μοντέρνα περίοδος των νευρωνικών δικτύων λέγεται ότι άρχισε με την πρωτοποριακή δουλειά των McCulloch και Pitts (1943). Ο πρώτος ήταν ψυχίατρος και ο δεύτερος μεγαλοφυής μαθηματικός. Σύμφωνα με τον Rall (1990), η κλασική εργασία των McCulloch και Pitts έγινε μέσα σε μια κοινωνία που ασχολούνταν με τους νευρώνες στο πανεπιστήμιο του Σικάγο για πάνω από 5 χρόνια. Αυτή η εργασία περιέγραφε το λογικό λογισμό των νευρωνικών δικτύων. Είναι αξιοσημείωτο το ότι ο Von Neumann χρησιμοποίησε ιδεατά στοιχεία καθυστέρησης τα οποία είχαν υπολογιστεί από τα ιδεατά στοιχεία νευρώνων των McCulloch και Pitts, για την κατασκευή του EDVAC (Electronic Discrete Variable Automatic Computer), ο οποίος κατάληξε στον ENIAC, τον πρώτο γενικού σκοπού υπολογιστή.

Η επόμενη μεγάλη ανάπτυξη πάνω στα νευρωνικά δίκτυα, ήρθε το 1949 με την έκδοση του βιβλίου του Hebb με τίτλο "The Organization of Behavior", στο οποίο μια ιδιαίτερη δήλωση ενός φυσιολογικού κανόνα μάθησης για συναπτικές τροποποιήσεις έγινε για πρώτη φορά. Πιο συγκεκριμένα ο Hebb πρότεινε ότι η συνδετικότητα του εγκεφάλου συνεχώς αλλάζει καθώς ο οργανισμός μαθαίνει διάφορες εργασίες, και ότι οι νευρωνικοί συγκεντρωτές δημιουργούνται από τέτοιες αλλαγές. Επίσης πρότεινε το διάσημο αίτημα μάθησης σύμφωνα με το οποίο η αποτελεσματικότητα μιας σύναψης μεταβλητής ανάμεσα σε δύο νευρώνες αυξάνεται από την επαναλαμβανόμενη ενεργοποίηση του ενός νευρώνα από τον άλλο κατά μήκος της σύναψης.

Η αναφορά των Rochester, Holland, Habit και Duda (1956) είναι ίσως η πρώτη προσπάθεια για χρήση υπολογιστή που εξομοιώνει, για να τεσταριστεί μια καλά σχηματισμένη νευρωνική θεωρία βασισμένη στο αίτημα μάθησης του Hebb. Η εξομοίωση έδειξε ότι χρειαζόταν να προστεθεί παρεμπόδιση ώστε η θεωρία πραγματικά να δουλέψει. Τον ίδιο χρόνο ο Uttley παρουσίασε την αποκαλούμενη "διαρρέουσα ολοκλήρωση" ή "νευρώνας φωτιάς" που αργότερα αναλύθηκε από τον Caianiello.

Το 1952 το βιβλίο του Ashby: "Design for a brain: The Origin of Adaptive Behavior" εκδόθηκε, το οποίο ασχολήθηκε με την βασική έννοια ότι η προσαρμοζόμενη συμπεριφορά δεν είναι έμφυτη αλλά μαθαίνεται. Το 1954 ο Minsky έγραψε ένα διδακτορικό με τίτλο "Theory of Neural-Analog Reinforcement Systems and its Application to the Brain-Model problem" και το 1961 ο ίδιος έγραψε μια εργασία με τίτλο "Steps Toward Artificial Intelligence". Επίσης το 1954 η ιδέα των μη γραμμικών προσαρμοζόμενων φίλτρων προτάθηκε από τον Gabor (εφευρέτης του ολογραφήματος). Έχτισε μια τέτοια μηχανή και η μάθηση επιτυγχανόταν με τροφοδότηση δειγμάτων στοχαστικών διαδικασιών στη μηχανή, μαζί με συναρτήσεις αποτελεσμάτων που η μηχανή ανέμενε να βγάλει.

Ένα θέμα ιδιαίτερου ενδιαφέροντος για τα νευρωνικά δίκτυα είναι αυτό της σχεδίασης ενός αξιόπιστου δικτύου με νευρώνες που μπορούν να θεωρηθούν σαν μη αξιόπιστα στοιχεία. Αυτό το σημαντικό πρόβλημα λύθηκε από τον Von Neumann (1956) χρησιμοποιώντας την ιδέα του πλεονασμού. 15 χρόνια μετά την έκδοση της εργασίας των McCulloch και Pitt μια νέα προσέγγιση πάνω στο πρόβλημα της αναγνώρισης προτύπων έγινε από τον Rosenblatt (1958) στην εργασία του πάνω στο αισθητήριο (perceptron). Το ιδιαίτερο επίτευγμα του ήταν το αποκαλούμενο θεώρημα σύγκλισης αισθητηρίου (perceptron convergence theorem). Το 1960 οι Widrow και Hoff πρότειναν τον αλγόριθμο ελάχιστου μέσου τετραγώνου (Least Mean-Square, LMS) και τον χρησιμοποίησαν για να σχηματίσουν το Adaline (adaptive linear element). Η διαφορά ανάμεσα στο αισθητήριο και στο Adaline βρίσκεται στον τρόπο μάθησης. Ένα από τα από τα πρόσφατα εκπαιδευσιμα νευρωνικά δίκτυα με πολλαπλά στοιχεία είναι η δομή Madaline (Widrow). Κατά την διάρκεια της κλασσικής περιόδου του perceptron ('60) πιστευόταν ότι τα νευρωνικά δίκτυα μπορούσαν να κάνουν τα πάντα. Αλλά τότε εκδόθηκε το βιβλίο των Minsky και Papert που με μαθηματικά απέδειξε ότι υπάρχουν όρια πάνω στο τι μπορεί να υπολογιστεί από τα αισθητήρια. Ένα σημαντικό πρόβλημα πάνω στη σχεδίαση ενός πολυεπίπεδου αισθητηρίου είναι το πρόβλημα της ανάθεσης εμπιστοσύνης (credit assignment problem), το οποίο βρήκε την λύση του μόλις την δεκαετία του 1980.

Κατά την δεκαετία του '70 λόγω των προβλημάτων εγκαταλείφτηκε το ενδιαφέρον πάνω στα νευρωνικά δίκτυα. Μια σημαντική ενέργεια την δεκαετία αυτή ήταν οι χάρτες αυτοοργάνωσης με ανταγωνιστική μάθηση.

Το 1980 έγιναν πολλές εργασίες πάνω στην θεωρία αλλά και στον σχεδιασμό των νευρωνικών δικτύων. Ο Grossberg (1980) ανέπτυξε μια καινούργια αρχή αυτοοργάνωσης που συνδυάζει φιλτράρισμα από "κάτω προς τα πάνω" και αντίθετη αύξηση σε μικρή μνήμη με από "πάνω προς τα κάτω" ταίριασμα προτύπων και σταθεροποίηση του κώδικα μάθησης. Δινόμενης μιας τέτοιας ικανότητας μία δυναμική κατάσταση καλούμενη υιοθετητική ηχώ λαμβάνει χώρα. Αυτό το φαινόμενο δίνει την βάση για μια νέα κατηγορία νευρωνικών δικτύων γνωστά σαν adaptive resonance theory (ART).

Το 1982 ο Hopfield χρησιμοποίησε την ιδέα μια συνάρτησης ενέργειας για να φτιάξει ένα νέο τρόπο κατανόησης του υπολογισμού που γίνεται από τα δίκτυα με συμμετρικές συναπτικές συνδέσεις. Επιπλέον καθιέρωσε τον ισομορφισμό ανάμεσα σε τέτοια περιοδικά δίκτυα και σε ένα Ising μοντέλο που χρησιμοποιείται στην στατιστική. Αυτή η αναλογία άνοιξε τον δρόμο για ένα κατακλυσμό θεωριών για τα νευρωνικά δίκτυα. Αυτή η συγκεκριμένη τάξη νευρωνικών δικτύων με ανατροφοδότηση χρήστηκε ιδιαίτερης προσοχής το '80 και με το χρόνο έγιναν γνωστά σαν δίκτυα Hopfield.

Το 1983 οι Cohen και Grossberg έδωσαν μια νέα αρχή για σχεδίαση μιας διευθυνσιοδοτούμενης μνήμης (content-addressable memory) που περιλαμβάνει την έκδοση συνεχούς χρόνου του δικτύου Hopfield σαν μια ιδιαίτερη περίπτωση. Μια ακόμα σημαντική ανάπτυξη το 1982 ήταν η έκδοση της εργασίας του Kohonen πάνω στους χάρτες αυτοοργάνωσης, χρησιμοποιώντας μιας ή δύο διαστάσεων δικτυωτές δομές. Το 1983 οι Kirkpatrick, Gallat και Vecchi περιέγραψαν μια νέα διαδικασία που λέγεται εξομοιωμένη ανόπτηση για λύση προβλημάτων συνδυαστικής βελτιστοποίησης. Η εξομοιωμένη ανόπτηση χρησιμοποιείται στην στατιστική θερμοδυναμική και βασίζεται σε μια απλή τεχνική. Την ίδια χρονιά εκδόθηκε μια εργασία από τους Burto, Sutton και Anderson πάνω στην ενισχυμένη μάθηση, η οποία

δημιούργησε μεγάλο ενδιαφέρον πάνω στην ενισχυμένη μάθηση και την εφαρμογή της. Το 1984 ο Braitenberg εξέδωσε ένα βιβλίο με τίτλο "Vehicles: Experiments in Synthetic Psychology" το οποίο περιγράφει διάφορες μηχανές με απλή εσωτερική αρχιτεκτονική, και το οποίο ενσωματώνει μερικές σημαντικές αρχές της αυτοοργανούμενης απόδοσης. Το 1986 η ανάπτυξη του αλγορίθμου για πίσω διάδοση (back-propagation algorithm) αναφέρθηκε από τον Rumelhart.

Το 1988 ο Linsker περιέγραψε μια νέα αρχή για την αυτοοργάνωση σε ένα δίκτυο αισθητηρίων. Η αρχή αυτή σχεδιάστηκε ώστε να διατηρεί μέγιστη πληροφορία σχετικά με την με τα πρότυπα ενεργειών, που υπόκεινται σε περιορισμούς όπως συναπτικές συνδέσεις και δυναμικές περιοχές σύναψης. Ο Linsker χρησιμοποίησε αφηρημένες έννοιες πάνω στη θεωρία πληροφοριών ώστε να σχηματίσει τη αρχή της διατήρησης μέγιστης πληροφορίας.

Επίσης το 1988 οι Broomhead και Lowe περιέγραψαν μία διαδικασία για το σχεδιασμό "προς τα εμπρός τροφοδότησης" (feedforward) δικτύων χρησιμοποιώντας συναρτήσεις ακτινικής βάσης (RBF), που είναι μια εναλλαγή των πολυεπίπεδων αισθητηρίων. Το 1989 εκδόθηκε το βιβλίο του Mead με τίτλο "VLSI and Neural Systems". Αυτό το βιβλίο δίνει μια ασυνήθιστη μίξη περιεχομένων από την νευροβιολογία και την τεχνολογία VLSI. Ίσως περισσότερο από κάθε άλλη έκδοση, η εργασία του Hopfield (1982) και το δίτομο βιβλίο των Rumelhart και McClelland (1986), να ήταν οι πιο σημαίνουσες εκδόσεις υπεύθυνες για την αναζωογόνηση του ενδιαφέροντος για τα νευρωνικά δίκτυα στην δεκαετία του '80. Τα νευρωνικά δίκτυα έχουν σίγουρα διανύσει πολύ δρόμο από την εποχή των McCulloch και Pitts. Πραγματικά έχουν εγκαθιδρυθεί σαν ενδοπειθαρχικό αντικείμενο με βαθιές ρίζες στην επιστήμη νευρώνων, στην ψυχολογία, στα μαθηματικά, στις φυσικές επιστήμες και στην μηχανική. Δεν είναι αναγκαίο να πούμε ότι είναι εδώ για να μείνουν και θα συνεχίσουν να αναπτύσσονται σε θεωρία, σχεδιασμό και εφαρμογές.

1.3 ΑΝΑΠΑΡΑΣΤΑΣΗ ΓΝΩΣΗΣ

Στην προηγούμενη ενότητα χρησιμοποιήσαμε τον όρο “γνώση” στον ορισμό ενός νευρωνικού δικτύου χωρίς μία ιδιαίτερη περιγραφή του τι εννοούμε με αυτό. Τώρα ενδιαφερόμαστε γι’ αυτό το θέμα δίνοντας τον ακόλουθο γενικό ορισμό:

«Με τον όρο γνώση αναφερόμαστε σε αποθηκευμένη πληροφορία ή σε μοντέλα χρησιμοποιούμενα από ένα άτομο ή μηχανή για να μεταφράσουν ,προβλέψουν και κατά προσέγγιση να αντιδράσουν στον εξωτερικό κόσμο» (Fischler and Firschein, 1987).

Η μέγιστη εργασία για ένα νευρωνικό δίκτυο είναι να μάθει ένα μοντέλο του κόσμου (περιβάλλον) στο οποίο είναι εγκατεστημένο και να συντηρήσει το μοντέλο ικανοποιητικά σύμφωνα με τον πραγματικό κόσμο, ώστε να επιτύχει τους στόχους μιας εφαρμογής που μας ενδιαφέρει. Η γνώση του κόσμου αποτελείται από δύο είδη πληροφορίας:

1. Τη γνωστή κατάσταση του κόσμου, αναπαραριστάμενη από παράγοντες σχετικά με το τι είναι και ήταν γνωστό. Αυτός ο τύπος γνώσης αναφέρεται και σαν από πριν (prior) πληροφορία.
2. Παρατηρήσεις (μετρήσεις) του κόσμου, που έχουν αποκτηθεί από αισθητήρες σχεδιασμένους να διερευνούν το περιβάλλον στο οποίο το νευρωνικό δίκτυο υποτίθεται ότι λειτουργεί. Συνήθως αυτές οι παρατηρήσεις περιέχουν από τη φύση τους θόρυβο, καθώς υπόκεινται σε λάθη εξαιτίας των θορύβων του αισθητήρα και των ατελειών του συστήματος. Σε οποιαδήποτε περίπτωση, οι παρατηρήσεις που αποκτώνται με αυτό τον τρόπο δίνουν το κυρίως μέρος των πληροφοριών, από τις οποίες παίρνονται τα παραδείγματα τα οποία χρησιμοποιούνται για την εκπαίδευση του νευρωνικού δικτύου.

Κάθε παράδειγμα αποτελείται από ένα ζεύγος εισόδου / εξόδου: ένα σήμα εισόδου και η αντίστοιχη επιθυμητή αντίδραση για το νευρωνικό δίκτυο. Γι' αυτό ένα σύνολο παραδειγμάτων αναπαριστά γνώση σχετικά με το περιβάλλον ενδιαφέροντος. Θεωρείστε για παράδειγμα το πρόβλημα της χειρόγραφης αναγνώρισης ψηφίων, στο οποίο η είσοδος αποτελείται από μία εικόνα από άσπρα και μαύρα pixels και κάθε εικόνα αναπαριστά ένα από τα δέκα ψηφία τα οποία διαχωρίζονται από το φόντο.

Σ' αυτό το παράδειγμα η επιθυμητή αντίδραση ορίζεται από την «ταυτότητα» ενός συγκεκριμένου ψηφίου του οποίου η εικόνα αναπαρίσταται στο νευρωνικό δίκτυο σαν ένα σήμα εισόδου. Τυπικά το σύνολο των παραδειγμάτων που χρησιμοποιείται για την εκπαίδευση του δικτύου αποτελείται από μια μεγάλη ποικιλία ψηφίων που είναι αντιπροσωπευτικά της πραγματικής εικόνας του δικτύου. Δοσμένου ενός τέτοιου συνόλου παραδειγμάτων η σχεδίαση ενός νευρωνικού δικτύου μπορεί να γίνει ως εξής:

Πρώτα μια κατάλληλη αρχιτεκτονική επιλέγεται για το νευρωνικό δίκτυο, με μία είσοδο αποτελούμενη από κόμβους εισόδου ισάριθμους προς τα pixels μιας εικόνας εισόδου και μια έξοδο αποτελούμενη από δέκα νευρώνες (έναν για κάθε ψηφίο). Ένα υποσύνολο παραδειγμάτων χρησιμοποιείται τότε για την εκπαίδευση του δικτύου μέσω ενός κατάλληλου αλγορίθμου. Αυτή η φάση λέγεται μάθηση.

Δεύτερο η απόδοση της αναγνώρισης του εκπαιδευμένου δικτύου τεστάρεται με δεδομένα που δεν έχουν εισαχθεί πριν. Συγκεκριμένα μια εικόνα εισόδου παρουσιάζεται στο δίκτυο αλλά αυτή τη φορά δεν του λέγεται η ταυτότητα του ψηφίου στο οποίο ανήκει αυτή η συγκεκριμένη εικόνα. Η απόδοση του δικτύου τότε παίρνεται συγκρίνοντας την αναγνώριση του ψηφίου που αναφέρεται από το δίκτυο με την πραγματική ταυτότητα του ψηφίου.

Αυτή η δεύτερη φάση της λειτουργίας του δικτύου ονομάζεται γενίκευση, ένας όρος που είναι δανεισμένος από την ψυχολογία. Σ' αυτό βρίσκεται μια θεμελιώδης διαφορά ανάμεσα στον σχεδιασμό ενός νευρωνικού δικτύου και στο μέρος της κλασικής επεξεργασίας πληροφορίας (pattern classifier). Στην τελευταία περίπτωση συνήθως προχωράμε σχηματίζοντας πρώτα ένα μαθηματικό μοντέλο περιβαλλοντικών παρατηρήσεων επαληθεύοντας το μοντέλο με πραγματικά δεδομένα και κατόπιν

χτίζοντας τον σχεδιασμό από τη βάση του μοντέλου. Αντίθετα ο σχεδιασμός ενός νευρωνικού δικτύου βασίζεται απ' ευθείας σε πραγματικά δεδομένα με το σύνολο των δεδομένων να κάνει απ' ευθείας όλη τη δουλειά. Γι' αυτό το νευρωνικό δίκτυο όχι μόνο δίνει ένα τέλειο μοντέλο του περιβάλλοντος στο οποίο εγκαθίσταται, αλλά παρουσιάζει μια εργασία επεξεργασίας πληροφοριών μεγάλου ενδιαφέροντος.

Τα παραδείγματα που χρησιμοποιούνται για την εκπαίδευση ενός νευρωνικού δικτύου μπορεί να αποτελούνται τόσο από θετικά όσο και από αρνητικά παραδείγματα. Για παράδειγμα στο πρόβλημα της παθητικής ανίχνευσης με σόναρ για θετικά παραδείγματα δίνονται σαν είσοδο δεδομένα εκπαίδευσης που περιέχουν τον στόχο ενδιαφέροντος (π.χ. υποβρύχιο). Τώρα σ' ένα παθητικό περιβάλλον σόναρ η πιθανή παρουσία υποβρύχιας ζωής στα δεδομένα τεσταρίσματος προκαλεί περιστασιακούς λάθος συναγερμούς. Για να ανακουφιστεί αυτό το πρόβλημα αρνητικά παραδείγματα (π.χ. ηχώ από υποβρύχια ζωή) περιλαμβάνονται στα δεδομένα εκπαίδευσης ώστε να μάθει το δίκτυο να μην μπερδεύει την υποβρύχια ζωή με το στόχο.

Σ' ένα νευρωνικό δίκτυο μιας συγκεκριμένης αρχιτεκτονικής η αναπαράσταση γνώσης του περιβάλλοντος ορίζεται από τις τιμές που παίρνουν οι ελεύθεροι παράμετροι (π.χ. συναπτικά βάρη και κατώφλια ενεργοποίησης) του δικτύου. Ο τύπος αυτής της αναπαράστασης γνώσης αποτελεί τους διαφορετικούς σχεδιασμούς ενός νευρωνικού δικτύου και γι' αυτό κρατάει το κλειδί της απόδοσης του.

Το θέμα της αναπαράστασης γνώσης μέσα σ' ένα τεχνητό νευρωνικό δίκτυο είναι παρ' όλα αυτά πολύ περιπλεγμένο. Το θέμα γίνεται ακόμα πιο πολύπλοκο όταν έχουμε πολλαπλές πηγές πληροφορίας να ενεργούν πάνω στο δίκτυο και αυτές οι πηγές αλληλεπιδρούν μεταξύ τους. Η παρούσα κατανόηση αυτού του σημαντικού θέματος είναι πραγματικά η πιο αδύναμη σύνδεση πάνω σ' αυτά που ξέρουμε σχετικά με τεχνητά νευρωνικά δίκτυα. Παρ' όλα αυτά υπάρχουν 4 κανόνες για αναπαράσταση γνώσης που είναι κοινής λογικής. Οι κανόνες αυτοί περιγράφονται ευθύς αμέσως.

ΚΑΝΟΝΑΣ 1: Παρόμοιες εισόδοι από παρόμοιες ομάδες πρέπει συνήθως να παράγουν παρόμοιες αναπαραστάσεις μέσα στο δίκτυο και πρέπει γι' αυτό να ταξινομούνται σαν να ανήκουν στην ίδια κατηγορία.

Υπάρχει μια πληθώρα μέτρων για να αποφασιστεί η «ομοιότητα» ανάμεσα σε εισόδους. Ένα συνηθισμένο μέτρο ομοιότητας βασίζεται στην έννοια της Ευκλείδειας απόστασης. Για να γίνουμε πιο συγκεκριμένοι έστω x_i ένα N-διάστατο πραγματικό διάνυσμα:

$$x_i = [x_{i1}, x_{i2}, \dots, x_{iN}]^T \quad (1.13)$$

που όλα τα στοιχεία είναι πραγματικά (το T δηλώνει αναστροφή πίνακα). Το διάνυσμα x_i ορίζει ένα σημείο σ' ένα N-διάστατο χώρο που καλείται Ευκλείδειος χώρος και συμβολίζεται με R^N . Η απόσταση μεταξύ x_i και x_j ορίζεται σαν:

$$d_{ij} = ||x_i - x_j|| = \left[\sum_{n=1}^N (x_{in} - x_{jn})^2 \right]^{1/2} \quad (1.14)$$

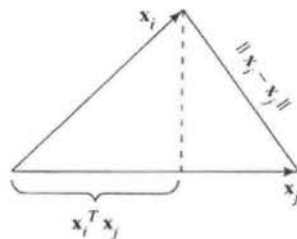
όπου x_{in} και x_{jn} είναι τα N-οστά στοιχεία του x_i και x_j . Αντίστοιχα η ομοιότητα ανάμεσα στις εισόδους που αναπαρίστανται από τα διανύσματα x_i και x_j ορίζεται σαν η παλινδρόμηση της Ευκλείδειας απόστασης d_{ij} . Όσο πιο κοντά είναι τα ξεχωριστά στοιχεία των διανυσμάτων εισόδου x_i και x_j , τόσο μικρότερη θα είναι η d_{ij} και τόσο μεγαλύτερη θα είναι η ομοιότητα των x_i και x_j . Ο κανόνας 1 δηλώνει ότι αν τα διανύσματα x_i και x_j είναι παρόμοια τότε θα πρέπει να τοποθετηθούν στην ίδια κατηγορία.

Ένα άλλο μέτρο της ομοιότητας βασίζεται στην ιδέα του εσωτερικού γινομένου (dot product) που το δανειζόμαστε από την γραμμική άλγεβρα. Δοθέντων δύο διανυσμάτων x_i και x_j ίδιας διάστασης το εσωτερικό του γινόμενο ορίζεται σαν:

$$x_i^T x_j = \sum_{n=1}^N x_{in} x_{jn}$$

(1.15)

Τα μέτρα ομοιότητας πραγματικά σχετίζονται μεταξύ τους όπως φαίνεται στο Σχήμα 4.12. Η Ευκλείδεια απόσταση $\|x_i - x_j\|$ ανάμεσα στα δύο διανύσματα x_i και x_j σκισάρεται σαν το μήκος της γραμμής που ενώνει τα άκρα των δύο αυτών διανυσμάτων, και το εσωτερικό του γινόμενο σκισάρεται σαν την προβολή του x_i πάνω στο διάνυσμα x_j . Το σχήμα δείχνει καθαρά ότι όσο μικρότερη είναι η Ευκλείδεια απόσταση, και γι' αυτό είναι πιο όμοια τα x_i και x_j , τόσο μεγαλύτερο θα είναι το εσωτερικό τους γινόμενο. Με όρους επεξεργασίας σημάτων το εσωτερικό γινόμενο μπορεί να ειπωθεί σαν μια συνάρτηση cross-correlation. Αναγνωρίζοντας ότι το εσωτερικό γινόμενο είναι βαθμωτό μπορούμε να δηλώσουμε ότι όσο πιο θετικό είναι τόσο πιο όμοια είναι τα διανύσματα x_i και x_j . Η συνάρτηση cross-correlation είναι ιδανική για εύρεση ηχούς σε ραντάρ και σόναρ συστήματα. Συγκεκριμένα περνώντας την ηχώ με συνάρτηση cross-correlate απ' το στόχο με ένα αντίγραφο του μεταδιδόμενου σήματος και βρίσκοντας την κορυφαία τιμή της συνάρτησης αποτελέσματος, είναι εύκολο να υπολογιστεί ο χρόνος άφιξης της ηχούς. Αυτή είναι βασική μέθοδος υπολογισμού της απόστασης του στόχου.



Σχήμα 1.12

Σχέση του εσωτερικού γινομένου και Ευκλείδειας απόστασης σαν ομοιότητα μεταξύ προτύπων.

ΚΑΝΟΝΑΣ 2: Αντικείμενα που κατηγοριοποιούνται σε διαφορετικές ομάδες θα πρέπει να έχουν διαφορετικές αναπαραστάσεις στο δίκτυο. Ο δεύτερος κανόνας είναι αντίθετος του πρώτου.

ΚΑΝΟΝΑΣ 3: Αν ένα συγκεκριμένο χαρακτηριστικό είναι σημαντικό τότε θα πρέπει να υπάρχει ένας μεγάλος αριθμός από νευρώνες που συμμετάσχουν στην αναπαράσταση αυτού του αντικειμένου στο δίκτυο.

Θεωρείστε για παράδειγμα την εφαρμογή ραντάρ για την ανίχνευση ενός στόχου με την παρουσία θορύβου. Σύμφωνα με το κριτήριο Neyman-Pearson η πιθανότητα ανίχνευσης μεγιστοποιείται αν το όριο της πιθανότητας λάθους συναγερμού δεν υπερβαίνει μια συγκεκριμένη τιμή. Σε μια τέτοια εφαρμογή η πραγματική παρουσία ενός στόχου στο λαμβανόμενο σήμα αναπαριστά ένα σημαντικό χαρακτηριστικό της εισόδου. Ο κανόνας 3 σαν αποτέλεσμα δηλώνει ότι θα πρέπει να υπάρχει ένας πολύ μεγάλος αριθμός νευρώνων ώστε να αποφασιστεί ότι ο στόχος είναι παρών όταν πράγματι είναι. Κατά τον ίδιο τρόπο θα πρέπει να υπάρχει ένας πολύ μεγάλος αριθμός νευρώνων ώστε να αποφασιστεί ότι υπάρχει μόνο θόρυβος. Και στις δύο περιπτώσεις ο μεγάλος αριθμός νευρώνων βεβαιώνει το μεγάλο βαθμό ακρίβειας της απόφασης και της απόκλισης εξαιτίας λανθασμένων νευρώνων.

ΚΑΝΟΝΑΣ 4: Οι από πριν (prior) πληροφορίες και μη μεταβλητότητες θα πρέπει να χτιστούν στον σχεδιασμό ενός νευρωνικού δικτύου απλοποιώντας έτσι τον σχεδιασμό αυτό του δικτύου, το οποίο δεν χρειάζεται να τις μάθει.

Ο κανόνας 4 είναι ιδιαίτερα σημαντικός διότι η κληρονομικότητα των αποτελεσμάτων τους δίνει ένα νευρωνικό δίκτυο με μια εξειδικευμένη δομή. Αυτό είναι ιδιαίτερα επιθυμητό για διάφορους λόγους:

1) Βιολογικά, οπτικά και ελεγκτικά δίκτυα είναι γνωστό ότι είναι πολύ εξειδικευμένα.

- 2) Ένα νευρωνικό δίκτυο με εξειδικευμένη δομή συνήθως έχει πολύ μικρότερο αριθμό ελεύθερων παραμέτρων διαθέσιμων για κανονικοποίηση απ' ό,τι ένα πλήρως συνδεδεμένο δίκτυο. Συνεπώς ένα εξειδικευμένο δίκτυο απαιτεί λιγότερα δεδομένα εισόδου για εκπαίδευση, μαθαίνει γρηγορότερα και συχνά γενικεύει καλύτερα.
- 3) Ο ρυθμός της μετάδοσης πληροφοριών μέσω ενός εξειδικευμένου δικτύου επιταχύνεται.
- 4) Το κόστος χτισίματος ενός εξειδικευμένου δικτύου μειώνεται εξαιτίας του μικρότερου μεγέθους σε σύγκριση μ' αυτό του πλήρως συνδεδεμένου.

1.4 ΜΑΘΗΣΗ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Η μάθηση είναι μια θεμελιακή ικανότητα των ΤΝΔ η οποία τους επιτρέπει να μαθαίνουν από το περιβάλλον τους και να βελτιώνουν τη συμπεριφορά τους με το πέρασμα του χρόνου. Η μάθηση είναι μια γενική έννοια της επιστήμης των συστημάτων και ορίζεται με ποικίλους τρόπους ανάλογα με το εκάστοτε πεδίο εφαρμογής. Στα ΤΝΔ, η μάθηση αναφέρεται στη διεργασία επίτευξης μιας επιθυμητής συμπεριφοράς μέσω ανανέωσης της τιμής των συναπτικών βαρών. Έτσι, ένα ΤΝΔ μαθαίνει για το περιβάλλον του μέσω μιας επαναληπτικής διαδικασίας ανανέωσης (αλλαγής) των συναπτικών βαρών και κατωφλίων.

Αλγόριθμος μάθησης ή εκπαίδευσης είναι κάθε προκαθορισμένο σύνολο καλά ορισμένων κανόνων επίλυσης του προβλήματος μάθησης (ή εκπαίδευσης) του ΤΝΔ. Γενικά, υπάρχουν πολλοί αλγόριθμοι εκπαίδευσης στα ΤΝΔ, καθένας από τους οποίους έχει συγκεκριμένα πλεονεκτήματα και μειονεκτήματα, προσφέροντας έναν άλλο τρόπο προσαρμογής (επιλογής / ανανέωσης) των βαρών. Περισσότερα για τους αλγόριθμους εκπαίδευσης θα αναφερθούν σε επόμενη ενότητα.

Ένα άλλο θέμα το οποίο πρέπει να ληφθεί υπ' όψη κατά τη σχεδίαση ΤΝΔ είναι το περιβάλλον στο οποίο εργάζεται κάθε δίκτυο. Έτσι, διαφορετικά μοντέλα του περιβάλλοντος οδηγούν σε διαφορετικά παραδείγματα μάθησης, δηλαδή την επιβλεπόμενη (ενεργή), την μη επιβλεπόμενη (αυτό-οργανούμενη) και την ενισχυτική μάθηση.

Τα τρία αυτά παραδείγματα μάθησης θα τα εξετάσουμε αφού πρώτα παρουσιαστεί ένας από τους βασικούς τρόπους νευρωνικής μάθησης, η μάθηση διόρθωσης σφάλματος που είναι και αυτή που θα μας απασχολήσει στη συνέχεια.

1.4.1 ΜΑΘΗΣΗ ΔΙΟΡΘΩΣΗΣ ΣΦΑΛΜΑΤΟΣ

Στον τύπο αυτό μάθησης η απαιτούμενη ανανέωση (αλλαγή, προσαρμογή) των συναπτικών βαρών υπολογίζεται παρουσιάζοντας στο ΤΝΔ δεδομένα πρότυπα (διανύσματα) εισόδου, συγκρίνοντας τις προκύπτουσες αποκρίσεις με τις (από πριν δοσμένες / a-priori) επιθυμητές αποκρίσεις και ακολούθως αλλάζοντας τα βάρη προς την κατεύθυνση μείωσης του σφάλματος.

Συγκεκριμένα έστω:

$d_k(t)$: η επιθυμητή έξοδος (απόκριση στόχου) του k νευρώνα στη διακριτή χρονική στιγμή t

$x(t)$: το διάνυσμα εισόδου (ερεθισμού) που εφαρμόζεται στο στρώμα εισόδου του ΤΝΔ

$y_k(t)$: η πρακτικά λαμβανόμενη απόκριση του k νευρώνα

Προφανώς το ζεύγος $(x(t), d_k(t))$ αποτελεί ένα παράδειγμα που παρουσιάζεται στο νευρώνα κατά τη χρονική στιγμή t . Το σφάλμα (διαφορά) μεταξύ της επιθυμητής εξόδου $d_k(t)$ και της πραγματικής εξόδου $y_k(t)$ είναι:

$$e_k(t) = d_k(t) - y_k(t)$$

(1.16)

με βάση το οποίο ορίζουμε το ακόλουθο κριτήριο (συνάρτηση κόστους):

$$I = E \left[\frac{1}{2} \sum_k e_k^2(t) \right]$$

(1.17)

το οποίο καλείται «κριτήριο (συνάρτηση) μέσου τετραγωνικού σφάλματος» (Mean Square Error criterion, MSE) και εκφράζει τη μέση τιμή του αθροίσματος των τετραγώνων των σφαλμάτων. Εδώ $E[]$ είναι ο στατιστικός τελεστής αναμενόμενης (μέσης) τιμής και θεωρείται (σιωπηρά) ότι το ΤΝΔ δρα σε ένα στάσιμο πιθανοτικό περιβάλλον με άγνωστες κατανομές πιθανότητας. Η άθροιση στο I εκτείνεται σε όλα τα k εξόδου δηλαδή σε όλους τους νευρώνες εξόδου.

Το πρόβλημα μάθησης (εκπαίδευσης) είναι τώρα: «Να επιλεγούν τα συναπτικά βάρη των νευρώνων έτσι ώστε να ελαχιστοποιηθεί το μέσο τετραγωνικό κριτήριο (1.17)».

Η ακριβής λύση του προβλήματος αυτού απαιτεί τη γνώση των στατικών ιδιοτήτων των στοχαστικών ανελίξεων που ενυπάρχουν σε κάθε περίπτωση. Γι' αυτό εδώ θα περιορισθούμε σε μια προσεγγιστική λύση ελαχιστοποιώντας το στιγμιαίο κριτήριο τετραγωνικού σφάλματος:

$$J = \frac{1}{2} \sum_k e_k^2(t)$$

(1.18)

ως προς τα συναπτικά βάρη w_{kj} του ΤΝΔ, όπου w_{kj} είναι το βάρος της σύναψης j του νευρώνα k . Ο κανόνας μάθησης ο οποίος βασίζεται στη διόρθωση του σφάλματος είναι τότε:

$$\Delta w_{kj}(t) = \gamma \cdot e_k(t) \cdot x_j(t)$$

(1.19)

όπου $\Delta w_{kj}(t)$ είναι η ποσότητα αλλαγής (προσαρμογής, ανανέωσης) στο χρόνο t του συναπτικού βάρους w_{kj} και γ είναι μια θετική σταθερά η οποία καθορίζει το ρυθμό μάθησης (learning rate). Ο κανόνας μάθησης (1.19) ο οποίος στη βιβλιογραφία είναι γνωστός ως «κανόνας δέλτα» λέει ότι η διόρθωση του συναπτικού βάρους είναι του σφάλματος $e_k(t)$ και του σήματος εισόδου $x_j(t)$ της υπό εξέταση σύναψης j .

Η ανανεωμένη (νέα) τιμή $w_{kj}(t+1)$ του θεωρούμενου συναπτικού βάρους δίνεται από τη σχέση:

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj}(t)$$

(1.20)

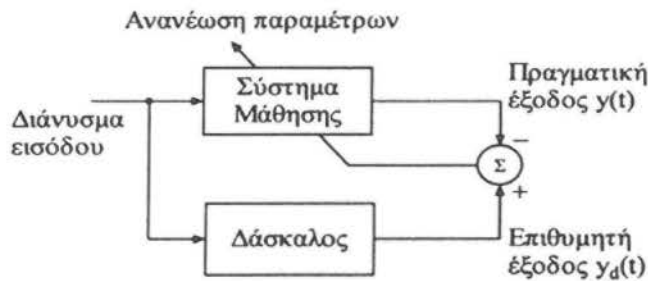
Για το ξεκίνημα του κανόνα μάθησης (1.20) χρειαζόμαστε τη γνώση κάποιων αρχικών τιμών για τα βάρη στη χρονική στιγμή $t = 0$. Αν το ΤΝΔ περιέχει μόνο γραμμικά στοιχεία επεξεργασίας, οπότε το J είναι ακριβώς τετραγωνικό, τότε ο αλγόριθμος οδηγεί βήμα – προς – βήμα στο ολικό ελάχιστο. Όταν όμως το ΤΝΔ περιέχει μη γραμμικότητες, τότε το ολικό ελάχιστο δεν μπορεί να ληφθεί πάντα γιατί ο αλγόριθμος μπορεί να εγκλωβισθεί σε κάποιο τοπικό ελάχιστο.

Επιπλέον, λόγω του ότι η μάθηση διόρθωσης σφάλματος συμπεριφέρεται σαν σύστημα κλειστής αναδρομής πρέπει να επιλεγεί με πολύ προσοχή η τιμή του ρυθμού μάθησης γ ώστε να εξασφαλισθεί η σταθερότητα της διαδικασίας. Αυτό γιατί ο ρυθμός μάθησης έχει μεγάλη επίδραση στην απόδοση της μεθόδου και επιδρά όχι μόνο στην ταχύτητα σύγκλισης της μάθησης αλλά και στην ίδια την κατάληξή της. Αν το γ έχει μικρή τιμή η διαδικασία προχωρά ομαλά αλλά μπορεί να πάρει πολύ χρόνο στο σύστημα να συγκλίνει σε σταθερή λύση. Αντίθετα, αν η τιμή του γ είναι μεγάλη η πορεία της μάθησης επιταχύνεται αλλά υπάρχει κίνδυνος η διαδικασία να αποκλίνει και το σύστημα να γίνει ασταθές.

1.4.2 ΠΑΡΑΔΕΙΓΜΑΤΑ ΜΑΘΗΣΗΣ.

ΕΠΙΒΛΕΠΟΜΕΝΗ ΜΑΘΗΣΗ

Η δομή της επιβλεπόμενης ή ενεργού μάθησης (supervised / active learning) έχει τη γενική μορφή του Σχήματος 1.13 και περιλαμβάνει δυο κύριες συνιστώσες, το δάσκαλο και το σύστημα μάθησης.



Σχήμα 1.13

Δομή της επιβλεπόμενης μάθησης.

Το κύριο χαρακτηριστικό της επιβλεπόμενης μάθησης είναι η ύπαρξη του εξωτερικού δασκάλου ο οποίος με βάση τη γνώση και την εμπειρία του είναι ικανός να διδάξει στο σύστημα μάθησης (δηλαδή στο ΤΝΔ) τις επιθυμητές εξόδους (εξόδους στόχου) για ένα σύνολο εισόδων εκπαίδευσης. Όταν ο δάσκαλος και το ΤΝΔ λάβουν ένα διάνυσμα εισόδου εκπαίδευσης, ο δάσκαλος δίνει στο ΤΝΔ μια επιθυμητή έξοδο η οποία παριστά τη βέλτιστη ενέργεια (δράση) που πρέπει να κάνει το ΤΝΔ. Οι παράμετροι του δικτύου ανανεώνονται με βάση τόσο το διάνυσμα εκπαίδευσης όσο και το σήμα σφάλματος (δηλαδή της διαφοράς μεταξύ της πραγματικής απόκρισης $y(t)$ και της επιθυμητής απόκρισης $y_d(t)$ του δικτύου). Ορίζοντας μια τετραγωνική συνάρτηση κόστους:

$$J(\mathbf{w}) = \frac{1}{2} E[(y(t) - y_d(t))^2]$$

(1.21)

όπου w είναι το διάνυσμα των προς επιλογή (ελεύθερων) παραμέτρων του συστήματος μάθησης (δηλαδή του ΤΝΔ), η ανανέωση / προσαρμογή των παραμέτρων παίρνει την μορφή του αλγόριθμου διόρθωσης σφάλματος που είδαμε στην προηγούμενη ενότητα. Η βήμα – προς – βήμα ανανέωση των παραμέτρων κάνει τελικά το ΤΝΔ μάθησης να μιμείται καλά τον δάσκαλο. Δυο περιπτώσεις αλγόριθμων επιβλεπόμενης μάθησης είναι ο αλγόριθμος Ελαχίστου Μέσου Τετραγώνου (Least Mean Square, LMS) και η γενίκευση του που είναι γνωστή ως Αλγόριθμος Ανάστροφης Διάδοσης (Back Propagation, BP) και θα μας απασχολήσει στα επόμενα.

ΕΝΙΣΧΥΤΙΚΗ ΜΑΘΗΣΗ

Στον τύπο αυτό μάθησης, το ΤΝΔ τροφοδοτείται και πάλι με δειγματικά πρότυπα εισόδου αλλά δεν τροφοδοτείται με τις επιθυμητές αποκρίσεις σε αυτές τις εισόδους. Εδώ χρησιμοποιείται ένα συνολικό μέτρο της επάρκειας της προκύπτουσας απόκρισης (δείκτης συμπεριφοράς) το οποίο μπορεί να οδηγήσει το δίκτυο στην επιθυμητή συμπεριφορά. Το μέτρο αυτό είναι γνωστό ως ενισχυτικό σήμα (reinforcement signal) και ανατροφοδοτείται στο ΤΝΔ έτσι ώστε να επιβραβεύσει (reward) τις ορθές συμπεριφορές και να τιμωρήσει (penalize) τις λανθασμένες.

Η ενισχυτική μάθηση στηρίζεται στο νόμο του Thorndike που λέει ότι «Αν μια ενέργεια ενός συστήματος μάθησης ακολουθείται από μια ικανοποιητική κατάσταση ή συμπεριφορά, τότε η τάση του συστήματος αυτού να παράγει τη συγκεκριμένη ενέργεια ενισχύεται (ενδυναμώνεται). Διαφορετικά, η τάση του συστήματος να παράγει την ενέργεια αυτή εξασθενείται».

Συνοπτικά, η ενισχυτική μάθηση λειτουργεί ως εξής:

1. Το ΤΝΔ υπολογίζει τις εξόδους που παράγονται από την παρούσα είσοδο με τις παρούσες τιμές των βαρών.
2. Το σύστημα αξιολογεί την έξοδο και το ενισχυτικό σήμα τροφοδοτείται στο δίκτυο.
3. Τα βάρη ανανεώνονται με βάση το ενισχυτικό σήμα αυξάνοντας τις τιμές των βαρών που συνέβαλαν σε καλή συμπεριφορά ή μειώνοντας τις τιμές αυτών που προκάλεσαν κακή συμπεριφορά.
4. Το δίκτυο ψάχνει να βρει ένα σύνολο βαρών τα οποία τείνουν να αποφύγουν αρνητικά ενισχυτικά σήματα.

Η διαφορά ανάμεσα στην ενισχυτική και την επιβλεπόμενη μάθηση είναι ότι στην ενισχυτική μάθηση το σύστημα μάθησης βελτιώνεται χρησιμοποιώντας ένα κριτήριο συμπεριφοράς οι τιμές του οποίου δίνονται από το περιβάλλον, ενώ στην επιβλεπόμενη μάθηση το κριτήριο συμπεριφοράς (συνάρτηση τετραγωνικού σφάλματος) καθορίζεται εσωτερικά με βάση τις επιθυμητές αποκρίσεις (στόχους). Συνεπώς ένα σύστημα ενισχυτικής μάθησης μπορεί να θεωρηθεί ως ένα «σύστημα αξιολόγησης με ανατροφοδότηση».

ΜΗ ΕΠΙΒΛΕΠΟΜΕΝΗ ΜΑΘΗΣΗ

Στον τύπο αυτό μάθησης, που καλείται και αυτό-οργανούμενη μάθηση, δεν χρησιμοποιείται εξωτερικός δάσκαλος για να επιβλέψει την εκπαίδευση του ΤΝΔ. Το δίκτυο αντί να μάθει συγκεκριμένα παραδείγματα (ζεύγη) εισόδου-εξόδου, μαθαίνει ένα «ανεξάρτητο-καθηκόντων μέτρο» της ποιότητας της παράστασης. Οι προς επιλογή ελεύθερες παράμετροι του δικτύου προσαρμόζονται ώστε να βελτιστοποιηθεί το μέτρο αυτό. Πρακτικά το μόνο που χρειάζεται ένα ΤΝΔ μη επιβλεπόμενης μάθησης είναι να συντονισθεί στις στατιστικές ομαλότητες των δεδομένων εισόδου και μετά να μπορέσει

να δημιουργήσει εσωτερικές παραστάσεις για την κωδικοποίηση των ιδιοτήτων εισόδου και να παράγει αυτόματα νέες κατηγορίες.

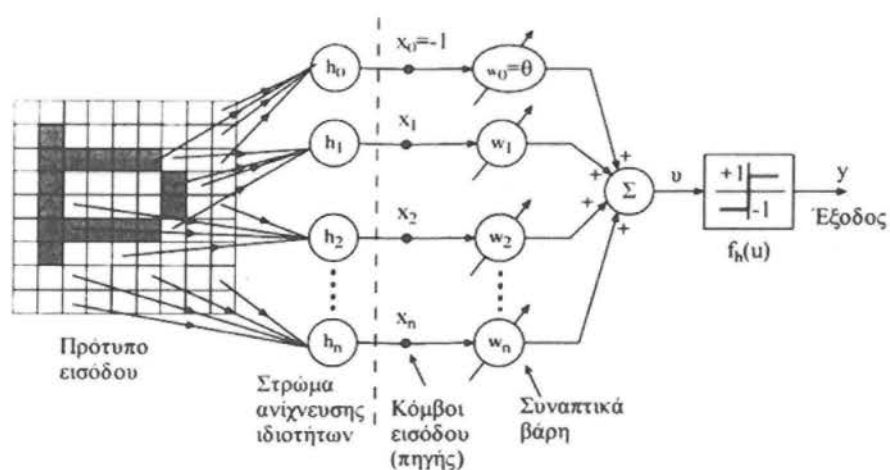
1.5. ΠΟΛΥΣΤΡΩΜΑΤΙΚΟ PERCEPTRON (MLP)

Τα νευρωνικά δίκτυα με τη βοήθεια των οποίων θα προσπαθήσουμε να λύσουμε το πρόβλημα του χρονοπρογραμματισμού εργασιών είναι τα ονομαζόμενα πολυστρωματικά perceptron (Multi-Layer Perceptrons, MLP), τα οποία ανήκουν στην κατηγορία των νευρωνικών δικτύων επιβλεπόμενης μάθησης (ανανεώνουν δηλαδή τις παραμέτρους τους με τον κανόνα επιβλεπόμενης μάθησης που παρουσιάστηκε νωρίτερα, ακολουθώντας την αρχιτεκτονική του σχήματος 1.13). Πριν παρουσιασθούν τα MLP και ο αλγόριθμος ανάστροφης διάδοσης που αυτά χρησιμοποιούν, κρίνεται απαραίτητη μια σύντομη παρουσίαση του απλού (μονοστρωματικού) perceptron.

1.5.1 ΤΟ ΑΠΛΟ PERCEPTRON

Το ΤΝΔ Perceptron αναπτύχθηκε από τον Roseblatt (1958) και αποτελεί μια από τις πρώτες προσπάθειες σχεδίασης και κατασκευής ευφυών συστημάτων με δυνατότητες αυτομάθησης. Το perceptron είναι το απλούστερο ΤΝΔ που χρησιμοποιείται για την ταξινόμηση γραμμικά διαχωρίσιμων προτύπων, δηλαδή προτύπων τα οποία διαχωρίζονται από ένα υπερεπίπεδο. Αποτελείται από έναν απλό νευρώνα με προσαρμόσιμα βάρη ο οποίος ακολουθείται από μια διπολική συνάρτηση ενεργοποίησης της μορφής του Σχήματος 1.7α. Η δομή του perceptron (ενός νευρώνα) φαίνεται στο σχήμα 1.14. Αυτό το δίκτυο μπορεί να πραγματοποιήσει ταξινόμηση στην

περίπτωση που έχουμε μόνο δυο κατηγορίες. Για να είναι δυνατή η ταξινόμηση περισσότερων γραμμικά διαχωρίσιμων κατηγοριών (κλάσεων) το στρώμα εξόδου χρειάζεται να έχει περισσότερους από έναν νευρώνες.



Σχήμα 1.14

Perceptron ενός νευρώνα. Οι εισοδοι του perceptron λαμβάνονται από ένα στρώμα ανίχνευσης ιδιοτήτων με σταθερά βάρη h_1, h_2, \dots, h_n .

Η έξοδος του perceptron δίνεται από τη σχέση:

$$y(t) = f_h(v(t))$$

(1.22)

$$v(t) = \sum_{i=0}^n w_i x_i = \mathbf{w}^T(t) \mathbf{x}(t)$$

(1.23)

όπου $f_h(u)$ είναι η συνάρτηση του σκληρού περιοριστή (διπολικής συνάρτησης) και

$$\mathbf{x}(t) = [-1, x_1(t), x_2(t), \dots, x_n(t)]^T$$

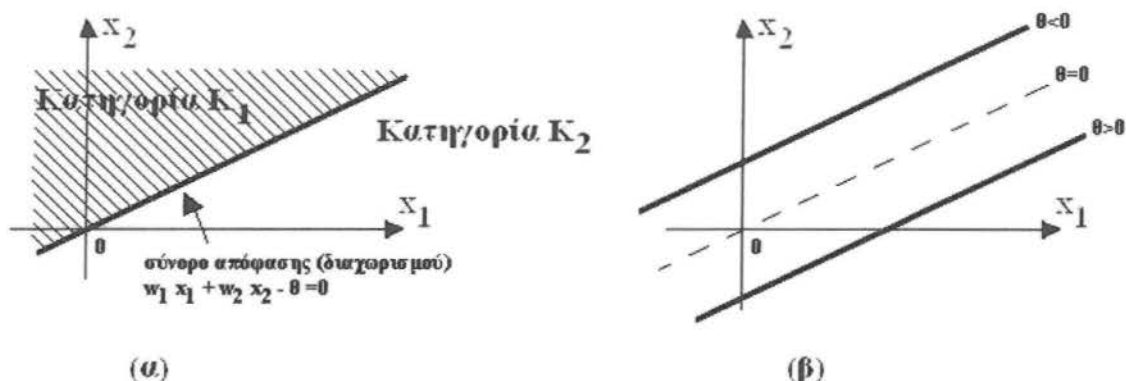
(4.24)

$$\mathbf{w}(t) = [\theta(t), w_1(t), w_2(t), \dots, w_n(t)]^T$$

(1.25)

Ο σκοπός του perceptron είναι να ταξινομήσει τα εξωτερικά πρότυπα x_1, x_2, \dots, x_n σε μια από δυο κατηγορίες K_1 ή K_2 . Τούτο γίνεται εκχωρώντας ένα σημείο $[x_1, x_2, \dots, x_n]^T$ στην κατηγορία K_1 αν η έξοδος y είναι $+1$ και στην κατηγορία K_2 αν είναι -1 . Η διδασκαλία του perceptron πρέπει να γίνει με καλά παραδείγματα κάθε μιας κατηγορίας. Αν οι προς διαχωρισμό κατηγορίες μοιάζουν πολύ μεταξύ τους τότε είναι δύσκολο να εκπαιδευτεί το perceptron να ταξινομεί τα πρότυπα που του παρουσιάζονται στη μια ή στην άλλη κατηγορία.

Για να κατανοήσουμε πως λειτουργεί ένας ταξινομητής προτύπων θεωρούμε την περίπτωση δυο μεταβλητών x_1 και x_2 όπου το σύνορο απόφασης (διαχωρισμού ανάμεσα στις κατηγορίες) είναι μια ευθεία γραμμή όπως δείχνει το Σχήμα 1.15:



Σχήμα 1.15

(α) Γραμμικός διαχωρισμός δυο κλάσεων K_1 και K_2 στην περίπτωση δυο διαστάσεων (το κατώφλι θ απλά μεταθέτει το σύνορο διαχωρισμού) (β) Σχετικός μετασχηματισμός που παράγεται από την παρουσία του κατωφλίου

Οποιοδήποτε σημείο βρίσκεται πάνω από τη διαχωριστική επιφάνεια ταξινομείται στην κατηγορία K_1 , διαφορετικά αποδίδεται στην κατηγορία K_2 . Ο αλγόριθμος επιλογής (ανανέωσης) των βαρών στηρίζεται στο γεγονός ότι αν οι κλάσεις K_1 και K_2 είναι γραμμικά διαχωρίσιμες, τότε υπάρχει ένα διάνυσμα βαρών w τέτοιο ώστε:

$$w^T x \geq 0 \quad \text{όταν το } x \text{ ανήκει στην } K_1$$

(1.26)

$$w^T x < 0 \quad \text{όταν το } x \text{ ανήκει στην } K_2$$

Συνεπώς όταν το perceptron λάβει ένα ζευγάρι εκπαίδευσης $(\mathbf{x}_1, \mathbf{x}_2)$ πρέπει να προσδιορίσει ένα διάνυσμα βάρους \mathbf{w} τέτοιο ώστε να ικανοποιούνται οι δυο ανισότητες (1.26).

Ο αλγόριθμος perceptron περιλαμβάνει τα ακόλουθα βήματα:

1. Διάλεξε μια αρχική τιμή $\mathbf{w}(0)$ του διανύσματος βάρους $\mathbf{w}(t)$. Συνήθως παίρνουμε

$\mathbf{w}(0)=0$ ή δίνουμε στα αρχικά βάρη $w_i(0)$ μικρές τυχαίες τιμές.

2. Παρουσίασε στο perceptron το διάνυσμα εισόδου \mathbf{x} και την επιθυμητή έξοδο $d(t)$.

3. Υπολόγισε την προκύπτουσα έξοδο με βάση τις (1.22) και (1.23):

$$y(t) = f_h(\mathbf{w}^T(t)\mathbf{x}(t)) = \text{sgn}\{\mathbf{w}^T(t)\mathbf{x}(t)\}$$

(1.27)

όπου $\text{sgn}\{\}$ είναι η συνάρτηση προσήμου:

$$f_h(v) = \text{sgn}(v) = \begin{cases} +1, v \geq 0 \\ -1, v < 0 \end{cases}$$

(1.28)

4. Αν η πραγματική έξοδος είναι σωστή, δηλαδή αν $\mathbf{w}^T(t)\mathbf{x}(t) \geq 0$ και το $\mathbf{x}(t)$ ανήκει στην

κατηγορία K_1 ή αν $\mathbf{w}^T(t)\mathbf{x}(t) < 0$ και το $\mathbf{x}(t)$ ανήκει στην κατηγορία K_2 , τότε δεν γίνεται καμία διόρθωση βάρους, δηλαδή:

$$\mathbf{w}(t+1) = \mathbf{w}(t)$$

(1.29)

Διαφορετικά γίνεται διόρθωση (ανανέωση) του βάρους ως εξής:

$$\mathbf{w}(t+1) = \mathbf{w}(t) - \gamma\mathbf{x}(t), \text{ αν } \mathbf{w}^T(t)\mathbf{x}(t) \geq 0 \text{ και } \mathbf{x}(t) \text{ ανήκει στην } K_2$$

(4.30)

ή

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \gamma\mathbf{x}(t), \text{ αν } \mathbf{w}^T(t)\mathbf{x}(t) < 0 \text{ και } \mathbf{x}(t) \text{ ανήκει στην } K_1$$

(1.31)

όπου $\gamma > 0$ είναι η παράμετρος που καθορίζει το ρυθμό μάθησης (ανανέωσης).

Ο Rosenblatt απέδειξε ότι ο αλγόριθμος αυτός συγκλίνει υπό την έννοια ότι το $\mathbf{w}(t^*) = \mathbf{w}(t^*+1) = \mathbf{w}(t^*+2) = \dots$ είναι ένα διάνυσμα βάρους λύσης για $t^* \leq t_{\max}$, όπου t^* είναι ένας πεπερασμένος αριθμός επαναλήψεων. Με άλλα λόγια, μετά από ένα πεπερασμένο αριθμό επαναλήψεων t^* σταματάει η μεταβολή των βαρών, ήτοι το perceptron έχει ταξινομήσει σωστά όλα τα διανύσματα εισόδου.

Ένας άλλος παρόμοιος αλγόριθμος δόθηκε από τους Widrow και Hoff και διαφέρει μόνο στο βήμα 4 που τώρα είναι το εξής:

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \gamma[d(t) - y(t)]\mathbf{x}(t)$$

(1.32)

όπου

$$d(t) = \begin{cases} +1, & \alpha \nu \quad \mathbf{x}(t) \in K_1 \\ -1, & \alpha \nu \quad \mathbf{x}(t) \in K_2 \end{cases}$$

(1.33)

Ο αλγόριθμος αυτός είναι γνωστός ως αλγόριθμος δέλτα (δ -algorithm) από το συμβολισμό του λάθους $\delta(t) = d(t) - y(t)$ που χρησιμοποίησαν οι Widrow και Hoff.

Η πρώτη σημαντική κριτική του perceptron του Rosenblatt, παρουσιάστηκε από τους Minsky και Selfridge (1961), οι οποίοι υπογράμμισαν ότι το perceptron δεν μπορεί ούτε καν να υλοποιήσει την συνάρτηση της ισοτιμίας δυο bits (πρόβλημα XOR). Τα υπολογιστικά όρια του perceptron καθορίστηκαν με αυστηρά μαθηματική διατύπωση στο βιβλίο των Minsky και Papert "Perceptrons" (1969, 1988). Στο βιβλίο αυτό απέδειξαν με αυστηρή μαθηματική ανάλυση ότι το perceptron, όπως αυτό ορίστηκε από τον Rosenblatt, είναι εγγενώς ανίκανο να αντιμετωπίσει δύσκολα προβλήματα ταξινόμησης. Στο τελευταίο κεφάλαιο του βιβλίου τους υποστήριξαν ότι οι περιορισμοί που είχαν ανακαλύψει για το perceptron θα έπρεπε λογικά να ισχύουν και για τις πολυστρωματικές επεκτάσεις του:

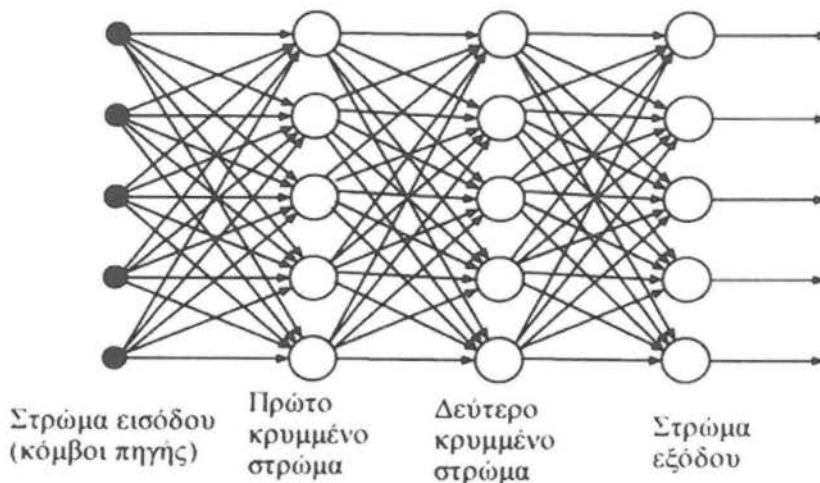
«Το perceptron μας έχει προσφέρει πολύτιμη γνώση εξαιτίας των πολλών περιορισμών του. Έχει πολλά χαρακτηριστικά που είναι άξια προσοχής: τη γραμμικότητα του, το ενδιαφέρον θεώρημα μάθησης, την απλότητά του ως ένα μοντέλο παράλληλου υπολογισμού. Ωστόσο, δεν υπάρχει κανένας λόγος να πιστέψουμε ότι κάποιο από αυτά τα προτερήματά του οδηγεί σε μια πολυστρωματική έκδοση. Ωστόσο, θεωρούμε ότι είναι σημαντικό ερευνητικό πρόβλημα η διαφώτιση (ή η

απόρριψη) της διαισθητικής μας κρίσης, ότι η επέκταση σε πολυστρωματικά συστήματα θα είναι άγονη».

Η παραπάνω εικασία δημιούργησε πολλές απορίες για τις υπολογιστικές ικανότητες όχι μόνο του perceptron αλλά και των νευρωνικών δικτύων. Εν τούτοις, η ιστορία έδειξε ότι η εικασία των Minsky και Papert δεν δικαιώνεται αφού σήμερα έχουμε διάφορους τύπους νευρωνικών δικτύων οι οποίοι είναι υπολογιστικά πολύ πιο ισχυροί από το perceptron του Rosenblatt. Χαρακτηριστικό παράδειγμα το πολυστρωματικό perceptron που εκπαιδεύεται με τον αλγόριθμο ανάστροφης διάδοσης σφάλματος (error backpropagation) και το οποίο παρουσιάζεται ευθύς αμέσως.

1.5.2 ΤΟ ΠΟΛΥΣΤΡΩΜΑΤΙΚΟ PERCEPTRON

Τα πολυστρωματικά perceptron (Multi-Layer Perceptron, MLP) είναι δίκτυα προστροφοδότησης (feedforward), αποτελούν γενίκευση του μονοστρωματικού perceptron και περιλαμβάνουν, εκτός από το στρώμα κόμβων εισόδου (πηγής) και το στρώμα εξόδων (αποκρίσεων), ένα ή περισσότερα στρώματα κρυμμένων κόμβων (hidden layers) όπως δείχνει το Σχήμα 1.16.



Σχήμα 1.16

Πολυστρωματικό perceptron με δυο κρυμμένα στρώματα.

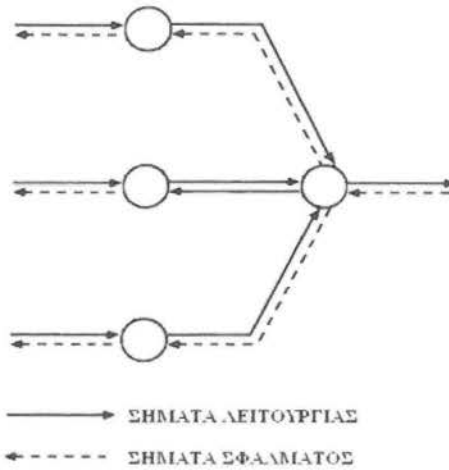
Τα πολυστρωματικά perceptron είναι κατάλληλα για την επίλυση πλήθους πολύπλοκων προβλημάτων και εκπαιδεύονται με επιβλεπόμενη μάθηση χρησιμοποιώντας τον αλγόριθμο ανάστροφης διάδοσης (BP: Back Propagation) ο οποίος στηρίζεται στον κανόνα μάθησης διόρθωσης σφάλματος. Η διαδικασία εκπαίδευσης στον αλγόριθμο back-propagation περιλαμβάνει υπολογισμούς που υλοποιούνται σε δυο περάσματα μέσω των στρωμάτων του δικτύου: ένα πέρασμα κατά την ευθεία φορά (από την είσοδο προς την έξοδο) και ένα κατά την ανάστροφη φορά (από την έξοδο προς την είσοδο). Στο ευθύ πέρασμα, εφαρμόζεται ένα πρότυπο στις εισόδους του δικτύου, πραγματοποιούνται οι υπολογισμοί κατά την ορθή φορά και στο τέλος παράγεται ένα σύνολο από εξόδους που αποτελούν και την πραγματική έξοδο του δικτύου. Κατά τη διάρκεια αυτού του περάσματος τα βάρη των συνδέσεων του δικτύου είναι σταθερά. Αντίθετα, κατά τη διάρκεια του αντίστροφου περάσματος, τα βάρη προσαρμόζονται σύμφωνα με τον κανόνα διόρθωσης σφάλματος. Συγκεκριμένα, η πραγματική τιμή της εξόδου αφαιρείται από την αντίστοιχη επιθυμητή και παράγεται το σήμα σφάλματος. Αυτό το σήμα στη συνέχεια προωθείται στο δίκτυο

κατά την ανάστροφη κατεύθυνση (από την έξοδο προς την είσοδο – απ' όπου και το όνομα του αλγορίθμου). Στη διαδικασία αυτή τα συναπτικά βάρη αναπροσαρμόζονται ώστε να μετατοπισθεί η απόκριση του δικτύου πλησιέστερα στην επιθυμητή.

Για να γίνει ευκολότερα κατανοητή η διαδικασία των δυο περασμάτων παρουσιάζεται στο Σχήμα 1.17 ένα τμήμα του δικτύου του Σχήματος 1.16. Σε αυτό διακρίνονται τα δυο ειδών σήματα:

Σήματα λειτουργίας: Τα σήματα αυτά εμφανίζονται στην είσοδο του δικτύου, ως ερέθισμα και προωθούνται κατά την ορθή φορά ανά στρώμα μέχρι την τελική έξοδο, όπου παράγονται τα σήματα εξόδου. Αναφερόμαστε σε αυτά με το όνομα σήματα λειτουργίας για δυο λόγους. Πρώτον, υποτίθεται ότι χρησιμοποιούνται κατά την κανονική λειτουργία του δικτύου και δεύτερον, κατά το πέρασμά τους από κάθε μονάδα νευρώνων, υλοποιείται ένας υπολογισμός ως συνάρτηση των εισόδων και των αντίστοιχων βαρών.

Σήματα σφάλματος: Αυτά τα σήματα δημιουργούνται (κατά την διάρκεια της εκπαίδευσης μόνο) στην έξοδο του δικτύου και προωθούνται κατά την ανάστροφη φορά διαμέσου του δικτύου.



Σχήμα 1.17

Ροή σημάτων σε τμήμα ενός MLP.

Ένα πολυστρωματικό perceptron έχει τα εξής χαρακτηριστικά:

1. Κάθε κρυμμένος νευρώνας περιέχει μια μη γραμμική συνάρτηση ενεργοποίησης. Μια

σημαντική διαφορά από το perceptron του Rosenblatt είναι η ομαλότητα αυτής της συνάρτησης σε σχέση με τη συνάρτηση κατωφλίου (σκληρό περιοριστή) και το γεγονός ότι είναι παραγωγίσιμη σε κάθε σημείο. Συναρτήσεις που ικανοποιούν αυτές τις απαιτήσεις και χρησιμοποιούνται συνήθως είναι η λογιστική συνάρτηση (βλ. εξίσωση 1.9):

$$y_k = f(v_k) = \frac{1}{1 + \exp(-\lambda v_k)}$$

(1.34)

ή η υπερβολική εφαπτομένη (βλ. εξίσωση 4.12):

$$y_k = f(v_k) = \frac{1 - \exp(-v_k)}{1 + \exp(-v_k)}$$

(1.35)

όπου $v_k = u_k - \theta_k$ είναι η καθαρή εσωτερική δραστηριότητα (κατάσταση) του k νευρώνα και y_k είναι η έξοδος του. Η ύπαρξη της μη-γραμμικότητας είναι πολύ σημαντική και είναι αυτή που προσδίδει στα δίκτυα MLP τις επιθυμητές υπολογιστικές δυνατότητες. Επιπλέον, η χρήση μια τέτοιας συνάρτησης υποστηρίζεται και βιολογικά από τη λειτουργία των φυσικών νευρώνων.

2. Το δίκτυο περιέχει ένα ή περισσότερα κρυμμένα στρώματα μη γραμμικών νευρώνων τα οποία δεν ανήκουν στα επίπεδα εισόδου ή εξόδου. Αυτοί οι νευρώνες καθιστούν το δίκτυο ικανό να μάθει πολύπλοκα πρότυπα εξάγοντας από αυτά κάποια ιδιαίτερα χαρακτηριστικά τους. Δεν υπάρχει σύνδεση μεταξύ νευρώνων του ίδιου επιπέδου (από τον ορισμό του επιπέδου) και συνήθως υπάρχει πλήρης διασύνδεση μεταξύ των νευρώνων δυο διαδοχικών επιπέδων.

Η υπολογιστική ισχύς του πολυστρωματικού perceptron προέρχεται από το συνδυασμό αυτών των χαρακτηριστικών και από την ικανότητα μάθησης μέσω εκπαίδευσης. Αυτά τα χαρακτηριστικά είναι όμως ταυτόχρονα υπεύθυνα για τις ελλείψεις που παρουσιάζουν οι γνώσεις μας γύρω από τη συμπεριφορά αυτών των

δικτύων. Η παρουσία κατανεμημένης μορφής μη-γραμμικότητας και η μεγάλη συνδεσιμότητα καθιστούν υπερβολικά δύσκολη τη θεωρητική ανάλυση τους. Επίσης, η χρήση κρυμμένων νευρώνων δυσκολεύει την οπτικοποίηση της διαδικασίας μάθησης.

1.5.3 Ο ΑΛΓΟΡΙΘΜΟΣ ΑΝΑΣΤΡΟΦΗΣ ΔΙΑΔΟΣΗΣ

Ο αλγόριθμος ανάστροφης διάδοσης (BP) αποτελεί επέκταση του κανόνα δέλτα που χρησιμοποιείται στο απλό perceptron και που παρουσιάστηκε στα προηγούμενα. Η ανάπτυξη του αποτελεί ορόσημο στην εξέλιξη των νευρωνικών δικτύων διότι παρέχει μια υπολογιστικά αποτελεσματική μέθοδο εκπαίδευσης των πολυστρωματικών perceptrons. Αν και δεν μπορούμε να ισχυριστούμε ότι ο αλγόριθμος αυτός μπορεί να επιλύσει όλα τα προβλήματα, πρέπει να διευκρινίσουμε ότι εξάλειψε κάθε απαισιοδοξία που είχαν εισαγάγει για τα MLP, με το βιβλίο τους "Perceptrons", οι Minsky και Papert το 1969. Αναλυτική μαθηματική παρουσίαση του αλγορίθμου ξεφεύγει από τα πλαίσια αυτής της εργασίας. Για χάρη πληρότητας όμως παρουσιάζονται εν συντομία τα βήματά του.

1. Διάλεξε τα αρχικά βάρη και κατώφλια χρησιμοποιώντας μικρές θετικές τυχαίες τιμές.
2. Παρουσίασε στο TND το διάνυσμα εκπαίδευσης

$$\mathbf{x}(t) = [x_0(t), x_1(t), \dots, x_n(t)]^T$$

και το επιθυμητό διάνυσμα εξόδου (στόχο)

$$\mathbf{d}(t) = [d_1(t), d_2(t), \dots, d_m(t)]^T$$

3. Υπολόγισε τα σήματα εξόδου όλων των νευρώνων του δικτύου νευρώνα-νευρώνα προς τα εμπρός χρησιμοποιώντας τις τρέχουσες τιμές των συναπτικών βαρών, δηλαδή:

$$y_j(t) = f_j(v_j(t)), \quad v_j(t) = \sum_i w_{ji}(t) y_i(t)$$

(1.36)

όπου $y_i(t)$ είναι η i είσοδος του j νευρώνα (δηλαδή η έξοδος του i νευρώνα) και w_{ji} είναι το συναπτικό βάρος που συνδέει τον i με τον j νευρώνα. Για τους νευρώνες j του πρώτου κρυμμένου στρώματος ισχύει $y_i(t) = x_i(t)$, $i = 1, 2, \dots, n$ όπου $x_i(t)$ είναι η i είσοδος (πρότυπο) ερεθισμού. Για τους νευρώνες j του στρώματος εξόδου, το $y_i(t)$ είναι η j πραγματική έξοδος (απόκριση) του δικτύου.

4. Ανανέωσε (βελτίωσε) τα βάρη αρχίζοντας από τους νευρώνες της εξόδου και προχωρώντας ανάστροφα προς το στρώμα εισόδου, χρησιμοποιώντας τον κανόνα:

$$w_{ji}(t+1) = w_{ji}(t) + \gamma \delta_j(t) y_i(t)$$

(1.37)

όπου το $\delta_j(t)$ ονομάζεται «τοπική κλίση» και αν ο θεωρούμενος νευρώνας ανήκει στο στρώμα εξόδου δίνεται από την:

$$\delta_j(t) = [d_j(t) - y_j(t)]y_j[1 - y_j(t)]$$

(1.38)

ενώ αν ο νευρώνας j ανήκει σε κρυμμένο στρώμα δίνεται από την:

$$\delta_j(t) = y_j(t)[1 - y_j(t)] \sum_m \delta_m(t) w_{mj}(t)$$

(1.39)

με το δείκτη m να εκτείνεται σε όλους τους νευρώνες του στρώματος εξόδου και το $w_{ji}(t)$ να παριστάνει το συναπτικό βάρος που συνδέει τον νευρώνα j ενός στρώματος με τον νευρώνα i του αμέσως προηγούμενου στρώματος (που κείται στα αριστερά του).

5. Επανάλαβε τη διαδικασία από το βήμα 2.

Σημείωση: Για να επιταχύνουμε τη σύγκλιση μπορούμε να προσθέσουμε στην εξίσωση (1.37) έναν πρόσθετο «όρο ορμής» (momentum term) $\alpha[w_{ji}(t) - w_{ji}(t-1)]$ όπου α είναι μια παράμετρος με τιμές στο διάστημα $0 < \alpha < 1$. Η ανανέωση των βαρών θα γίνεται σε αυτή την περίπτωση με βάση την εξίσωση:

$$w_{ji}(t+1) = w_{ji}(t) + \gamma \delta_j(t) y_i(t) + \alpha [w_{ji}(t) - w_{ji}(t-1)]$$

(1.40)

1.5.4 ΤΡΟΠΟΙ ΕΚΠΑΙΔΕΥΣΗΣ

Στην εφαρμογή του αλγορίθμου back propagation, η εκπαίδευση επιτυγχάνεται με την παρουσίαση στο δίκτυο ενός συνόλου παραδειγμάτων εκπαίδευσης (Training Data Set, TDS). Η παρουσίαση όλων των προτύπων του συνόλου εκπαίδευσης μια φορά το καθένα ονομάζεται εποχή (epoch). Η διαδικασία εκπαίδευσης εκτελείται σε επαναλήψεις εποχών ως ότου τα βάρη του δικτύου σταθεροποιηθούν σε συγκεκριμένες τιμές, για τις οποίες η μέση τιμή του σφάλματος (για το σύνολο προτύπων εκπαίδευσης) συγκλίνει στην ελάχιστη τιμή της. Είναι καλό σε κάθε εποχή να παρουσιάζεται με τυχαία σειρά το σύνολο προτύπων, έτσι ώστε η διαδικασία αναζήτησης στο χώρο των βαρών να είναι περισσότερο στοχαστική. Για ένα συγκεκριμένο σύνολο εκπαίδευσης, η διαδικασία εκπαίδευσης μπορεί να εκτελεστεί με δυο διαφορετικούς τρόπους:

1. Εκπαίδευση ανά πρότυπο (on-line ή incremental training). Στην περίπτωση αυτή, τα βάρη ενημερώνονται έπειτα από την παρουσίαση κάθε προτύπου στο δίκτυο. Πιο συγκεκριμένα, θεωρούμε μια εποχή που αποτελείται από πρότυπα εκπαίδευσης τοποθετημένα στη σειρά $(x(1),d(1)), \dots, (x(N),d(N))$ (στη γενική περίπτωση τα $x(i),d(i)$ είναι διανύσματα). Το πρώτο πρότυπο $(x(1),d(1))$ παρουσιάζεται στο δίκτυο και εκτελούνται οι ευθείς και οι αντίστροφοι υπολογισμοί οι οποίοι οδηγούν στη μεταβολή των συναπτικών βαρών. Στη συνέχεια παρουσιάζεται το δεύτερο πρότυπο και επαναλαμβάνεται η ίδια διαδικασία που οδηγεί σε νέα μεταβολή των βαρών κ.ο.κ. Η εποχή τελειώνει με την παρουσίαση και του τελευταίου προτύπου.

2.Εκπαίδευση ανά εποχή (o -line ή batch training). Στην περίπτωση αυτή, τα βάρη ενημερώνονται έπειτα από την παρουσίαση στο δίκτυο ολόκληρου του συνόλου προτύπων της εποχής.

Καθεμία από τις δυο αυτές μορφές εκπαίδευσης έχει τα πλεονεκτήματα και τα μειονεκτήματά της. Το ποια είναι η καταλληλότερη καθορίζεται από το είδος του προς επίλυση προβλήματος.

1.5.5 ΕΦΑΡΜΟΓΕΣ – ΔΥΝΑΤΟΤΗΤΕΣ ΤΟΥ ΠΟΛΥΣΤΡΩΜΑΤΙΚΟΥ PERCEPTRON

ΕΦΑΡΜΟΓΕΣ ΤΟΥ MLP

Η βασική λειτουργία που επιτελεί ένα πολυστρωματικό perceptron (MLP) είναι η υλοποίηση απεικόνισης (mapping) από το χώρο των εισόδων στο χώρο των εξόδων, χρησιμοποιώντας τα ζεύγη εκπαίδευσης και τους κατάλληλους αλγορίθμους εκπαίδευσης. Έτσι το MLP έχει χρησιμοποιηθεί με ιδιαίτερη επιτυχία για την κατασκευή συστημάτων πρόβλεψης (prediction), για την ταξινόμηση προτύπων (classification), για την κατασκευή μοντέλων από δεδομένα (data fitting), για τον έλεγχο συστημάτων, μέχρι και για την επίλυση μερικών διαφορικών εξισώσεων.

ΔΥΝΑΤΟΤΗΤΕΣ ΤΟΥ MLP

Κλείνοντας αυτή την ενότητα θα πρέπει να γίνει σύντομη αναφορά σε δυο πολύ βασικές δυνατότητες των πολυστρωματικών perceptron: τη δυνατότητα γενίκευσης και την ανεκτικότητα σε βλάβες.

• Γενίκευση

Γενίκευση (generalization) είναι η δυνατότητα ενός ΤΝΔ να ταξινομεί πρότυπα τα οποία ποτέ δεν γνώρισε. Η γενίκευση είναι μια από τις πιο σημαντικές δυνατότητες των ΤΝΔ. Υποτίθεται βεβαίως ότι οι προς ταξινόμηση είσοδοι (πρότυπα) λαμβάνονται από τον ίδιο πληθυσμό προτύπων ο οποίος χρησιμοποιήθηκε για την εκπαίδευση του δικτύου. Η γενίκευση επιτυγχάνεται ανιχνεύοντας ιδιότητες του προτύπου εισόδου που είναι σημαντικές και έχουν εμφυτευθεί (κωδικοποιηθεί) στους εσωτερικούς κόμβους. Ένα γνωστό πρότυπο ταξινομείται τότε μαζί με τα άλλα πρότυπα που έχουν τις ίδιες χαρακτηριστικές ιδιότητες. Το ΤΝΔ έχει επίσης την δυνατότητα να διακρίνει και να ταξινομεί (σχετικά καλά) ελλιπή, παραμορφωμένα και διαταραχθέντα με θόρυβο δεδομένα. Αν το δίκτυο έχει εκπαιδευτεί όπως πρέπει, τότε είναι ικανό να δώσει σωστή έξοδο έστω και αν η είσοδος που εφαρμόστηκε είναι ελαφρά διαφορετική από τις εισόδους των παραδειγμάτων με τα οποία εκπαιδεύτηκε.

Η γενίκευση ενός ΤΝΔ επηρεάζεται από τα ακόλουθα:

- α) Το μέγεθος και την καταλληλότητα του συνόλου εκπαίδευσης.
- β) Την αρχιτεκτονική του δικτύου.
- γ) Την πολυπλοκότητα του θεωρούμενου προβλήματος.

Στην πράξη λοιπόν προσπαθούμε να βελτιώσουμε την ποιότητα της γενίκευσης επιλέγοντας όσο το δυνατόν αντιπροσωπευτικότερο (representative) σύνολο εκπαίδευσης. Η καλύτερη μορφή (αρχιτεκτονική) του δικτύου επιλέγεται συνήθως με προσεκτική μελέτη του υπό εξέταση προβλήματος. Περισσότερα πάνω στην

καταλληλότητα ενός δικτύου να λύσει ένα συγκεκριμένο πρόβλημα και την επιλογή της βέλτιστης μορφής θα συζητηθούν σε επόμενο κεφάλαιο, όταν θα προσπαθήσουμε δηλαδή να βρούμε το δίκτυο που θα ταιριάζει καλύτερα στο πρόβλημα του χρονοπρογραμματισμού εργασιών.

• **Ανεκτικότητα Σε Βλάβες**

Το πολυστρωματικό perceptron είναι ενδογενώς ανεκτικό σε βλάβες γιατί είναι ένα σύστημα παράλληλης επεξεργασίας στο οποίο κάθε κόμβος επεξεργασίας (νευρώνας) συμβάλλει στο τελικό αποτέλεσμα της εξόδου. Αν χαθεί ή χαλάσει ένας νευρώνας ή υποστούν βλάβη τα συναπτικά του βάρη, η «ανάκτηση» πέφτει σε ποιότητα, αλλά η κατανεμημένη φύση της πληροφορίας εξασφαλίζει ότι η βλάβη θα πρέπει να είναι πολύ μεγάλη για να υποβιβασθεί η ποιότητα της απόκρισης σε μεγάλο βαθμό (επικίνδυνα). Συνεπώς, το δίκτυο παρουσιάζει ελαφρά πτώση της συμπεριφοράς του και όχι καταστροφική. Επίσης το πολυστρωματικό perceptron αντέχει σε θορύβους λόγω της δυνατότητας γενίκευσης που έχει. Η βλάβη στο δίκτυο αντιμετωπίζεται με επανεκπαίδευση και η «ανάρρωση» αυτή του δικτύου είναι συνήθως πολύ γρήγορη.

1.6 ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΚΑΙ MATLAB

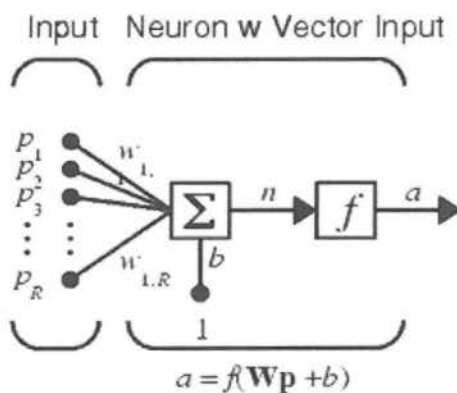
Το Neural Network Toolbox, που προσφέρει η MATLAB, είναι μια εξαιρετική συλλογή από συναρτήσεις για σχεδιασμό, εκπαίδευση και προσομοίωση νευρωνικών δικτύων. Υποστηρίζει πληθώρα αρχιτεκτονικών δικτύων, με τον αριθμό των υπολογιστικών μονάδων (νευρώνων) και των διασυνδέσεων να περιορίζεται μόνο από την υπολογιστική ισχύ του συστήματος του χρήστη. Οι αρχιτεκτονικές και μορφές εκπαίδευσης που υποστηρίζονται περιλαμβάνουν μεταξύ άλλων την υπό επίβλεψη εκπαίδευση feedforward δικτύων με χρήση του κανόνα διόρθωσης σφάλματος, πολλές παραλλαγές του αλγορίθμου ανάστροφης διάδοσης για να επιλέξει ο χρήστης αυτή που θα ανταποκριθεί καλύτερα στο πρόβλημά του, δίκτυα ακτινικών συναρτήσεων βάσης (RBF), αναδρομικά δίκτυα, αυτό-οργανούμενους χάρτες κ.α. Το Neural Network Toolbox περιλαμβάνει ποιοτικότερο οδηγό χρήσης, διδακτικά βοηθήματα και πολλά παραδείγματα για ευκολότερη κατανόηση της λογικής, του σχεδιασμού και της χρησιμότητας κάθε διαφορετικής αρχιτεκτονικής. Το Toolbox παραδίδεται υπό τη μορφή M-files επιτρέποντας στο χρήστη να δει με περισσότερη λεπτομέρεια τους αλγορίθμους, να κάνει αλλαγές ή να δημιουργήσει νέες συναρτήσεις ώστε να διευθετηθούν συγκεκριμένες εφαρμογές.

Ο συνδυασμός όλων αυτών των χαρακτηριστικών κάνουν την MATLAB ένα από τα πλέον διαδεδομένα εργαλεία στο χώρο των νευρωνικών δικτύων και είναι το εργαλείο που και εμείς χρησιμοποιήσαμε στα πλαίσια αυτής της εργασίας. Κρίνεται λοιπόν απαραίτητη, πριν κλείσουμε αυτό το κεφάλαιο, μια σύντομη παρουσίαση του Neural Network Toolbox και συγκεκριμένα των όσων αφορούν τα πολυστρωματικά perceptrons (MLP), ώστε να γίνει ομαλότερα η εισαγωγή του αναγνώστη στο κεφάλαιο

της υλοποίησης MLP για λύση του προβλήματος χρονοπρογραμματισμού εργασιών που ακολουθεί.

1.6.1 ΤΡΟΠΟΣ ΠΑΡΑΣΤΑΣΗΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Στο σχήμα 1.18 φαίνεται ένας νευρώνας με διάνυσμα εισόδου ένα διάνυσμα R -στοιχείων. Οι εισοδοί του νευρώνα p_1, p_2, \dots, p_R πολλαπλασιάζονται με τα βάρη $w_{1,1}, w_{1,2}, \dots, w_{1,R}$ και τροφοδοτούνται στον κόμβο άθροισης. Στην έξοδο του αθροιστή (ουσιαστικά στο εσωτερικό γινόμενο του διανύσματος των εισόδων \mathbf{p} με το διάνυσμα των βαρών \mathbf{W}) προστίθεται η τιμή της πόλωσης b και το αποτέλεσμα, n , είναι το όρισμα της συνάρτησης ενεργοποίησης f που δίνει έξοδο $a = f(\mathbf{W}\mathbf{p} + b) = f(w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b)$. Ο νευρώνας του Σχήματος 1.18 λοιπόν, είναι στην πραγματικότητα το βασικό μοντέλο τεχνητού νευρώνα που παρουσιάστηκε στο Σχήμα 1.4.



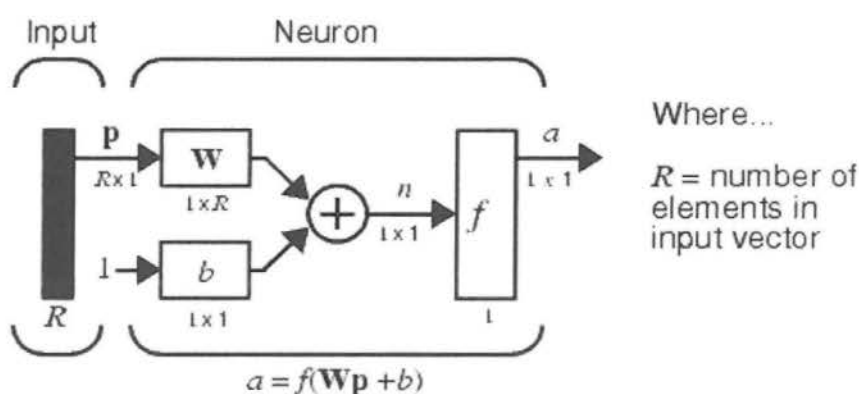
Where...

R = number of elements in input vector

Σχήμα 1.18

Βασικό μοντέλο τεχνητού νευρώνα.

Επειδή η παράσταση του παραπάνω σχήματος περιέχει πολλές λεπτομέρειες και στην περίπτωση μεγάλου μεγέθους δικτύων γίνεται ασύμφορη, στη MATLAB χρησιμοποιείται μια εναλλακτική παράσταση, αυτή του Σχήματος 1.19.

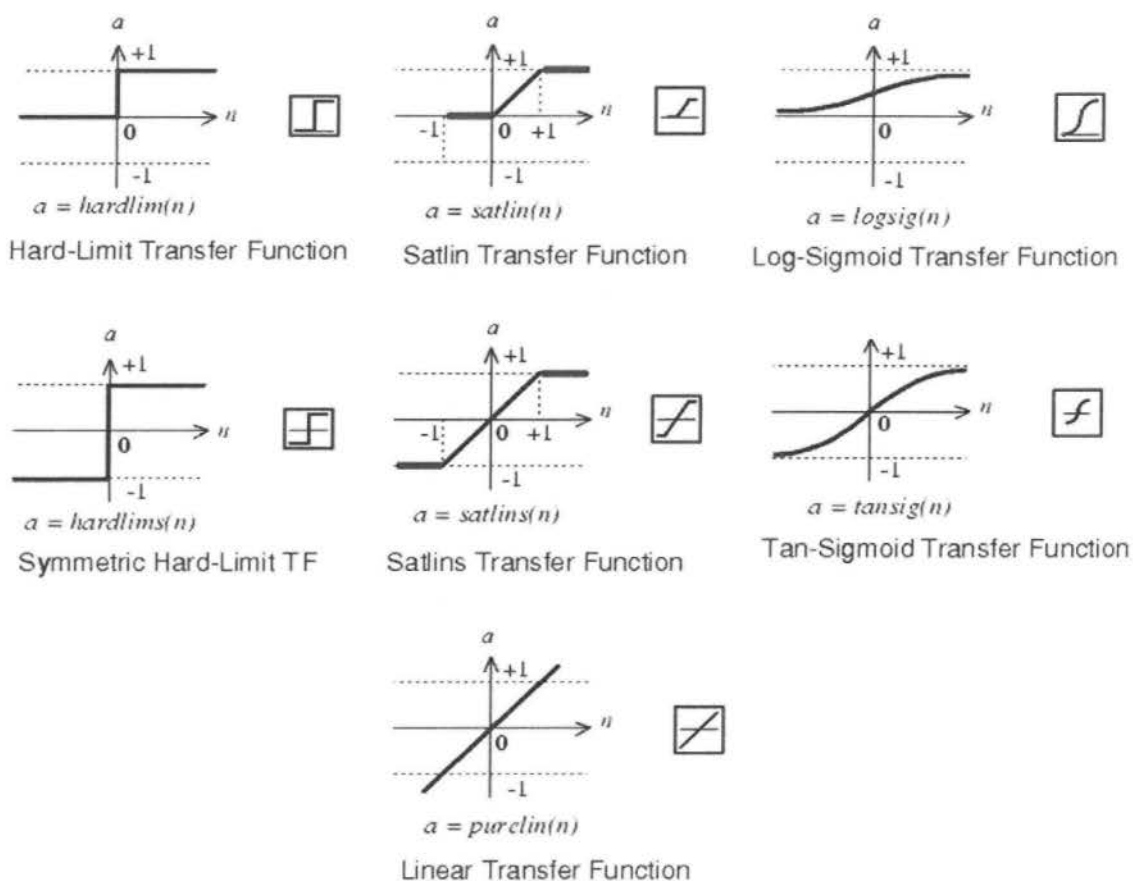


Σχήμα 1.19

Βασικό μοντέλο τεχνητού νευρώνα / Εναλλακτική παράσταση.

Εδώ το διάνυσμα εισόδου \mathbf{p} παριστάνεται με τη συμπαγή κάθετη ράβδο στα αριστερά. Η διάσταση του \mathbf{p} φαίνεται κάτω από τη ράβδο και εδώ είναι $R \times 1$. Το διάνυσμα \mathbf{p} πολλαπλασιάζεται με το διάνυσμα των βαρών \mathbf{W} (διαστάσεως $1 \times R$) και το άθροισμα του εσωτερικού αυτού γινομένου και της πόλωσης b (που είναι βαθμωτό μέγεθος, 1×1) δίνεται σαν είσοδος στη συνάρτηση ενεργοποίησης f . Η έξοδος θα είναι και πάλι $a = f(w_{1,1}p_1 + w_{1,2}p_2 + \dots + w_{1,R}p_R + b)$ (βαθμωτό μέγεθος, 1×1). Στην γενικότερη περίπτωση που έχουμε περισσότερες από μια εξόδους το a θα είναι διάνυσμα.

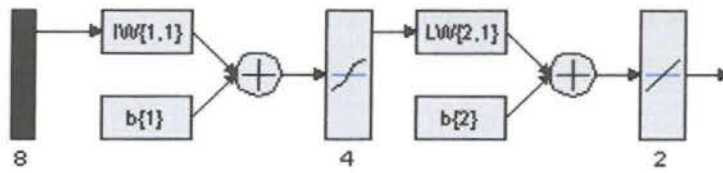
Για τις συναρτήσεις ενεργοποίησης δεν χρειάζεται να ειπωθεί κάτι διαφορετικό από αυτά που συζητήθηκαν στην παράγραφο 1.1.5. Απλά ξαναδίνουμε τις 6 συναρτήσεις των Σχημάτων 1.6 και 1.7 στο σχήμα που ακολουθεί, με τα ονόματα που χρησιμοποιεί για αυτές η MATLAB και μια εναλλακτική απεικόνιση της καθεμίας που δίνεται εντός των τετραγώνων – δεξιά της γραφικής τους παράστασης. Αυτή η απεικόνιση θα αντικαθιστά το γενικό f που είδαμε στα Σχήματα 1.18 και 1.19, ώστε να μπορούμε εύκολα να δούμε ποια ακριβώς συνάρτηση χρησιμοποιείται. Στο σχήμα αυτό παρουσιάζεται και μια συνάρτηση πέρα από αυτές που γνωρίσαμε στην υποενότητα 1.1.5, η γραμμική συνάρτηση μεταφοράς, η οποία χρησιμοποιείται συχνά στα MLP, σε περιπτώσεις που δεν θέλουμε να περιορίσουμε τις εξόδους των νευρώνων σε κάποιο διάστημα. Πρέπει να αναφερθεί ότι πέρα από αυτές τις 7 συναρτήσεις ενεργοποίησης η MATLAB προσφέρει έτοιμες για χρήση και αρκετές άλλες με τις οποίες ο χρήστης μπορεί να πειραματιστεί.



Σχήμα 1.20

Οι πλέον χρησιμοποιούμενες συναρτήσεις ενεργοποίησης στα πολυστρωματικά perceptrons.

Το feedforward MPL 8-4-2 που παρουσιάστηκε στο Σχήμα 1.9, αν θεωρήσουμε ότι οι κρυμμένοι νευρώνες χρησιμοποιούν σαν συνάρτηση ενεργοποίησης την $\text{tansig}()$ και οι νευρώνες εξόδου την $\text{purelin}()$, μπορεί με βάση τα προηγούμενα να εικονισθεί όπως στο Σχήμα 1.21:



Σχήμα 1.21

Feedforward MLP 8-4-2.

Με $IW\{1,1\}$ συμβολίζεται ο πίνακας βαρών του στρώματος εισόδου (Input Weight Matrix), δηλαδή τα βάρη των συνάψεων μεταξύ εισόδου και πρώτου κρυμμένου στρώματος, ενώ με $LW\{2,1\}$ συμβολίζεται ο πίνακας βαρών του κρυμμένου στρώματος (Layer Weight Matrix), δηλαδή τα βάρη των συνάψεων από το πρώτο στο δεύτερο κρυμμένο στρώμα (που εδώ βέβαια είναι το στρώμα εξόδου).

1.6.2 ΔΗΜΙΟΥΡΓΙΑ, ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΠΡΟΣΟΜΟΙΩΣΗ ΤΟΥ ΔΙΚΤΥΟΥ

Για τη δημιουργία, εκπαίδευση και προσομοίωση του δικτύου σε άγνωστα δεδομένα, έχουμε τη δυνατότητα να γράψουμε κώδικα καλώντας έτοιμες συναρτήσεις της MATLAB ή εναλλακτικά να χρησιμοποιήσουμε το nntool, ένα Graphical User Interface σχεδιασμένο με τέτοιο τρόπο ώστε και ο μη εξοικειωμένος χρήστης να μπορεί να πειραματιστεί με τα νευρωνικά δίκτυα και να αρχίσει να κατανοεί τον τρόπο λειτουργία τους. Μια σύντομη παρουσίαση του nntool θα γίνει στο κεφάλαιο που θα προσπαθήσουμε να λύσουμε το πρόβλημα του χρονοπρογραμματισμού εργασιών με χρήση των MLP. Εδώ παρουσιάζουμε εν συντομία της 3 βασικές εντολές που πρέπει να

γνωρίζει κανείς για να γράψει κώδικα χειρισμού MLP, και τις οποίες θα χρησιμοποιήσουμε επίσης στο επερχόμενο κεφάλαιο.

• **new** (Δημιουργία feed-forward δικτύου) Το πρώτο βήμα, πριν την εκπαίδευση του δικτύου είναι η δημιουργία του αντικειμένου του. Η συνάρτηση newff δημιουργεί ένα feed-forward δίκτυο. Παίρνει 4 ορίσματα σαν ορίσματα εισόδου και επιστρέφει το αντικείμενο του δικτύου. Το πρώτο όρισμα είναι ένας $R \times 2$ πίνακας με τις ελάχιστες και μέγιστες τιμές για κάθε ένα από τα R στοιχεία του διανύσματος εισόδου (δοθέντος του πίνακα εισόδων, έστω P , μπορούμε να καλέσουμε τη συνάρτηση `minmax(P)` η οποία αναλαμβάνει να βρει αυτές τις ακραίες τιμές). Το δεύτερο όρισμα είναι ένας πίνακας που περιέχει το μέγεθος του κάθε στρώματος (αριθμό νευρώνων ανά στρώμα). Το τρίτο όρισμα είναι ένα cell array που περιέχει τα ονόματα των συναρτήσεων ενεργοποίησης για καθένα από τα χρησιμοποιούμενα στρώματα και το τέταρτο όρισμα είναι το όνομα του αλγόριθμου εκπαίδευσης που θέλουμε να χρησιμοποιήσουμε. Για παράδειγμα η εντολή:

```
net = newff(minmax(P),[4,2],{'tansig','purelin'},'traingd')
```

δημιουργεί το αντικείμενο του feedforward δικτύου 8-4-2 που παρουσιάστηκε στα Σχήματα 1.9 και 1.21. Το μέγεθος του διανύσματος εισόδου «διαβάζεται» από τον πίνακα των εισόδων P , το πρώτο κρυμμένο στρώμα έχει 4 νευρώνες με συνάρτηση ενεργοποίησης την `tansig` και το δεύτερο στρώμα (στρώμα εξόδου) έχει 2 νευρώνες με συνάρτηση ενεργοποίησης την `purelin`. Ο αλγόριθμος εκπαίδευσης που ορίζουμε για αυτό το δίκτυο είναι ο `traingd` (ο οποίος θα περιγραφεί σε επόμενη ενότητα). Η `newff` εκτός από το ότι δημιουργεί το δίκτυο αρχικοποιεί και τις παραμέτρους του (βάρη και πολώσεις) σε μικρές τυχαίες τιμές.

Αντί της `newff`, μπορούμε να χρησιμοποιήσουμε εναλλακτικά την συνάρτηση `newcf` η οποία λειτουργεί με τον ίδιο τρόπο, με μόνη διαφορά ότι το δίκτυο που δημιουργεί είναι `cascade-forward`, πολυστρωματικό `perceptron` δηλαδή, με τους κόμβους εισόδου αλλά και του κάθε στρώματος να συνδέονται με τους κόμβους των επόμενων προς τα εμπρός στρωμάτων, και όχι μόνο με το γειτονικό.

- **train (Εκπαίδευση δικτύου)** Εφόσον το δίκτυο έχει δημιουργηθεί και οι παράμετροι του έχουν αρχικοποιηθεί μπορούμε να προχωρήσουμε στην εκπαίδευση του. Η εντολή `train` καλείται π.χ. ως:

```
[net,TR,Y,E,] = train(net,P,T)
```

Σαν ορίσματα εισόδου παίρνει το αντικείμενο του δικτύου που μόλις δημιουργήσαμε, τον πίνακα με τις εισόδους των παραδειγμάτων εκπαίδευσης, έστω `P`, και τον πίνακα με τους αντίστοιχους στόχους, έστω `T`. Τα ορίσματα που επιστρέφει είναι το εκπαιδευμένο δίκτυο, το `training record` `TR` με πληροφορίες για την απόδοση του δικτύου ανά εποχή και τους πίνακες `Y` και `E` με τις εξόδους που δίνει το εκπαιδευμένο δίκτυο στα παραδείγματα εκπαίδευσης και τα αντίστοιχα λάθη.

Πρέπει να σημειωθεί ότι η εντολή `train` χρησιμοποιείται για `batch training` του δικτύου και οι πίνακες `P` και `T` πρέπει να είναι μεγέθους (Αριθμός Εισόδων) \times (Αριθμός Παραδειγμάτων Εκπαίδευσης) και (Αριθμός Εξόδων) \times (Αριθμός Παραδειγμάτων Εκπαίδευσης) αντίστοιχα. Με άλλα λόγια τα στοιχεία των παραδειγμάτων εκπαίδευσης πρέπει να είναι τοποθετημένα στους πίνακες `P`, `T` κατά στήλες.

- **sim (Προσομοίωση δικτύου – Απάντηση σε άγνωστα δεδομένα)** Με δεδομένο ότι η εκπαίδευση έχει ολοκληρωθεί το δίκτυο μπορεί να χρησιμοποιηθεί για να μας δώσει απάντηση σε νέα δεδομένα. Αν υποθέσουμε ότι τα διανύσματα εισόδου των οποίων

την έξοδο θέλουμε να μάθουμε περιέχονται στον πίνακα `P_new`, κλήση της `sim(net,P_new)` επιστρέφει τις εξόδους-απαντήσεις του δικτύου.

1.6.3 ΑΛΓΟΡΙΘΜΟΙ ΕΚΠΑΙΔΕΥΣΗΣ

Αναλυτική παρουσίαση των διαφορετικών εκδοχών του back-propagation αλγόριθμου που προσφέρει η MATLAB ξεφεύγει από τα πλαίσια αυτής της εργασίας. Στην ενότητα αυτή κάνουμε μια γρήγορη αναφορά στα ιδιαίτερα χαρακτηριστικά τους και επικεντρωνόμαστε σε μια προσπάθεια θεωρητικής σύγκρισής της ταχύτητας και των απαιτήσεών τους σε μνήμη. Σε επόμενο κεφάλαιο θα προχωρήσουμε πιο πέρα κάνοντας πρακτική σύγκριση με βάση τα αποτελέσματα που θα δώσουν στο πρόβλημα του χρονοπρογραμματισμού εργασιών μεγέθους 6x6.

Οι αλγόριθμοι λοιπόν που προσφέρει σαν έτοιμες συναρτήσεις για την εκπαίδευση πολυστρωματικών perceptron η MATLAB είναι οι ακόλουθοι:

traingd, traingdm: Οι συναρτήσεις αυτές υλοποιούν τον βασικό αλγόριθμο ανάστροφης διάδοσης που περιγράψαμε στην υποενότητα 1.5.3. Η διαφορά της `traingdm` από την `traingd` είναι ότι συνυπολογίζει και τον όρο ορμής (momentum term). Ο `traingd` είναι γενικά πολύ αργός καθώς απαιτεί μικρές τιμές του ρυθμού μάθησης ώστε η διαδικασία να μην αποκλίνει. Ο `traingdm` είναι συνήθως γρηγορότερος αλλά και αυτός θεωρείται πολύ αργός για πρακτικά προβλήματα. Αυτοί οι δυο αλγόριθμοι προτείνεται να χρησιμοποιούνται μόνο σε περιπτώσεις που θέλουμε η εκπαίδευση του δικτύου να γίνει ανά πρότυπο (incremental training). Για batch training νευρωνικού δικτύου είναι προτιμότερο ο χρήστης να επιλέξει κάποια από τις παραλλαγές του back-propagation αλγόριθμου.

traingda, traingdx: Το ιδιαίτερο χαρακτηριστικό αυτών των αλγορίθμων είναι ο μεταβλητός ρυθμός μάθησης. Στον βασικό αλγόριθμο ανάστροφης διάδοσης (traingd ή traingdm) η τιμή του ρυθμού μάθησης διατηρείται σταθερή σε όλη τη διάρκεια της εκπαίδευσης με συνέπεια η απόδοση του αλγορίθμου να εξαρτάται σε μεγάλο βαθμό από την τιμή που θα επιλέξει ο χρήστης. Μεγάλη τιμή του ρυθμού μάθησης μπορεί να κάνει το σύστημα ασταθές ενώ από την άλλη μικρή τιμή οδηγεί σε πολλή αργή σύγκλιση. Αυτό ακριβώς το πρόβλημα προσπαθούν να λύσουν οι αλγόριθμοι traingda και traingdx. Μεταβάλλουν κατά τη διάρκεια της εκπαίδευσης το ρυθμό μάθησης προσπαθώντας να τον κρατήσουν στη μεγαλύτερη δυνατή τιμή που θα εξασφαλίζει ευσταθές σύστημα. Η διαφορά της traingdx από την traingda είναι ότι συνυπολογίζει και τον όρο ορμής.

trainrp: Στα πολυστρωματικά perceptrons συνήθως χρησιμοποιούμε σιγμοειδείς συναρτήσεις ενεργοποίησης για τα κρυμμένα στρώματα. Οι συναρτήσεις αυτές, όπως είδαμε και στα προηγούμενα, καλούνται και συναρτήσεις σύνθλιψης (squashing functions) καθώς περιορίζουν (συνθλίβουν) το επιτρεπόμενο πλάτος του σήματος εισόδου (θεωρητικά άπειρο) σε κάποια πεπερασμένη τιμή (συνήθως στο διάστημα $[0,1]$ ή $[-1,1]$). Η κλίση δηλαδή μιας σιγμοειδούς συνάρτησης πρέπει να προσεγγίζει το μηδέν όταν η είσοδος γίνεται πολύ μεγάλη. Αυτό δημιουργεί πρόβλημα όταν εκπαιδεύουμε ένα MLP με τον αλγόριθμο ανάστροφης διάδοσης, καθώς η μερική παράγωγος δ_k περιορίζεται σε ένα μικρό διάστημα με αποτέλεσμα οι αλλαγές στα βάρη και τις πολώσεις του δικτύου να είναι μικρές ακόμα και όταν βρισκόμαστε πολύ μακριά από τη βέλτιστη λύση. Το πρόβλημα αυτό μπορεί να ξεπεραστεί με χρήση της trainrp η οποία λαμβάνει υπόψη της μόνο το πρόσημο της μερικής παραγώγου και όχι το μέγεθός της.

Η trainrp είναι η γρηγορότερη συνάρτηση σε προβλήματα αναγνώρισης προτύπων και μια από τις συναρτήσεις που γενικότερα προτιμάται για μεγάλου μεγέθους δίκτυα καθώς οι απαιτήσεις της σε μνήμη είναι σχετικά μικρές. Αντίθετα, σε

προβλήματα προσέγγισης συναρτήσεων δεν αποδίδει καθόλου καλά. Πρέπει να τονισθεί επίσης ότι όσο το error goal που επιλέγουμε για το πρόβλημα μας μειώνεται τόσο μειώνεται και η απόδοση του αλγορίθμου.

traincgf, traincgp, traincgb, trainscg: Ο βασικός αλγόριθμος ανάστροφης διάδοσης ψάχνει τις βέλτιστες τιμές των παραμέτρων στην αντίθετη κατεύθυνση από αυτή της παραγώγου, που είναι και η κατεύθυνση κατά την οποία η συνάρτηση απόδοσης μειώνεται περισσότερο απότομα. Έχει αποδειχτεί ότι αυτό δεν οδηγεί απαραίτητα σε γρηγορότερη σύγκλιση. Το βασικό χαρακτηριστικό αυτής της «οικογένειας» συναρτήσεων (που καλούνται conjugate gradient algorithms) είναι ότι επιτυγχάνουν γρηγορότερη σύγκλιση μέσω της χρήσης line-search ρουτινών που παρέχονται έτοιμες από τη MATLAB. Ενδεικτικά αναφέρουμε τις Charalambous' Search (srchcha), Brent's Search (srchbre) και Golden Section Search (srchgol). Ο χρήστης σε περίπτωση που διαλέξει να εκπαιδεύσει το δίκτυο του με αυτούς τους αλγόριθμους μπορεί είτε να κρατήσει τις default επιλογές ανά αλγόριθμο είτε να πειραματιστεί ώστε να βρει ποια ταιριάζει καλύτερα στο προς επίλυση πρόβλημα.

Ιδιαίτερη αναφορά πρέπει να γίνει στον αλγόριθμο trainscg. Ο αλγόριθμος αυτός ανήκει στους conjugate gradient αλγόριθμους αλλά σε αντίθεση με όλους τους άλλους δεν χρησιμοποιεί line-search routine. Αυτό έχει σαν αποτέλεσμα να χρειάζεται συνήθως περισσότερες επαναλήψεις από τους άλλους μέχρι να συγκλίνει, με τον αριθμό των υπολογισμών ανά επανάληψη όμως να είναι πολύ μικρότερος. Ο trainscg αποδίδει ικανοποιητικότερα σε πολλών ειδών προβλήματα και κυρίως σε δίκτυα με μεγάλο αριθμό παραμέτρων (η MATLAB ρητά προτείνει για μεγάλου μεγέθους δίκτυα την trainscg και την trainrp). Η ταχύτητά του είναι ανάλογη της ταχύτητας του trainlm (που θα εξεταστεί στη συνέχεια) σε προβλήματα προσέγγισης συναρτήσεων και ανάλογη της ταχύτητας του trainrp σε προβλήματα αναγνώρισης προτύπων. Σε αντίθεση με τον trainrp μάλιστα δεν επηρεάζεται από την τιμή του στόχου και η απόδοση του δεν μειώνεται όσο μειώνεται αυτός.

trainbfg, trainoss: Σαν απάντηση στους conjugate gradient αλγόριθμους έρχεται η μέθοδος του Newton με τον παράγοντα διόρθωσης των βαρών να εξαρτάται από τον αντίστροφο του πίνακα Hessian (ο πίνακας Hessian περιέχει τις δεύτερες παραγώγους του δείκτη απόδοσης ως προς τα βάρη). Η μέθοδος αυτή συγκλίνει συχνά πολύ γρηγορότερα από τους conjugate gradient αλγόριθμους, η ανάγκη υπολογισμού του Hessian πίνακα όμως συνεπάγεται ιδιαίτερα αυξημένο κόστος υπολογισμού. Μια κλάση αλγορίθμων που βασίζονται στη λογική της μεθόδου Newton αλλά δεν απαιτούν τον υπολογισμό των δεύτερων παραγώγων είναι οι λεγόμενοι quasi-Newton αλγόριθμοι με δυο από αυτούς (trainbfg, trainoss) να παρέχονται έτοιμοι από τη MATLAB.

Η απόδοση των αλγορίθμων αυτών είναι παρόμοια με αυτή του trainlm (που εξετάζεται αμέσως μετά). Δεν χρειάζονται τόσο πολύ μνήμη αλλά οι απαιτούμενοι υπολογισμοί αυξάνονται γεωμετρικά με το μέγεθος του δικτύου, καθώς σε κάθε επανάληψη πρέπει να υπολογίζεται ο αντίστροφος ενός πίνακα.

trainlm: Σαν μια προσπάθεια γεφύρωσης των conjugate gradient και των quasi-Newton αλγορίθμων παρέχεται η συνάρτηση trainlm η οποία αντί του Hessian πίνακα χρησιμοποιεί μια προσέγγιση του, τον πίνακα Jacobian. Η συνάρτηση trainlm είναι και η default συνάρτηση της MATLAB για τη δημιουργία νευρωνικού δικτύου, είναι πολύ γρήγορη αλλά ταυτόχρονα πολύ απαιτητική σε μνήμη. Προτείνεται να χρησιμοποιείται μόνο για μικρού ή μεσαίου μεγέθους δίκτυα (λίγες εκατοντάδες παραμέτρων) ή στην περίπτωση που η μνήμη του υπολογιστικού μας συστήματος δεν αποτελεί περιοριστικό παράγοντα. Να σημειωθεί τέλος ότι είναι ο αλγόριθμος που αποδίδει καλύτερα σε πολύ μικρές τιμές του error goal

Με βάση τα όσα συζητήθηκαν σε αυτό το κεφάλαιο είναι σαφές ότι το να γνωρίζει κανείς εκ των προτέρων ποιο νευρωνικό δίκτυο είναι αυτό που θα ταιριάζει

καλύτερα σε μια δεδομένη εφαρμογή είναι αν όχι αδύνατο, εξαιρετικά δύσκολο. Οι διαφορετικές επιλογές που μπορούν να γίνουν αφορούν την αρχιτεκτονική του δικτύου, τον αριθμό των στρωμάτων και των νευρώνων με τις συναρτήσεις ενεργοποίησης που αυτοί θα χρησιμοποιούν, το σύνολο των παραδειγμάτων εκπαίδευσης, τον τρόπο παράστασης των παραδειγμάτων αυτών, τον αλγόριθμο εκπαίδευσης με τις επιλογές των παραμέτρων του, και πολλές άλλες. Το όλο πρόβλημα αποτελεί πραγματική σχεδιαστική πρόκληση στην οποία θα προσπαθήσουμε να αντεπεξέλθουμε ώστε να λύσουμε όσο το δυνατόν αποδοτικότερα το πρόβλημα του χρονοπρογραμματισμού εργασιών. Περισσότερα πάνω σε όλα αυτά θα αναφερθούν λοιπόν σε επόμενο κεφάλαιο.

2. ΤΕΧΝΗΤΗ ΑΙΣΘΗΣΗ ΚΑΙ ΕΦΑΡΜΟΓΕΣ

2.1 ΕΙΣΑΓΩΓΗ

Η προσπάθεια κατασκευής ευφυών μηχανών που διαθέτουν την ικανότητα να κατανοούν τον κόσμο αποτελεί πάντα μια ενδιαφέρουσα ενασχόληση των επιστημόνων πληροφορικής. Ένα μεγάλο μέρος του ανθρωπίνου εγκεφάλου είναι αφιερωμένο στη λειτουργία των αισθήσεων. Οι άνθρωποι λύνουν τα προβλήματα αίσθησης χωρίς προσπάθεια, παρ' όλο που λίγοι κατανοούν το μηχανισμό της. Ο Alan Turing, ένας από τους πατέρες των Ψηφιακών Υπολογιστών και του πεδίου της Τεχνητής Νοημοσύνης, πίστευε ότι είναι δυνατό ο ψηφιακός υπολογιστής να αποκτήσει ευφυΐα και την ικανότητα να κατανοεί το περιβάλλον. Αν και αυτός ο στόχος δεν έχει επιτευχθεί ακόμα σε βαθμό συγκρίσιμο με του ανθρώπου, έχει επιτευχθεί εκπληκτική πρόοδος. Στο μάθημα αυτό εξετάζεται το γνωστικό αντικείμενο της Τεχνητής Αίσθησης με μια σύγχρονη και εμπειριστατωμένη προσέγγιση.

2.2 ΑΝΙΧΝΕΥΣΗ ΑΚΜΩΝ

Μια προσέγγιση για τον προσδιορισμό των συνόρων αντικειμένων σε μια εικόνα έγκειται στην εξέταση κάθε εικονοστοιχείου και της άμεσης περιοχής του για να αποφασιστεί αν το εικονοστοιχείο ανήκει πραγματικά στο σύνορο ενός αντικειμένου. Τέτοια εικονοστοιχεία προσδιορίζονται ως εικονοστοιχεία ακμών (edge pixels). Οι παρακάτω αλγόριθμοι προσφέρουν διαφορετικούς τρόπους για την υλοποίηση αυτής της διαδικασίας.

Ο ανιχνευτής ακμών του Roberts: Είναι ένας τοπικός διαφορικός τελεστής που δέχεται μια εικόνα εισόδου $A(m,n)$ και παράγει την εικόνα εξόδου:

$$B(m,n) = \{[\sqrt{A(m,n)} - \sqrt{A(m+1,n+1)}]^2 + [\sqrt{A(m+1,n)} - \sqrt{A(m,n+1)}]^2\}^{\frac{1}{2}}$$

Οι εσωτερικές τετραγωνικές ρίζες κάνουν την διαδικασία παρόμοια με την επεξεργασία που λαμβάνει χώρα στο ανθρώπινο σύστημα όρασης.

-1	0
0	1

0	-1
1	0

Ο
ανιχνευ
τής
ακμών
του

Sobel: Χρησιμοποιεί τους συνελκτικούς πυρήνες των επόμενων πινάκων. Κάθε εικονοστοιχείο συνελίσσεται και με τους δύο πυρήνες. Ο ένας πυρήνας έχει μέγιστη απόκριση σε μια κατακόρυφη ακμή, ενώ ο άλλος σε μια οριζόντια. Τελικά χρησιμοποιείται η μέγιστη τιμή μεταξύ των δύο συνελίξεων και το αποτέλεσμα είναι μια εικόνα του μεγέθους των ακμών.

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

Ο ανιχνευτής ακμών του Prewitt: Η διαδικασία είναι ακριβώς η ίδια όπως και με τον ανιχνευτή ακμών του Sobel, αλλά χρησιμοποιούνται οι επόμενοι δύο πυρήνες:

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

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

Ο ανιχνευτής ακμών του Kirsch: Χρησιμοποιεί 8 συνελκτικούς πυρήνες (βιβλίο «Ανάλυση εικόνας, σελίδα 92). Κάθε εικονοστοιχείο συνελίσσεται και με τις οκτώ μάσκες. Κάθε μάσκα έχει μέγιστη απόκριση σε ακμές κάποιου συγκεκριμένου

προσανατολισμού. Επιλέγεται το ολικό μέγιστο και τελικά δημιουργείται μια εικόνα του μεγέθους των ακμών

Χρησιμοποιώντας τις 4 προαναφερθέντες μεθόδους, μπορούμε να δημιουργήσουμε ένα αρχείο που θα χρησιμοποιεί κάθε μία από αυτές για να εντοπίζει ακμές. Ο κώδικας του αρχείου αυτού (με πλήρη επεξήγησή της κάθε γραμμής) βρίσκεται παρακάτω.

```
function edge_detection (image )  
  
pic = imread ( image );  
  
pic = im2bw ( pic );  
  
pic = double ( pic );  
  
image1 = edge(pic,'prewitt');  
  
image2 = edge(pic,'roberts');  
  
image3 = edge(pic,'sobel');  
  
figure, imshow(image1);title('Prewitt');  
  
figure, imshow(image2);title('Roberts');  
  
figure, imshow(image3);title('Sobel');  
  
[image4]=edge_detection_kirsch(pic);  
  
figure;imshow(image4);title('Kirsch');  
  
end
```

Πρόκειται για το αρχείο `edge_detection` που περιέχει την ομώνυμη συνάρτηση. Η `edge_detection` παίρνει ως όρισμα το όνομα μιας εικόνας. Καταρχάς φορτώνεται η εικόνα στην μεταβλητή `pic` με την εντολή

```
pic = imread ( image );
```

Έπειτα με τις εντολές

```
pic = im2bw ( pic );
```

```
pic = double ( pic );
```

η εικόνα μετατρέπεται σε ασπρόμαυρη (για να μπορέσει να δουλέψει ο αλγόριθμος) με την εντολή `im2bw` και τα στοιχεία της γίνονται τύπου `double` με την εντολή `double`.

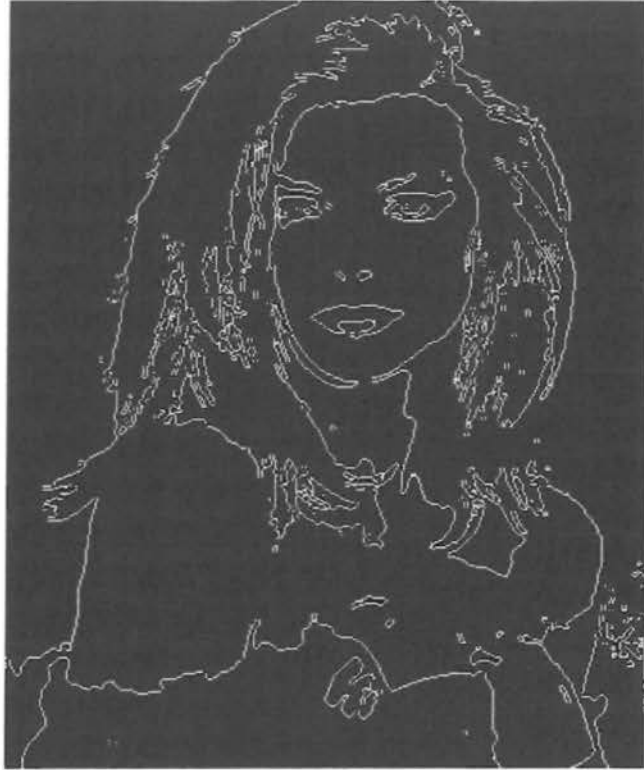
Έπειτα αρχίζει το κομμάτι της ανίχνευσης ακμών.

Οι αλγόριθμοι ανίχνευσης ακμών είναι ήδη υλοποιημένοι στην βιβλιοθήκη συναρτήσεων του Matlab. Η εντολή που χρησιμοποιείται είναι η `edge`, η οποία παίρνει

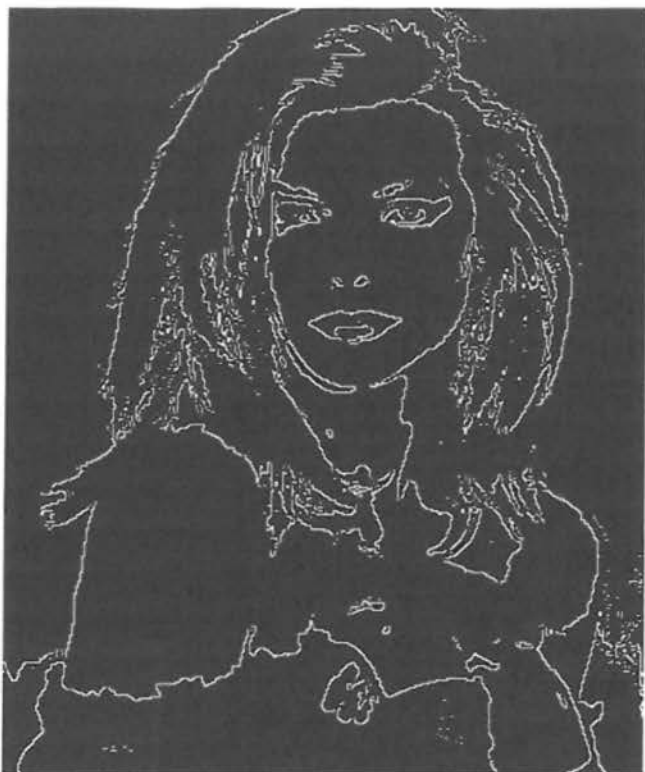
ως παράμετρο την εικόνα και έπειτα τον αλγόριθμο που θα χρησιμοποιηθεί (παράμετροι sobel, prewitt, roberts). Η εντολή edge επιστρέφει μια δυαδική εικόνα όπου οι άσσοι αντιπροσωπεύουν τις ακμές και τα μηδενικά τα εσωτερικά σημεία. Στο παραπάνω κομμάτι κώδικα φαίνεται η δοκιμή των αλγορίθμων αυτών. Επάνω στην μεταβλητή pic ενεργούνται οι τρεις αυτοί αλγόριθμοι με την χρήση της εντολής edge και τελικά εμφανίζονται οι εικόνες με την εντολή figure, imshow.



Η αρχική εικόνα



Το αποτέλεσμα του αλγόριθμου Prewitt



Το αποτέλεσμα του αλγορίθμου Roberts



Το αποτέλεσμα του αλγορίθμου Sobel

Ο αλγόριθμος Kirsch: Για την υλοποίηση του αλγορίθμου `kirsch` χρησιμοποιείται το παρακάτω βοηθητικό αρχείο. Πρόκειται για μια συνάρτηση που χρησιμοποιεί τις μήτρες του `kirsch`. Η συνάρτηση ονομάζεται `edge_detection_kirsch` και επιστρέφει την μεταβλητή `e`.

```
function [e] = edge_detection_kirsch(x)
```

```
k1 = [5 5 5; -3 0 -3; -3 -3 -3];
```

```
k2 = [-3 -3 5; -3 0 5; -3 -3 -3];
```

```
k3 = [-3 5 5; -3 0 5; -3 -3 5];
```

```
k4 = [-3 -3 -3; -3 0 5; -3 5 5];
```

```
k5 = [-3 -3 -3; -3 0 -3; 5 5 5];
```

```
k6 = [-3 -3 -3; 5 0 -3; 5 5 -3];
```

```
k7 = [5 -3 -3; 5 0 -3; 5 -3 -3];
```

```
k8 = [5 5 -3; 5 0 -3; -3 -3 -3];
```

Με αυτές τις γραμμές κώδικα δημιουργούνται οι 8 πρώτες συνελκτικοί πίνακες (ή πυρήνες) που χρησιμοποιεί ο αλγόριθμος του kirsch. Έπειτα οι τιμές της εικόνας γίνονται τύπου double για να μπορούν να επεξεργαστούν.

```
x = double(x);
```

```
x = double(x);
```

```
C1 = conv2 ( x , k1 );
```

```
C2 = conv2 ( x , k2 );
```

```
C3 = conv2 ( x , k3 );
```

```
C4 = conv2 ( x , k4 );
```

```
C5 = conv2 ( x , k5 );
```

```
C6 = conv2 ( x , k6 );
```

```
C7 = conv2 ( x , k7 );
```

```
C8 = conv2 ( x , k8 );
```

```
max_temp = max ( C1, C2 );
```

```
max_temp = max ( max_temp, C3 );
```

```
max_temp = max ( max_temp, C4 );
```

```
max_temp = max ( max_temp, C5 );
```

```
max_temp = max ( max_temp, C6 );
```

```
max_temp = max ( max_temp, C7 );
```

```
max_temp = max ( max_temp, C8 );
```

```
e = max_temp;
```

```
end
```

Με τη βοήθεια της συνάρτησης `conv2`, πραγματοποιούμε τη συνέλιξη ανάμεσα στην εικόνα και σε κάθε πίνακα ξεχωριστά. Επιλέγεται το ολικό μέγιστο από τις συνελίξεις και δημιουργείται μια εικόνα του μεγέθους των ακμών.

Kirsch



το αποτέλεσμα του αλγορίθμου του Kirsch

2.3 ΥΦΗ - ΑΛΓΟΡΙΘΜΟΣ ΤΟΥ LAWS

Ο αλγόριθμος του Laws, αποτελεί μια προσέγγιση για την δημιουργία χαρακτηριστικών υφής. Σύμφωνα με τον συγκεκριμένο αλγόριθμο, για τον υπολογισμό της ενέργειας της υφής χρησιμοποιείται ένα σύνολο συνελκτικών μασκών που αναπαριστάνεται με ένα διάνυσμα αριθμών για κάθε εικονοστοιχείο.

Αρχείο LawsTexture.m

Το αρχείο .m που υλοποιήθηκε ο αλγόριθμος του Laws ονομάζεται LawsTexture.m και περιγράφεται παρακάτω

$L5=[1,4,6,4,1];$

$E5=[-1,-2,0,2,1];$

$S5=[-1,0,2,0,-1];$

$R5=[1,-4,6,-4,1];$

Με τις 4 αυτές εντολές δημιουργούνται τα 4 διανύσματα που χρειαζόμαστε για να υλοποιηθούν οι μάσκες του Laws. Το διάνυσμα L5 είναι για τον κεντραρισμένο τοπικό μέσο, το E5 ανιχνεύει ακμές, το S5 ανιχνεύει κηλίδες και το R5 ανιχνεύει κυματισμούς.

$L5E5=L5'*E5;$

$E5L5=E5'*L5;$

$L5R5=L5'*R5;$

$R5L5=R5'*L5;$

$E5S5=E5'*S5;$

$S5E5=S5'*E5;$

$S5S5=S5'*S5;$

$R5R5=R5'*R5;$

$L5S5=L5'*S5;$

$S5L5=S5'*L5;$

$E5E5=E5'*E5;$

$E5R5=E5'*R5;$

```
RSE5=R5'*E5;
```

```
S5R5=S5'*R5;
```

```
R5S5=R5'*S5;
```

Με τις παραπάνω 15 γραμμές κώδικα υπολογίζεται κάθε φορά και η αντίστοιχη μάσκα που είναι το εξωτερικό γινόμενο ζευγών διανυσμάτων. Οι μεταβλητές έχουν το όνομα των 2 διανυσμάτων που χρησιμοποιούνται (πχ. αν χρησιμοποιηθούν τα διανύσματα S5 και E5 το όνομα της μεταβλητής θα είναι S5E5). Για τον υπολογισμό των масκών χρησιμοποιείται ο ανάστροφος του πρώτου πίνακα (πχ ο A' είναι ο ανάστροφος του A) και πολλαπλασιάζεται με τον άλλον πίνακα (πολλαπλασιασμός πινάκων).

```
Image=imread('michelle.png');
```

```
subplot(5,5,1);
```

```
imshow(Image);
```

με την εντολή imread διαβάζεται η εικόνα. Η εντολή subplot χωρίζει την εικόνα σε 25 (5x5) κελιά, και εμφανίζει στο κελί no.1 (την τρίτη παράμετρο) την πρώτη εικόνα.

Με την εντολή imshow εμφανίζεται η εικόνα.

```
d=5;
```

```
X = preprocess(Image, d);
```

Έπειτα δηλώνεται το μέγεθος του παραθύρου που θα χρησιμοποιηθεί για την προεπεξεργασία της εικόνας (την καταπίεση των επιδράσεων από τον τρόπο φωτισμού). Αυτό επιτυγχάνεται κινώντας ένα μικρό παράθυρο πάνω στην εικόνα και αφαιρώντας τον τοπικό μέσο από κάθε εικονοστοιχείο. Το πλάτος του παραθύρου αποθηκεύεται στην μεταβλητή d που εδώ παίρνει την τιμή 5. Έπειτα χρησιμοποιείται η συνάρτηση preprocess, η οποία είναι υλοποιημένη σε ξεχωριστό αρχείο, και χρησιμοποιείται ακριβώς για την προεπεξεργασία της εικόνας παίρνοντας ως παραμέτρους την εικόνα Image και το πλάτος παραθύρου d. Το αρχείο preprocess.m είναι το εξής:

Αρχείο preprocess.m

```
function prepimg=preprocess(Image,d)
```

```
[m n]=size(Image);
```

```

prepimg=double(Image);

for j=1:d:m

    if j+d-1>m

        break;

    end

    for i=1:d:n

        if i+d-1>n

            break;

        end

        prepimg(j:j+d-1,i:i+d-1)=prepimg(j:j+d-1,i:i+d-1)- mean(mean(prepimg(j:j+d-1,i:i+d-1)));

    end

end

```

Το αρχείο `preprocess.m` περιέχει μια συνάρτηση με όνομα `preprocess`, η οποία παίρνει ως ορίσματα μια εικόνα `Image` και ένα πλάτος παραθύρου `d` και επιστρέφει την μεταβλητή `prepimg`. Αρχικά, υπολογίζονται το πλάτος και το ύψος της εικόνας `Image` και αποθηκεύονται στις μεταβλητές `m` και `n` αντίστοιχα με την εντολή `[m n]=size(Image);`. Έπειτα η μεταβλητή `prepimg` παίρνει τις τιμές της εικόνας `Image`. Ύστερα ξεκινάει ένας επαναληπτικός βρόγχος για `j` από 1 έως το πλάτος της εικόνας και με βήμα το πλάτος τους παραθύρου. Γίνεται κάθε φορά ένας έλεγχος εάν το παράθυρο έρευνας έχει υπερβεί το πλάτος της εικόνας (εντολή `if j+d-1>m break;`). Έπειτα ακολουθεί δεύτερος, εμφωλιασμένος βρόγχος από 1 έως το ύψος της εικόνας με βήμα πάλι `d`, και επίσης γίνεται έλεγχος εάν το παράθυρο έρευνας ξεπερνάει το ύψος της εικόνας. Εάν το παράθυρο βρίσκεται μέσα στην εικόνα, τότε σε κάθε επανάληψη αφαιρείται από το κάθε σημείο που βρίσκεται το παράθυρο ο τοπικός μέσος. Αυτό γίνεται με την εντολή

```

prepimg(j:j+d-1,i:i+d-1)=prepimg(j:j+d-1,i:i+d-1)- mean(mean(prepimg(j:j+d-1,i:i+d-1)));

```

Η `prepimg` κρατάει τις τιμές της εικόνας, και αφαιρούνται από αυτήν (για τα στοιχεία από `j` έως `j+d-1` και από `i` έως `i+d-1`, δηλαδή στον χώρο του παραθύρου) ο τοπικός μέσος των ίδιων εικονοστοιχείων, που βρίσκεται από τις εντολές `mean(mean(...))`.

Στην συνέχεια του αρχείου LawsTextures, το αποτέλεσμα της συνάρτησης preprocess αποθηκεύεται στην μεταβλητή X (εντολή X=preprocess(Image, d);).

```
for kernel = 1:15
```

```
disp(kernel);
```

Έπειτα ξεκινάει ένας επαναληπτικός βρόγχος ο οποίος θα τρέξει 15 φορές (μία για κάθε μάσκα που πρέπει να δημιουργηθεί). Χρησιμοποιείται ο μετρητής kernel ο οποίος εκτυπώνεται και κάθε φορά στο παράθυρο διαλόγου. Για κάθε επανάληψη επιλέγεται και η διαφορετική case της εντολής switch, και κατασκευάζεται και κάθε μία από τις 15 μάσκες. Οι μάσκες κατασκευάζονται παίρνοντας την προεπεξεργασμένη εικόνα X και χρησιμοποιώντας την εντολή

```
fk = conv2(double(X),double(L5E5));
```

Η εντολή conv2 εφαρμόζει συνέλιξη ανάμεσα στα δύο ορίσματα που δέχεται ως είσοδος. Το αποτέλεσμα είναι ένας πίνακας διαστάσεων όσο το άθροισμα του μεγέθους κάθε διάστασης κάθε διανύσματος μείον 1 ($X_a + X_b - 1$, $Y_a + Y_b - 1$). Ουσιαστικά με την εντολή αυτή δημιουργείται και κάθε μάσκα, αφού το κάθε διάνυσμα από τα 15 εμπλέκεται με την εικόνα X.

```
[m n] = size(fk);
```

```
energyMap = Ek(fk, m, n);
```

Στη συνέχεια, υπολογίζεται σε κάθε επανάληψη και ο χάρτης ενέργειας υφής. Ο υπολογισμός πραγματοποιείται με βοήθεια της συνάρτησης energyMap η οποία βρίσκεται σε ξεχωριστό αρχείο και περιγράφεται αναλυτικά παρακάτω. Η συνάρτηση αυτή δέχεται ως ορίσματα την κάθε φιλτραρισμένη εικόνα fk που δημιουργείται κάθε φορά και τα μεγέθη των διαστάσεων της εικόνας αυτής, που βρίσκονται από την εντολή size και αποθηκεύονται στις μεταβλητές m και n για το πλάτος και το ύψος αντίστοιχα.

Αρχείο ek.m

```
function ek=Ek(fk,m,n)
```

```
for i=8:m-7
```

```
for j=8:n-7
```

```

ek(i,j)=sum(sum(abs(fk(i-7:i+7,j-7:j+7))));

end

end

```

το αρχείο ek.m αποτελείται από την συνάρτηση Ek που παίρνει ως ορίσματα αυτά που περιγράφηκαν παραπάνω και επιστρέφει τον κάθε χάρτη ενέργειας υφής, ο οποίος είναι μια πλήρης εικόνα που αναπαριστά την εφαρμογή της κ-οστής μάσκας στην αρχική εικόνα (μεταβλητή ek). Έπειτα το κύριο κομμάτι κώδικα του αρχείου είναι ουσιαστικά η υλοποίηση σε προγραμματιστική γλώσσα του μαθηματικού τύπου υπολογισμού της ενέργειας υφής που βρίσκεται στο βιβλίο «Ανάλυση Εικόνας» σελίδα 122.

Η συνάρτηση Ek επιστρέφει κάθε φορά τον χάρτη ενέργειας υφής, και έπειτα με τις εντολές

```

subplot(5, 5, kernel + 1);

imshow(energyMap, []);

```

εμφανίζεται σε κάθε ένα από τα 25 κελιά στα οποία χωρίστηκε η εικόνα στην αρχή και ο αντίστοιχος χάρτης ενέργειας υφής. Με κάθε επανάληψη ο μετρητής kernel προχωράει κατά ένα και πηγαίνει στο επόμενο κελί, ενώ με την εντολή imshow εμφανίζεται η εικόνα της energyMap.

Έπειτα ξεκινάει ένας βρόγχος ο οποίος τρέχει 9 φορές και υπολογίζει κάθε φορά τον τελικό χάρτη ενέργειας υφής του Laws.

```

for counter = 1:9

disp(counter);

switch counter

case 1

[m n] = size(L5E5fi);

energyMapL5E5 = Ek(L5E5fi, m, n);

[m n] = size(E5L5fi);

energyMapE5L5 = Ek(E5L5fi, m, n);

```

```
finalEnergyMap = (energyMapL5E5 + energyMapE5L5) / 2;
```

case 2

```
[m n] = size(L5R5fi);
```

```
energyMapL5R5 = Ek(L5R5fi, m, n);
```

```
[m n] = size(R5L5fi);
```

```
energyMapR5L5 = Ek(R5L5fi, m, n);
```

```
finalEnergyMap = (energyMapL5R5 + energyMapR5L5) / 2;
```

case 3

```
[m n] = size(E5S5fi);
```

```
energyMapE5S5 = Ek(E5S5fi, m, n);
```

```
[m n] = size(S5E5fi);
```

```
energyMapS5E5 = Ek(S5E5fi, m, n);
```

```
finalEnergyMap = (energyMapE5S5 + energyMapS5E5) / 2;
```

case 4

```
[m n] = size(S5S5fi);
```

```
energyMapS5S5 = Ek(S5S5fi, m, n);
```

```
finalEnergyMap = energyMapS5S5;
```

case 5

```
[m n] = size(R5R5fi);
```

```
energyMapR5R5 = Ek(R5R5fi, m, n);
```

```
finalEnergyMap = energyMapR5R5;
```

case 6

[m n] = size(L5S5fi);

energyMapL5S5 = Ek(L5S5fi, m, n);

[m n] = size(S5L5fi);

energyMapS5L5 = Ek(S5L5fi, m, n);

finalEnergyMap = (energyMapL5S5 + energyMapS5L5) / 2;

case 7

[m n] = size(E5E5fi);

energyMapE5E5 = Ek(E5E5fi, m, n);

finalEnergyMap = energyMapE5E5;

case 8

[m n] = size(E5R5fi);

energyMapE5R5 = Ek(E5R5fi, m, n);

[m n] = size(R5E5fi);

energyMapR5E5 = Ek(R5E5fi, m, n);

finalEnergyMap = (energyMapE5R5 + energyMapR5E5) / 2;

case 9

[m n] = size(S5R5fi);

energyMapS5R5 = Ek(S5R5fi, m, n);

[m n] = size(R5S5fi);

energyMapR5S5 = Ek(R5S5fi, m, n);

```
finalEnergyMap = (energyMapSSR5 + energyMapR5SS) / 2;
```

```
end
```

```
subplot(5, 5, counter + 16);
```

```
imshow(finalEnergyMap, []);
```

```
title ( counter );
```

```
end
```

Σε κάθε ένα από τα κομμάτια κώδικα που τρέχουν για κάθε επανάληψη υπολογίζονται οι διαστάσεις της εικόνας με την εντολή

```
[m n] = size(LSE5fi);
```

Και έπειτα ο χάρτης ενέργειας υφής

```
energyMapL5E5 = Ek(LSE5fi, m, n);
```

Αυτό γίνεται μία ή δύο φορές, ανάλογα με το τι χρειάζεται ο συγκεκριμένος χάρτης που κατασκευάζεται. Μόλις γίνει αυτό, τότε υπολογίζεται ο τελικός χάρτης ο οποίος αποτελείται από τον μέσο όρο των δύο προκατασκευασθέντων χαρτών

```
finalEnergyMap = (energyMapL5E5 + energyMapE5LS) / 2;
```

Στο τέλος του βρόγχου οι εντολές

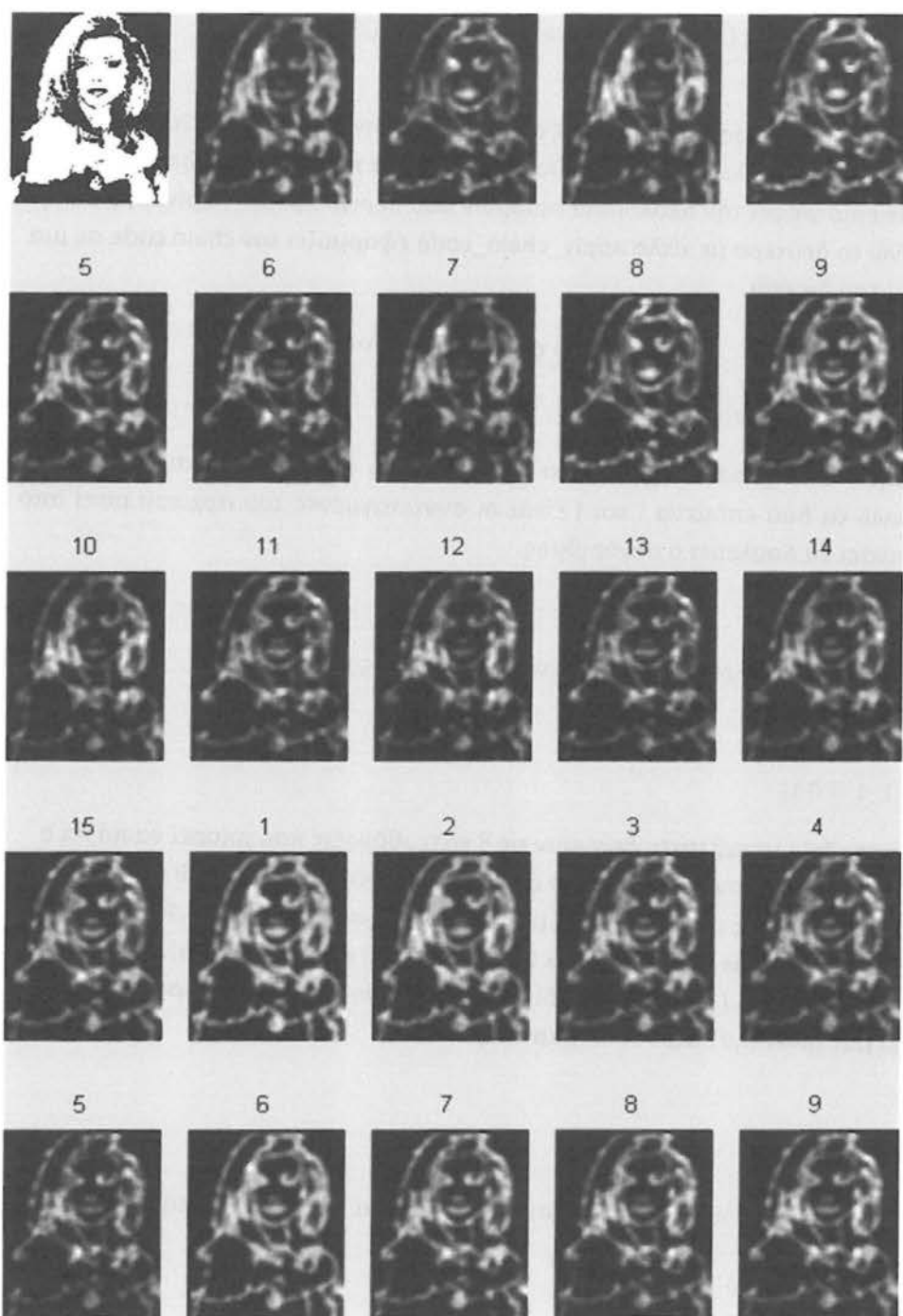
```
subplot(5, 5, counter + 16);
```

```
imshow(finalEnergyMap, []);
```

```
title ( counter );
```

εμφανίζουν τις εικόνες.

Παρακάτω παρουσιάζονται οι εικόνες που προκύπτουν από τον αλγόριθμο Laws. Η πρώτη είναι η κανονική εικόνα, οι επόμενες 15 παρουσιάζουν τις εικόνες αφού έχουν συνελιχθεί με τις μάσκες και οι τελευταίες 9 παρουσιάζουν τις εικόνες αφού έχουν συνελιχθεί με τους 9 ενεργειακούς χάρτες υφής του Laws.



2.4 ΑΝΑΛΥΣΗ ΣΧΗΜΑΤΟΣ-ΑΛΓΟΡΙΘΜΟΣ CHAIN CODE

Για την υλοποίηση της ανάλυσης σχήματος με την μέθοδο του chain code και differential chain code θα χρησιμοποιηθούν δύο αρχεία matlab. Το πρώτο με τίτλο chain_code επιστρέφει την ακολουθία αριθμών που περιγράφει το περίγραμμα της εικόνας, ενώ το δεύτερο με τίτλο apply_chain_code εφαρμόζει τον chain code σε μια εικόνα που του δίνεται.

Αρχείο chain_code.m

```
function code=chain_code(x,i,j)
```

Η συνάρτηση chain_code παίρνει τρία ορίσματα. Το πρώτο (x) είναι ο πίνακας της εικόνας, ενώ τα δύο επόμενα i και j είναι οι συντεταγμένες του αρχικού pixel από το οποίο ξεκινάει να δουλεύει ο αλγόριθμος.

```
[M,N]=size(x);
```

Στις μεταβλητές M και N αποθηκεύονται οι διαστάσεις του πίνακα X.

```
di=[0 -1 -1 -1 0 1 1 1];
```

```
dj=[1 1 0 -1 -1 -1 0 1];
```

Οι δύο παραπάνω μεταβλητές δείχνουν τις 8 κατευθύνσεις που μπορεί να πάρει ο αλγόριθμος και σχετίζονται αντίστοιχα οι τιμές των δύο μεταβλητών di και dj. Για παράδειγμα οι πρώτες τιμές των di, dj μας δίνουν το ζευγάρι τιμών [0,1] που δηλώνει την κατεύθυνση **επάνω** (το di για τον x άξονα και τα dj για τον y άξονα). Το δεύτερο ζευγάρι τιμών είναι το [-1,1] που μας δίνει την κατεύθυνση **πάνω-αριστερά**, το τρίτο το [-1,0] που μας δίνει την κατεύθυνση **αριστερά** κλπ.

```
first=1
```

```
val=x(i,j);
```

Διαβάζεται η τιμή του αρχικού pixel και αποθηκεύεται στην μεταβλητή val.

```
n=0;
```

το n κρατάει το μέγεθος της αλυσίδας

```
h1=0; w1=0;
```

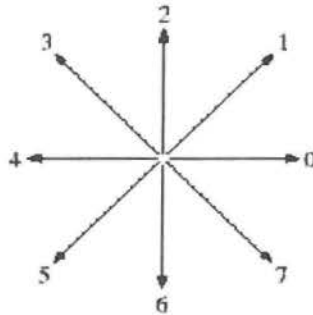
οι παραπάνω μεταβλητές χρησιμοποιούνται για να ελέγξουν την ροή του κώδικα

```
q = i; r = j;
```

οι μεταβλητές q και r θα αποθηκεύουν το τρέχον pixel (τώρα παίρνουν τις τιμές του αρχικού Pixel i, j).

```
lastdir = 4;
```

Στην μεταβλητή lastdir αποθηκεύεται η τελευταία κατεύθυνση κάθε φορά (αρχικά, η κατεύθυνση 4, η οποία έχει κατεύθυνση προς τα αριστερά).



Οι 8 κατευθύνσεις

```
flag=1;
```

Η flag είναι μεταβλητή ελέγχου ροής του κώδικα.

```
while ((q~=i) || (r~=j) || (flag==1))
```

Στο σημείο αυτό, ξεκινάει ένας βρόγχος που επαναλαμβάνεται για όσο δεν έχουμε επιστρέψει στο αρχικό εικονοστοιχείο ή όσο η μεταβλητή flag έχει την τιμή 1

```
first=0;
```

```
m = 0;
```

```
dii = -1;
```

η m είναι η μεταβλητή που καθορίζει εάν βρέθηκε γειτονικό πίξελ

```
for ii=lastdir+1:lastdir+7
```

```
jj = mod(ii,8);
```

```
if ((di(jj+1)+q)<=M) && ((dj(jj+1)+r)<=N) && (di(jj+1)+q>0) && (dj(jj+1)+r)>0
```

```
if ( x(di(jj+1)+q,dj(jj+1)+r) == val)
```

```
dii = jj; m = 1;
```

```
break;
```

Στον παραπάνω βρόγχο, εξετάζονται με την σειρά όλες οι γειτονικές κατευθύνσεις. Ξεκινάμε από την κατεύθυνση 7 και συνεχίζουμε κατά τη φορά των δεικτών του ρολογιού μέχρι την τελευταία κατεύθυνση. Ο μετρητής jj εφαρμόζεται στα di και dj και μας δίνει κάθε φορά την νέα κατεύθυνση. Η γραμμή

```
if ((di(jj+1)+q)<=M) && ((dj(jj+1)+r)<=N) && (di(jj+1)+q>0) && (dj(jj+1)+r)>0
```

αποτελεί έναν έλεγχο, ο οποίος μας δείχνει εάν το εικονοστοιχείο που μάς υποδηλώνει η νέα κατεύθυνση είναι εντός ορίων της εικόνας (πρέπει να είναι μέσα στον χώρο $(0,M)$ και $(0,N)$). Έπειτα με την γραμμή

```
if ( x(di(jj+1)+q,dj(jj+1)+r) == val)
```

```
    dii = jj; m = 1;
```

ελέγχουμε εάν το εικονοστοιχείο στο οποίο ψάχνουμε έχει τιμή ίση με το val . Εάν αυτό ισχύει, τότε ο μετρητής dii (που θα καθορίσει την κατεύθυνση) παίρνει την τιμή του jj και η m γίνεται 1, που σημαίνει ότι βρέθηκε γειτονικό εικονοστοιχείο.

```
end;
```

```
end;
```

```
end;
```

```
if m==1
```

```
    if n<200
```

```
        n=n+1;
```

```
        c(n) = dii;
```

```
        q =q+ di(dii+1);
```

```
        r =r+ dj(dii+1);
```

```
        lastdir = mod(dii+4,8);
```

```
end;
```

```
    else break;
```

```
end;
```

```
    if n>=200 break; end;
```

```
end;
```

Αν υπήρχε επιτυχία στην εύρεση εικονοστοιχείου και η m πήρε την τιμή 1, τότε οι μεταβλητές q και r παίρνουν τις τιμές του νέου εικονοστοιχείου και η $lastdir$ της προηγούμενης ακριβώς κατεύθυνσης.

```
if(n>0)
```

```
code=c;
```

```
else error(sprintf('Μήκος Αλυσίδας: 0'));
```

```
end;
```

Οι τελευταίες γραμμές κώδικα επιστρέφουν την αλυσίδα ενώ καλύπτουν και την περίπτωση όπου το μήκος αυτής είναι 0.

Αρχείο dif_chain_code.m

```
function dif=dif_chain_code(cc)
```

```
[M,N]=size(cc);
```

```
for i=1:N-1
```

```
    dif(i)=mod((cc(i+1)-cc(i)),8);
```

```
end
```

```
dif(N)=mod((cc(1)-cc(end)),8);
```

Στο αρχείο αυτό υλοποιείται ο differential chain code. Η μεταβλητή dif επιστρέφεται και περιέχει τον κώδικα τροποποιημένο. Πιο συγκεκριμένα, δέχεται ως είσοδο την ακολουθία των αριθμών που προέκυψαν με τη χρήση του αλγορίθμου chain code. Στη συνέχεια κάθε αριθμός αφαιρείται από τον επόμενο του και υπολογίζεται το υπόλοιπο του αποτελέσματος με την τιμή 8. Με αυτόν τον τρόπο, προκύπτει η νέα και τροποποιημένη ακολουθία αριθμών.

Αρχείο apply_chain_code.m

Το επόμενο αρχείο ονομάζεται apply_chain_code και είναι μια συνάρτηση που προεπεξεργάζεται μια εικόνα και έπειτα χρησιμοποιεί τους αλγορίθμους chain code και differential chain code για να κάνει ανάλυση σχήματος

```
function apply_chain_code
```

```
I=imread('michelle.jpg');
```

```
I=rgb2gray(I);
```

Πρώτα γίνεται η μετατροπή του σχήματος σε ασπρόμαυρο (κλίμακα του γκριζου)

```
BW=edge(I,'canny');
```

Έπειτα χρησιμοποιείται η συνάρτηση edge για ανίχνευση ακμών, με τελεστή τον canny.

```
BW1=bwareaopen(BW,35);
```

Η συνάρτηση του Matlab bwareaopen, η οποία εφαρμόζεται στη δυαδική πλέον εικόνα BW1, αφαιρεί όλα τα συνεκτικά στοιχεία (αντικείμενα) από μια δυαδική εικόνα τα οποία έχουν λιγότερα από 35 pixels.

```
L=bwlabel(BW1);
```

Έπειτα η bwlabel επιστρέφει τον πίνακα L ο οποίος έχει το ίδιο μέγεθος με τον πίνακα BW, ο οποίος περιέχει ετικέτες για τα συνεκτικά αντικείμενα (8 στον αριθμό στο συγκεκριμένο παράδειγμα) στον BW.

```
figure;
```

```
subplot(221);
```

```
imshow(I)
```

```
xlabel('input image(gray)');
```

```
axis equal;
```

```
subplot(222);
```

```
imshow(BW)
```

```
xlabel('Canny Filter');
```

```
axis equal;
```

```

subplot(223);
imagesc(L)
xlabel('Connected components larger than 35 Pixel');
axis equal;
%-----

```

```
figure;imshow(BW);
```

```
hold on;
```

το παραπάνω κομμάτι κώδικα δείχνει την εικόνα και πώς αυτή χωρίζεται σε περιοχές.

```
%-----Do chain code-----
```

```
k=max(max(L));
```

```
for i=1:k%για καθε περιοχι i
```

```
    [xc,yc]=find(L==i);
```

```
    cc=chain_code(L,xc(1),yc(1));%chain code
```

```
    dcc=dif_chain_code(cc);%differential chain code
```

```
    plot(xc(1),yc(1),'-r*');xlabel('Start point');
```

```
    disp(sprintf('\nΚωδικας αλυσιδας:')); cc
```

```
    disp(sprintf('\nΔιαφορικος κωδικας αλυσιδας:')); dcc
```

```
    disp(sprintf('\n-----'));

```

```
    clear cc dcc;
```

```
end;
```

Το παραπάνω κομμάτι κώδικα έχει ως αποτέλεσμα, για καθε περιοχή της εικόνας να εφαρμόζονται οι αλγόριθμοι chain code και differential chain code.

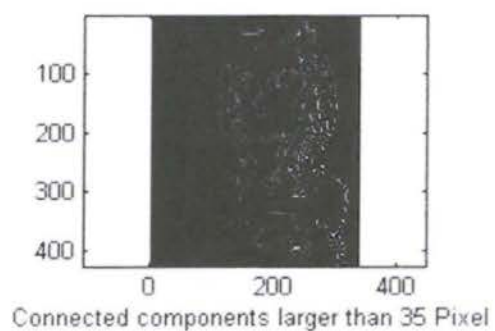
Παρακάτω φαίνονται οι εικόνες που παράγει το αρχείο apply_chain_code.



input image(gray)



Canny Filter



Connected components larger than 35 Pixel

Η εικόνα εισόδου, η εικόνα μετά την ανίχνευση ακμών και οι περιοχές της εικόνας.



Η εικόνα μετά την εφαρμογή του chain code.

Chain Code:

cc =

Columns 1 through 21

6 6 6 6 7 6 7 0 7 0 0 0 1 1 0 7 6 6 6 6 6

Columns 22 through 42

6 5 5 5 5 5 6 6 6 7 6 6 7 6 6 6 6 6 7 6

Columns 43 through 51

6 7 6 6 6 6 7 6 7

Differential Chain Code:

dcc =

Columns 1 through 21

0 0 0 1 7 1 1 7 1 0 0 1 0 7 7 7 0 0 0 0 0

Columns 22 through 42

7 0 0 0 0 1 0 0 1 7 0 1 7 0 0 0 0 0 1 7 0

Columns 43 through 51

1 7 0 0 0 1 7 1 7

Ενδεικτικές τιμές (κατευθύνσεις) που προκύπτουν με την εφαρμογή του chain code και differential chain code αντίστοιχα σε μία περιοχή της εικόνας.

2.5 ΑΝΑΛΥΣΗ ΣΧΗΜΑΤΟΣ-ΚΕΝΤΡΙΚΕΣ ΚΑΙ ΑΝΑΛΟΙΩΤΕΣ ΡΟΠΕΣ

Το σύνολο των ροπών moments μιας φραγμένης συνάρτησης $f(x,y)$ ορίζεται από την σχέση

$$M_{jk} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^j y^k f(x,y) dx dy$$

Όπου οι δείκτες j και k λαμβάνουν όλες τις μη-αρνητικές τιμές. Το σύνολο όλων των ροπών $\{M_{jk}, j,k = 0,1,2,\dots\}$ προσδιορίζει πλήρως και μονοσήμαντα την συνάρτηση $f(x,y)$. Υποθέτουμε ότι η συνάρτηση $f(x,y)$ λαμβάνει την τιμή 1 στο εσωτερικό ενός αντικειμένου, και 0 παντού αλλού. Η συνάρτηση αυτή δηλαδή αντανακλά μόνο το σχήμα ενός αντικειμένου αγνοώντας χαρακτηριστικά του εσωτερικού του.

Η παράμετρος $j+k$ ονομάζεται τάξη (order) της ροπής M_{jk} . Υπάρχει μόνο μια μηδενική ροπή μηδενικής τάξης

$$M_{00} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy$$

Και αντίστοιχα περισσότερες υψηλότερης τάξης (2 ροπές πρώτης τάξης κλπ).

Με τη βοήθεια των συντεταγμένων **κέντρου βάρους**, υπολογίζονται οι κεντρικές ροπές από την εξής σχέση:

$$\mu_{jk} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^j (y - \bar{y})^k f(x,y) dx dy$$

όπου \bar{x} , \bar{y} οι συντεταγμένες του κέντρου βάρους και άκρα ολοκλήρωσης $+\infty$ και $-\infty$. Στην συνάρτηση αυτή λοιπόν, περνάμε σαν παράμετρο μια εικόνα και τις τιμές των j,k .

Το βασικότερο πλεονέκτημα των κεντρικών ροπών είναι το γεγονός ότι παραμένουν αναλλοίωτες σε μεταβολές της θέσης του αντικειμένου.

Η αναζήτηση ουσιαστικών χαρακτηριστικών αναλλοίωτων και σε μετασχηματισμούς περιστροφής οδήγησε στη δημιουργία των ροπών του Hu.

Πρόκειται για μια σειρά από μεγέθη, τα $\phi_1, \phi_2, \dots, \phi_7$, που υπολογίζονται συναρτήσει των κεντρικών κανονικοποιημένων ροπών της περιοχής. Οι έξι πρώτες από αυτές παραμένουν αναλλοίωτες και σε κατοπτρικούς μετασχηματισμούς. Η ϕ_7 δίνει τιμές περίπου αντίθετες για μια συμμετρική περιοχή και την κατοπτρική της.

$$\phi_1 = I_{20} + I_{02}$$

$$\phi_2 = (I_{20} - I_{02})^2 + 4I_{11}^2$$

$$\phi_3 = (I_{30} - 3I_{12})^2 + (3I_{21} - I_{03})^2$$

$$\phi_4 = (I_{30} + I_{12})^2 + (I_{21} + I_{03})^2$$

$$\phi_5 = (I_{30} - 3I_{12})(I_{30} + I_{12}) \left[(I_{30} + I_{12})^2 - 3(I_{21} + I_{03})^2 \right] + \\ + (3I_{21} - I_{03})(I_{21} + I_{03}) \left[3(I_{30} + I_{12})^2 - (I_{21} + I_{03})^2 \right]$$

$$\phi_6 = (I_{20} - I_{02}) \left[(I_{30} + I_{12})^2 - (I_{21} + I_{03})^2 \right] + 4I_{11}(I_{30} + I_{12})(I_{21} + I_{03})$$

$$\phi_7 = (3I_{21} - I_{03})(I_{30} + I_{12}) \left[(I_{30} + I_{12})^2 - 3(I_{21} + I_{03})^2 \right] + \\ + (3I_{12} - I_{30})(I_{21} + I_{03}) \left[3(I_{30} + I_{12})^2 - (I_{21} + I_{03})^2 \right]$$

Οι ροπές του Hu

Για τον υπολογισμό των αναλλοίωτων ροπών, θα χρησιμοποιηθεί πρώτα το αρχείο `centralmoments.m` για τον υπολογισμό των κεντρικών ροπών και μετά το αρχείο `apply.m` για τον υπολογισμό των κανονικοποιημένων κεντρικών ροπών καθώς και των κανονικοποιημένων ροπών 2ης και 3ης τάξης. Από αυτές υπολογίζονται οι 7 ροπές του Hu.

Αρχείο `centralmoments.m`

```
function muv = centralmoments (Image,u,v)
```

πρόκειται για την συνάρτηση `centralmoments` η οποία επιστρέφει την μεταλητή `mun`, παίρνοντας ως ορίσματα την εικόνα `Image` και τις παραμέτρους `u` και `v`.

```
Image = double(Image);
```

```
Image = im2bw(Image);
```

Η εικόνα γίνεται διπλής ακρίβειας και ασπρόμαυρη με τις δύο παραπάνω εντολές.

```
[r,c] = find(Image==1);
```

Στις μεταβλητές `r` και `c` αποθηκεύονται τα σημεία στα οποία η εικόνα περιέχει άσσους. Στην `r` αποθηκεύονται τα `x` αυτών των σημείων και στην `c` τα `y` αυτών των σημείων.

```
rbar = mean(r);
```

```
cbar = mean(c);
```

Με τις δύο αυτές εντολές υπολογίζεται ο μέσος όρος των `r` και `c` και αποθηκεύεται στις μεταβλητές `rbar` και `cbar` αντίστοιχα.

```
n = length(r);
```

Στην μεταβλητή `n` αποθηκεύεται το μέγεθος των σειρών (της μεταβλητής `r`).

```
momlist = zeros(n,1);
```

Στην μεταβλητή `momlist` δημιουργείται ο πίνακας διαστάσεων `n x 1` και αρχικοποιείται με μηδενικά.

```
for i = 1 : n
```

```
    momlist(i) = (r(i) - rbar)^u * (c(i) - cbar)^v;
```

```
end
```

Σε αυτό το βρόγχο πραγματοποιείται ο μαθηματικός τύπος υπολογισμού κεντρικών ροπών, όπως αυτός φαίνεται στο βιβλίο «Ανάλυση Εικόνας», σελίδα 112.

```
mun = sum(momlist);
```

```
end
```

Το άθροισμα των στοιχείων της μεταβλητής `momlist` επιστρέφεται από την συνάρτηση.

Αρχείο `apply.m`

Με το αρχείο apply.m υπολογίζονται οι αναλλοίωτες ροπές.

```
function [vec]=apply(Image)
```

η συνάρτηση apply δέχεται ως όρισμα την εικόνα Image

```
area = bwarea(Image);
```

με την εντολή αυτή υπολογίζεται ο χώρος που καλύπτεται στην εικόνα και αποθηκεύεται στην μεταβλητή area.

```
u11 = centralmoments(Image,1,1) / (area^2);
```

Χρησιμοποιείται η συνάρτηση centralmoments για να υπολογιστούν οι κεντρικές ροπές μηδενικού, 1^{ου} και 2^{ου} βαθμού αντίστοιχα. Έπειτα οι ροπές κανονικοποιούνται αφού διαιρούνται με την μεταβλητή area στο τετράγωνο.

```
u20 = centralmoments(Image,2,0) / (area^2);
```

```
u02 = centralmoments(Image,0,2) / (area^2);
```

... το ίδιο για τις μεταβλητές 1^{ου} βαθμού

```
u30 = centralmoments(Image,3,0) / (area^2.5);
```

```
u03 = centralmoments(Image,0,3) / (area^2.5);
```

```
u21 = centralmoments(Image,2,1) / (area^2.5);
```

```
u12 = centralmoments(Image,1,2) / (area^2.5);
```

...και τις μεταβλητές 3^{ου} βαθμού

Στην συνέχεια υπολογίζονται οι συναρτήσεις του hu, οι οποίες δίνουν τις αναλλοίωτες ροπές και είναι 7 στον αριθμό.

```
h1 = u20 + u02;
```

```
h2 = (u20-u02)^2 + 4*u11^2;
```

```
h3 = (u30-3*u12)^2 + (u03-3*u21)^2;
```

```
h4 = (u30+u12)^2 + (u03+u21)^2;
```

```
h5 = (u30-3*u12)*(u30+u12)*((u30+u12)^2-3*(u03+u21)^2) ...  
+ (3*u21-u03)*(u03+u21)*(3*(u30+u12)^2 - (u03+u21)^2);
```

$$h6 = (u20-u02)*((u30+u12)^2-(u03+u21)^2) + 4*u11*(u30+u12)*(u03+u21);$$

$$h7 = (3*u21-u03)*(u30+u12)*((u30+u12)^2-3*(u03+u21)^2) ...$$

$$- (u30-3*u12)*(u03+u21)*(3*(u30+u12)^2 - (u03+u21)^2);$$

Η συνάρτηση επιστρέφει το διάνυσμα vec , όπου βρίσκονται οι συναρτήσεις του h

$$vec = [h1, h2*100, h3*100, h4*10^3, h5*10^5, h6*10^4, h7*10^5];$$

`disp(vec)`

αρχείο test.m

Τέλος, στο αρχείο test.m φαίνεται η εφαρμογή των κεντρικών και αναλλοίωτων ροπών.

```
function [centroids]=test()
Image = imread('michelle.png');
figure, title('Original Image'),imshow(Image);
Image=~im2bw(Image);
centroids=apply(Image);
centroids
```

```
J = imresize(Image, 0.5);
figure,title('michelle'),imshow(Image);
Image=~im2bw(J);
centroids=apply(Image);
centroids
```

```
Image = imread('michelle1.png');
figure, imshow(Image);
Image=~im2bw(Image);
centroids=apply(Image);
centroids
```

```
mage = imread('michelle2.png');
figure, imshow(Image);
Image=~im2bw(Image);
centroids=apply(Image);
```

```
centroids
```

```
Image = imread('michelle3.png');
```

```
figure, imshow(Image);
```

```
Image=~im2bw(Image);
```

```
centroids=apply(Image);
```

```
centroids
```

Σε κάθε block κώδικα, διαβάζεται και μια διαφορετική εικόνα. Η αυθεντική, μειωμένη στις διαστάσεις κατά 50%, περιστραμμένη κατά 45 μοίρες δεξιά, περιστραμμένη κατά 45 μοίρες αριστερά, και περιστραμμένη κατά 180 μοίρες. Κάθε φορά παρουσιάζονται τα αποτελέσματα των αναλλοίωτων ροπών με την χρήση της συνάρτησης `apply`. Είναι σημαντικό ότι οι τιμές τους δεν αλλάζουν σημαντικά στην περίπτωση της περιστροφής. Ωστόσο, οι μεταβολές αυτές οφείλονται στο θόρυβο που εισέρχεται στις εικόνες κατά τη δημιουργία των περιστραμμένων αντιγράφων καθώς και στα σφάλματα των πράξεων κατά τον υπολογισμό των ροπών.

Μεταβολή	Φ1	Φ2	Φ3	Φ4	Φ5	Φ6	Φ7
Αρχική εικόνα	0.3626	0.2998	0.2364	0.2072	0.0126	0.1082	0.0072
Σμίκρυνση	0.2813	0.9214	0.6172	0.0641	0.0029	0.0604	0.0028
Περιστροφή 45° δεξιά	0.3633	0.1876	0.1379	0.1797	0.0059	0.0618	0.0067
Περιστροφή 45° αριστερά	0.3633	0.1876	0.1379	0.1797	0.0059	0.0618	0.0067
Περιστροφή 180°	0.3626	0.2998	0.2364	0.2072	0.0126	0.1082	0.0072



Η αρχική εικόνα.



Περιστροφή κατά 45 μοίρες προς τα αριστερά.



Περιστροφή κατά 45 μοίρες προς τα δεξιά.



Περιστροφή κατά 180 μοίρες.

2.6 ΧΑΡΤΗΣ ΒΑΘΟΥΣ-ΧΑΡΤΗΣ ΒΕΛΟΝΩΝ

Ο χάρτης βελονών υλοποιείται σε δύο ξεχωριστά αρχεία, το `needleplotst.m` και το `needle.m` που κάνει χρήση του προηγούμενου. Θα γίνει επίδειξη του χάρτη βελονών επάνω σε μία σφαίρα.

Αρχείο `needle.m`

```
function needle
```

η βασική μας συνάρτηση λέγεται `needle` και χρησιμοποιεί επίσης την συνάρτηση `needleplotst`.

```
[a,b,c] = sphere(25);
```

Η εντολή αυτή δημιουργεί μια σφαίρα με την χρήση 25 grids για οδηγούς και αποθηκεύει τις τιμές των σημείων των πλεγμάτων σε τρεις πίνακες $n+1 \times n+1$ όπου n ο αριθμός των πλεγμάτων (εδώ 25). Είναι οι τρεις πίνακες $[a,b,c]$.

```
figure,surf(a,b,c);
```

με την εντολή `surf` σχεδιάζουμε την σφαίρα

```
z=[a b c];
```

```
[dx, dy] = gradient(z);
```

Έπειτα αποθηκεύονται οι τιμές των πινάκων a , b και c στην μεταβλητή z και ύστερα με την εντολή `Gradient` παίρνουμε το `gradient` της z και το αποθηκεύουμε στις μεταβλητές dx και dy αντίστοιχα.

```
tilt = atan2(-dy, -dx);
```

η εντολή `atan2(-dx, -dy)` επιστρέφει ένα array (το `Tilt` εδώ) με το ίδιο μέγεθος με το dx και το dy το οποίο περιέχει τις εφαπτομένες όλων των μερών (των πραγματικών μερών) των dx , dy

```
grad = sqrt(dx.^2 + dy.^2)+eps;
```

η μεταβλητή `grad` ισούται με την τετραγωνική ρίζα του αθροίσματος των dx^2 και dy^2 (δηλαδή την απόσταση μεταξύ των dx και dy) συν την απόσταση `eps`, που είναι μια συνάρτηση του `Matlab` για διόρθωση σφαλμάτων.

```
slant = atan(grad);
```

Τέλος, η μεταβλητή `slant` είναι ίση την τιμή της εφαπτομένης του `grad`.

```
figure,needleplotst(slant,tilt,5,2), axis('off')
```

Αρχείο needleplotst.m

Εδώ παρουσιάζεται η συνάρτηση needleplotst, η οποία δημιουργεί το χάρτη βελονών από τα δεδομένα των μεταβλητών slant, tilt. Η μεταβλητή len που εισάγεται ως παράμετρος είναι το μέγεθος των βελονών. Η μεταβλητή spacing καθορίζει το διάστημα με βάση το οποίο επιλέγουμε βελόνες για να εμφανιστούν. Οι τιμές αυτές στο παράδειγμά μας έχουν πάρει τιμή ήδη από την συνάρτηση needle και είναι len = 5 και spacing = 2.

```
function needleplotst(slant, tilt, len, spacing)
```

```
lw = 5;
```

η μεταβλητή lw είναι ίση με το μήκος της βελόνας

```
s_slant = slant(1:spacing:end, 1:spacing:end);
```

```
s_tilt = tilt(1:spacing:end, 1:spacing:end);
```

Οι μεταβλητές s_slant και s_tilt παίρνουν τα στοιχεία των πινάκων slant και tilt αντίστοιχα, ξεκινώντας από το πρώτο στοιχείο τους κάθε φορά και προχωρώντας με βήμα spacing. Αυτό γίνεται γιατί δεν θέλουμε να εμφανίσουμε όλες τις βελόνες (η εικόνα θα φορτωνόταν υπερβολικά) αλλά κάποιες από αυτές που χωρίζονται μεταξύ τους με βήμα spacing.

```
[s_rows, s_cols] = size(s_slant);
```

Έπειτα οι μεταβλητές s_rows και s_cols παίρνουν τις τιμές των μεγεθών των διαστάσεων (γραμμές και στήλες αντίστοιχα) της slant (εντολή size).

```
projlen = len*sin(s_slant);
```

```
dx = projlen.*cos(s_tilt);
```

```
dy = projlen.*sin(s_tilt);
```

Με τις τρεις αυτές εντολές προβάλλεται το μήκος της κάθε βελόνας στο επίπεδο xy.

```
clf
```

```
for r = 1:s_rows
```

```
for c = 1:s_cols
```

```
    x = (c-1)*spacing+1;
```

```

y = (r-1)*spacing+1;

plot(x,y,'r. '); hold on

line([x x+dx(r,c)],[y y+dy(r,c)],'color',[0 1 1],'linewidth',lw);

end

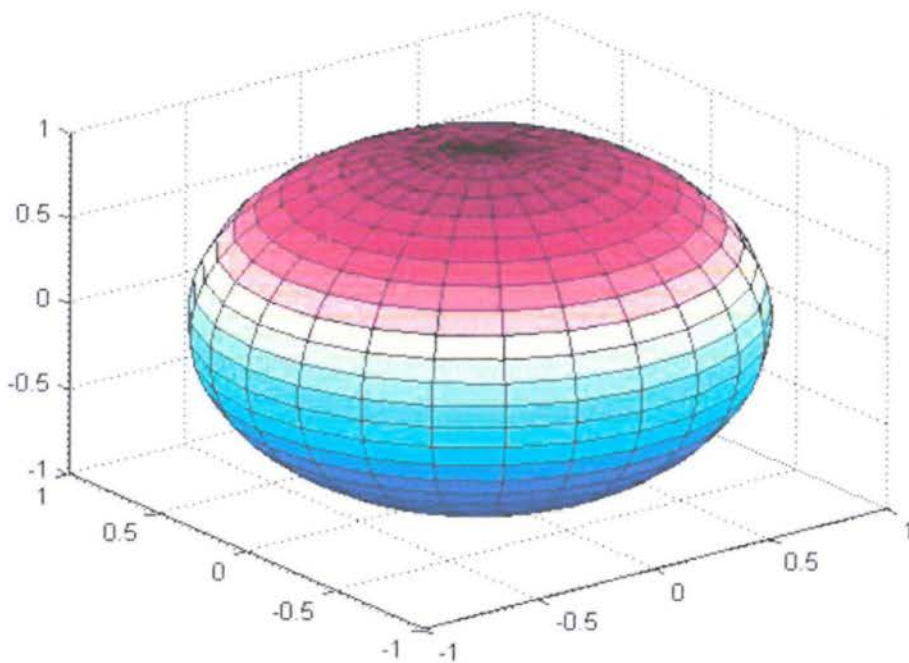
end

axis('equal')

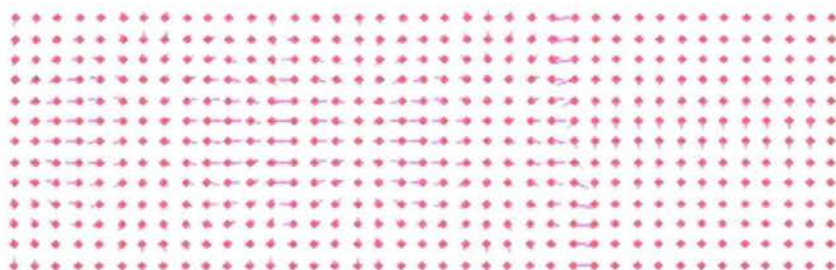
hold('off')

```

Με τον παραπάνω επαναληπτικό βρόγχο σχεδιάζονται όλες οι βελόνες. Εμφανίζονται πρώτα οι κουκίδες με την εντολή plot και έπειτα οι βελόνες με την εντολή line.



Το αρχικό μας σχήμα



Ο χάρτης βελονών

2.7 ΕΥΡΕΣΗ ΑΝΤΙΣΤΟΙΧΩΝ ΣΗΜΕΙΩΝ ΣΕ ΖΕΥΓΗ ΣΤΕΡΕΟΣΚΟΠΙΚΩΝ ΕΙΚΟΝΩΝ (ΑΛΓΟΡΙΘΜΟΣ ΤΟΥ SHIRAI)

Ο αλγόριθμος Shirai επιτρέπει την ανακατασκευή ορατών επιφανειακών σημείων αντικειμένων στο τρισδιάστατο χώρο σκηνής. Στη γενική περίπτωση, έχουμε να υπολογίσουμε αντίστοιχα σημεία εικόνας καθώς και το χάρτη ανισοτιμιών για ζεύγη στερεοσκοπικών εικόνων επιπέδου γκρίζου.

Αρχικά, υπολογίζουμε την εικόνα των ακμών της αριστερής εικόνας, δηλαδή η αναζήτηση αρχίζει πάντα με αρχικό σημείο στην αριστερή εικόνα. Για κάθε εικονοστοιχείο της αριστερής εικόνας ξεκινάει μια διαδικασία αναζήτησης για αντίστοιχο εικονοστοιχείο στη δεξιά εικόνα.

Κατά τη διάρκεια της αναζήτησης, υπολογίζεται ένα μέτρο ομοιότητας (similarity measure) για κάθε εικονοστοιχείο στη δεξιά εικόνα που αποτελεί υποψήφιο με βάση το διάστημα αναζήτησης. Οι τιμές επιπέδου γκρίζου σε παράθυρα της αριστερής και της δεξιάς εικόνας συγκρίνονται μεταξύ τους για να προσδιοριστεί το μέτρο ομοιότητας SIMILARITY (p, q). Στην αριστερή εικόνα ένα σταθερό παράθυρο F (E_{left}, p) συμπεριφέρεται ως πίνακας σύγκρισης. Κατά την αναζήτηση, το παράθυρο εικόνας F (E_{right}, q) συγκρίνεται σε όλες τις δυνατές θέσεις στη γραμμή αναζήτησης y με το σταθερό παράθυρο εικόνας χρησιμοποιώντας το τετραγωνικό σφάλμα (squared error).

$$SE(p, q) = \sum_{i=-k}^k \sum_{j=-k}^k (E_{left}(x_{left} + i, y + j) - E_{right}(x_{right} + i, y + j))^2$$

Αυτή η σύγκριση κανονικοποιείται σε σχέση με την ποσότητα

$$\begin{aligned} VARIANCE(p) &= \frac{1}{(2k+1)^2} \sum_{i=-k}^k \sum_{j=-k}^k [E_{left}(x_{left} + i, y + j) - AVERAGE(p)]^2 \\ &= \frac{1}{(2k+1)^2} \sum_{i=-k}^k \sum_{j=-k}^k E_{left}^2(x_{left} + i, y + j) - AVERAGE^2(p) \end{aligned}$$

που αποτελεί την εκτίμηση της διακύμανσης του σταθερού παραθύρου εικόνας στην αριστερή εικόνα. Η τιμή AVERAGE (p) υποδηλώνει τον αριθμητικό μέσο του παραθύρου εικόνας F (E_{left}, p) και, κατά συνέπεια, μια εκτίμηση της αναμενόμενης τιμής του επιπέδου γκρίζου σε αυτό το παράθυρο εικόνας.

Το μέτρο ομοιότητας του Shirai ορίζεται ως

$$SIMILARITY(p, q) = \frac{SE(p, q)}{VARIANCE(p) + 1}$$

Ο αλγόριθμος είναι υλοποιημένος σε 5 αρχεία.

Apply.m: Είναι η κύρια συνάρτηση, η οποία χρησιμοποιείται για την ανάγνωση των εικόνων και την κλήση των επιμέρους συναρτήσεων για την εφαρμογή του αλγόριθμου Shirai.

Shirai.m: Η συνάρτηση που περιέχει το βασικό αλγόριθμο Shirai.

Se.m: Η συνάρτηση η οποία επιστρέφει την τιμή του τετραγωνικού σφάλματος (squared error).

Variance.m: Η συνάρτηση που επιστρέφει την τιμή την εκτίμηση της διακύμανσης του σταθερού παραθύρου εικόνας στην αριστερή εικόνα.

Mean.m: Η συνάρτηση που επιστρέφει τον τοπικό μέσο του παραθύρου εικόνας.

Αρχείο apply.m

```
leftim= imread('left.jpg');  
rightim= imread('right.jpg');  
imshow(leftim);  
title('Left Picture')  
figure,imshow(rightim);  
title('Right Picture')  
  
left=im2bw(leftim);  
right=im2bw(rightim);
```

```
left = double ( left );  
edgeimage = edge(left,'sobel');
```

Αρχικά, διαβάζουμε τις εικόνες left.jpg και right.jpg και αποθηκεύουμε τις τιμές των pixel τους στις μεταβλητές leftid και rightim. Εμφανίζουμε τις αρχικές εικόνες στην οθόνη. Μετατρέπουμε τις εικόνες σε ασπρόμαυρες για να μειώσουμε την πολυπλοκότητα του αλγορίθμου. Μετατρέπουμε τις τιμές του πίνακα left σε μορφή διπλής ακρίβειας για να είναι δεκτός ο πίνακας ως όρισμα στη συνάρτηση edge. Πραγματοποιούμε ανίχνευση ακμών με τη μέθοδο Sobel.

```
maxinterval = 3;  
d1 = 1;  
d2 = 5;  
d3 = 5;  
out_image=(1);  
out_image=SHIRAI(left,right,edgeimage,d1,d2,d3,maxinterval);  
figure,imshow(out_image);  
title('Output Picture')
```

Καλούμε τη συνάρτηση SHIRAI με τις προκαθορισμένες τιμές των μεταβλητών d1, d2, d3 και maxinterval. Η αλλαγή των τιμών αυτών προσδιορίζει τα κατώφλια που θέτουμε και επηρεάζει το συνολικό αριθμό αντιστοίχισης σημείων της αριστερής και της δεξιάς εικόνας. Τα σημεία αυτά είναι εμφανή κατά την εμφάνιση της εικόνας out_image που επιστρέφει η συνάρτηση SHIRAI.

Αρχείο SHIRAI.m

```
function out_image = SHIRAI(left,right,edgeimage,d1,d2,d3,maxinterval)  
[imageWidth imageHeight] = size(left);
```

Υπολογίζουμε το μέγεθος της εικόνας.

```
maximumDisparity = 0;
```

Στη συνέχεια διατρέχουμε την αριστερή εικόνα.

```
for yPos = 1:imageHeight-1
    for xPos = 1:imageWidth-1
        out(xPos,yPos) = 0;
        if (edgeimage(xPos,yPos)== 1)
```

Ελέγχουμε αν το τρέχον pixel είναι σημείο ακμής.

```
        k = 1;
        rightBorder = xPos;
        leftBorder = rightBorder -(imageWidth*maxinterval);
        if leftBorder < 1      % έλεγχος αν το αριστερό σύνορο είναι έγκυρο
                                leftBorder = 1;      % διόρθωση αριστερού συνόρου
        end
        continueLoop=true;
```

Ορίζουμε το μέγεθος του παραθύρου ίσο με την τιμή $n = 2 * k + 1$

```
        while continueLoop
            n = 2*k+1;
            belowD1Count = 0;
            aboveD2Count = 0;
            minPos = -1;
```

Υπολογίζουμε την τιμή της διακύμανσης για το παράθυρο.

```
            windowVariance=VARIANCE(left, k, n, xPos, yPos);
            if (not(windowVariance == 0))
```

```
for i = rightBorder :-1: leftBorder
```

Υπολογίζουμε το μέτρο ομοιότητας (ή πιο σωστά ανομοιότητας) διαιρώντας το τετραγωνικό σφάλμα με την τιμή της διακύμανσης και ελέγχουμε αν οι τελικές τιμές είναι στα επιτρεπτά όρια που έχουμε θέσει με τις μεταβλητές d1, d2 και d3.

```
simvalues(i) = SE(left, right, k, xPos, yPos, i) / windowVariance;
```

```
if ((simvalues(i) < d1) && (simvalues(i) >= 0)) %
```

έλεγχος αν το μέτρο ομοιότητας είναι κάτω από το d1

```
belowD1Count = belowD1Count +1;
```

```
minPos = i;
```

```
end
```

```
if ((simvalues(i) > d2) || (simvalues(i) < 0))
```

```
aboveD2Count = aboveD2Count +1;
```

```
end
```

```
end
```

```
end
```

```
if (belowD1Count == 1)
```

```
out(xPos,yPos) = xPos - minPos;
```

```
if (out(xPos,yPos) > maximumDisparity )
```

```
maximumDisparity = out(xPos,yPos);
```

```
end
```

```
continueLoop = false;
```

```
elseif (aboveD2Count == rightBorder - leftBorder + 1)
```

```
out(xPos,yPos) = 0;
```

```

continueLoop = false;

elseif (((floor(n/2)+2) >= xPos) || ((xPos+ceil(n/2)+2) >= imageWidth) || ((floor(n/2)+2) >=
yPos) || ((yPos+ ceil(n/2)+2) >= imageHeight) || (k==5) )

    out(xPos,yPos) = 0;

    continueLoop = false;

else

    k = k + 1;

    while (((simvalues(leftBorder) > d3) || ( floor(n/2) >= leftBorder)) &&(leftBorder <
rightBorder))

        leftBorder = leftBorder + 1;

    end

    while (simvalues(rightBorder) > d3) && (leftBorder < rightBorder)

        rightBorder = rightBorder - 1;           % μετακίνηση του δεξιού ορίου end

    if (floor(n/2) >= leftBorder)

        out(xPos,yPos) = 0;

        continueLoop = false;

    end

end

end

end

end

end

```

```

if (not(maximumDisparity == 0))
    factor = 255/maximumDisparity;
    for y = 1:imageHeight-1
        for x = 1:imageWidth-1
            out(x,y)= floor((out(x,y)*factor)+0.5);
        end
    end
end

for y= 1:(imageHeight-1)
    for x= 1:(imageWidth-1)
        out_image(x, y) = out(x,y);
    end
end
end
end

```

Αρχείο MEAN.m

Αρχικά, αποθηκεύουμε τις διαστάσεις (μήκος και ύψος αντίστοιχα) της αριστερής εικόνας στις μεταβλητές width και height. Διατρέχουμε την εικόνα με τη χρήση παραθύρου με τιμές από $-k$ μέχρι $+k$, αφού το μέγεθος του παραθύρου είναι $2 * k + 1$. Προσθέτουμε τις τιμές των pixel του παραθύρου και διαιρούμε το αποτέλεσμα με τη διάσταση n^2 του παραθύρου. Η τελική τιμή, που επιστρέφει η συνάρτηση, αποτελεί και τον αριθμητικό μέσο του παραθύρου.

```

function mean_value=MEAN(leftImage,k,n,xLeft,y)

[width height] = size(leftImage);

mean_value = 0;

```



```

for i = -k:k
    for j = -k:k
        if( xLeft+i < width && xLeft+i >= 1 && y+j >= 1 && height < y+j)
            mean_value = mean_value + leftImage(xLeft+i,y+j);
        end
    end
end

end

end

mean_value = mean_value / (n*n);

```

Αρχείο SE.m

Αρχικά, υπολογίζουμε τις διαστάσεις των δύο εικόνων. Διατρέχουμε τη δεξιά εικόνα στην ίδια γραμμή που διατρέχουμε και την αριστερή με τη χρήση του παραθύρου μεγέθους $n = 2 * k + 1$ και υπολογίζουμε το τετραγωνικό σφάλμα, το οποίο επιστρέφει από τη συνάρτηση.

```

function se_val=SE(leftImage, rightImage, k,xLeft,y,xRight)

se_val=0;

[width height] = size(leftImage);

[width2 height2] = size(rightImage);

for i=-k:k

    for j=-k:k

        if( xLeft+i < width && xLeft+i >= 1 && y+j >= 1 && height > y+j)

            if( xRight+i < width2 && xRight+i >= 1 && y+j >= 1 && height2 > y+j)

```

```

        se_val = se_val + ( (leftImage(xLeft+i,y+j) -
rightImage(xRight+i,y+j))*(leftImage(xLeft+i,y+j) - rightImage(xRight+i,y+j)));
    end
end
end
end
end

```

Αρχείο VARIANCE.m

Η συνάρτηση variance επιστρέφει την εκτίμηση της διακύμανσης του σταθερού παραθύρου εικόνας στην αριστερή εικόνα.

```

function variance_val = VARIANCE(leftImage,k, n, xLeft, y)
variance_val=0;
[width height] = size(leftImage);
mean = MEAN(leftImage, k, n, xLeft, y);
for i = -k:k
    for j = -k:k
        if( xLeft+i < width && xLeft+i >= 1 && y+j >= 1 && height > y+j)
            variance_val = variance_val + ((leftImage(xLeft+i,y+j)*leftImage(xLeft+i,y+j)) -
(mean*mean));
        end
    end
end
variance_val = variance_val /(n*n);

```

Οι δύο εικόνες που χρησιμοποιούνται καθώς και η εικόνα με τα αντίστοιχα σημεία τους φαίνονται παρακάτω.



Αριστερή εικόνα



Δεξιά εικόνα

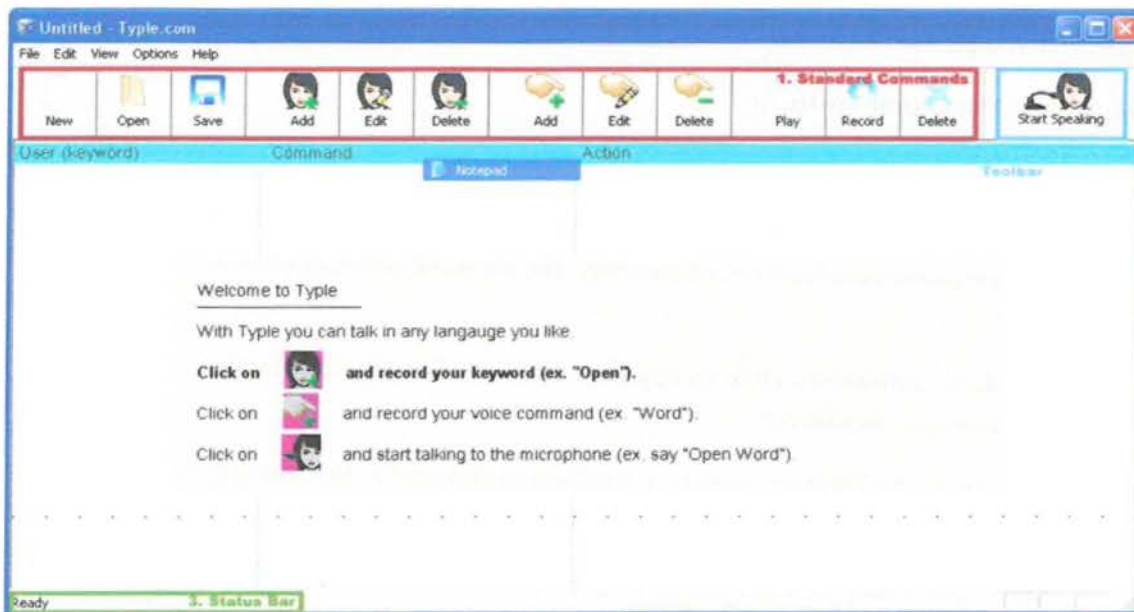


Τελική εικόνα με τα αντίστοιχα σημεία των δύο εικόνων

2.8 ΛΟΓΙΣΜΙΚΟ ΑΝΑΓΝΩΡΙΣΗΣ ΦΩΝΗΣ

Το λογισμικό που χρησιμοποιήθηκε για την εκπόνηση της εργασίας είναι το Typle. Υπάρχει ελεύθερη έκδοση, την οποία μπορεί κανείς να κατεβάσει από το σύνδεσμο <http://www.typle.com/>. Κατεβάσαμε, λοιπόν, το αρχείο Typle-2.0n-MultiLingual-setup.exe, το οποίο είναι υπεύθυνο για την εγκατάσταση του προγράμματος.

Μετά την εγκατάσταση του προγράμματος, εκτελούμε το πρόγραμμα και βλέπουμε την κεντρική οθόνη.



Κεντρική οθόνη του Typle

Η κεντρική οθόνη της εφαρμογής περιέχει το βασικό μενού με τις επιλογές File, Edit, View, Options και Help.

Η επιλογή File περιέχει τις εξής υποκατηγορίες:

New: Δημιουργία νέου αρχείου

Open: Άνοιγμα υπάρχοντος αρχείου

Save: Αποθήκευση αρχείου

Save as: Αποθήκευση αρχείου με διαφορετική ονομασία

Start Speaking: Ενεργοποίηση της εφαρμογής

Η επιλογή Edit περιέχει τις συνήθεις υποκατηγορίες:

Undo: Αναίρεση της τελευταίας ενέργειας

Cut: Αποκοπή μιας επιλογής

Copy: Αντιγραφή μιας επιλογής

Paste: Επικόλληση μιας επιλογής

Η επιλογή View σχετίζεται με τα αντικείμενα τα οποία επιθυμεί ο χρήστης να είναι ορατά στο κεντρικό παράθυρο της εφαρμογής. Πιο συγκεκριμένα, οι υποκατηγορίες είναι οι εξής:

1. Standard Commands: είναι τα εργαλεία που περιέχονται στο κόκκινο περίγραμμα της οθόνης.
2. Start Speaking Toolbar: είναι το εικονίδιο που βρίσκεται στο πάνω δεξιά μέρος της κεντρικής οθόνης, μέσα στο μπλε περίγραμμα.
3. Status bar: είναι η μπάρα στο κάτω μέρος της οθόνης που βρίσκεται μέσα στο πράσινο περίγραμμα.

Η επιλογή Options σχετίζεται με τις ρυθμίσεις της εφαρμογής. Περιέχει τις εξής υποκατηγορίες:

Set Volume: ρύθμιση της έντασης των ηχείων και του μικροφώνου.

Set Default Font: χρήσης της προκαθορισμένης γραμματοσειράς της εφαρμογής.

Align Right To Left: διάταξη του κειμένου από δεξιά προς τα αριστερά.

Languages: Επιλογή γλώσσας της εφαρμογής. Να σημειώσουμε ότι η επιλογή αυτή δεν αφορά τη γλώσσα που μπορεί να χρησιμοποιήσει προφορικά ένας χρήστης για τη δημιουργία εντολών αλλά αφορά τη μετάφραση των επιλογών του προγράμματος σε μία από τις δυνατές υποστηριζόμενες γλώσσες.

Τέλος, η επιλογή Help αποτελείται από τις επιλογές:

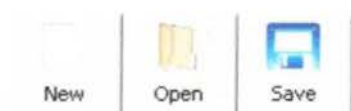
Online Help: προσφέρει βοήθεια μέσω διαδικτύου για τη χρήση του προγράμματος.

FAQ: εμφάνιση συχνά χρησιμοποιούμενων ερωτήσεων

Purchase Online: Αγορά της πλήρους έκδοσης της εφαρμογής μέσω διαδικτύου.

About Typle: Εμφάνιση πληροφορικών σχετικά με την εφαρμογή, όπως αριθμός έκδοσης και πνευματικά δικαιώματα.

Κάτω από το βασικό μενού υπάρχουν τα βασικά εργαλεία του προγράμματος, δηλαδή τα standard commands που είδαμε και προηγουμένως. Τα τρία πρώτα (New, Open, Save) είναι ισοδύναμα με τις αντίστοιχες υποκατηγορίες της επιλογής File του βασικού μενού.



Τα επόμενα τρία εργαλεία (Add, Edit, Delete) με το εικονίδιο της κοπέλας αφορούν την προσθήκη, επεξεργασία ή διαγραφή ενός χρήστη.



Στη συνέχεια, τα εργαλεία Add, Edit, Delete με το εικονίδιο του χεριού αποσκοπούν στη δημιουργία νέας εντολής, επεξεργασία και διαγραφή υπάρχουσας εντολής αντίστοιχα.



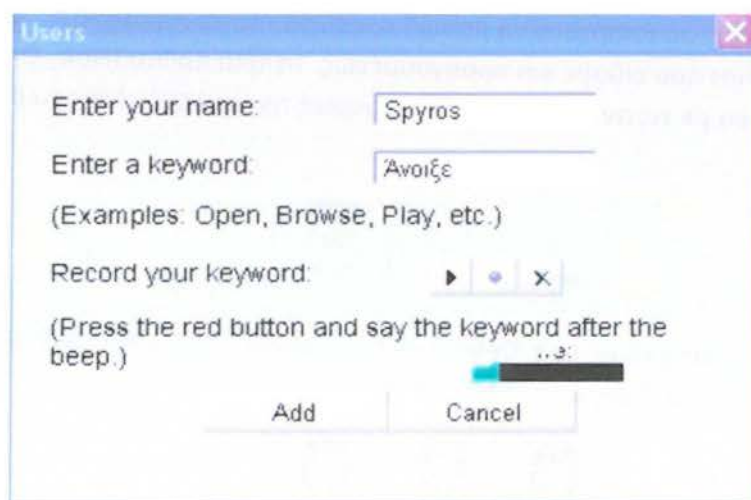
Τέλος, τα εργαλεία Play, Record και Delete χρησιμοποιούνται για την αναπαραγωγή, εγγραφή ή διαγραφή ενός αρχείου ήχου σε μορφή wav.



Εκπαίδευση της εφαρμογής

Για την εκπαίδευση της εφαρμογής ακολουθούντα τα εξής βήματα:

1. Κλικ στο εικονίδιο προσθήκης νέου χρήστη. Βλέπουμε τότε στην οθόνη το παρακάτω παράθυρο.



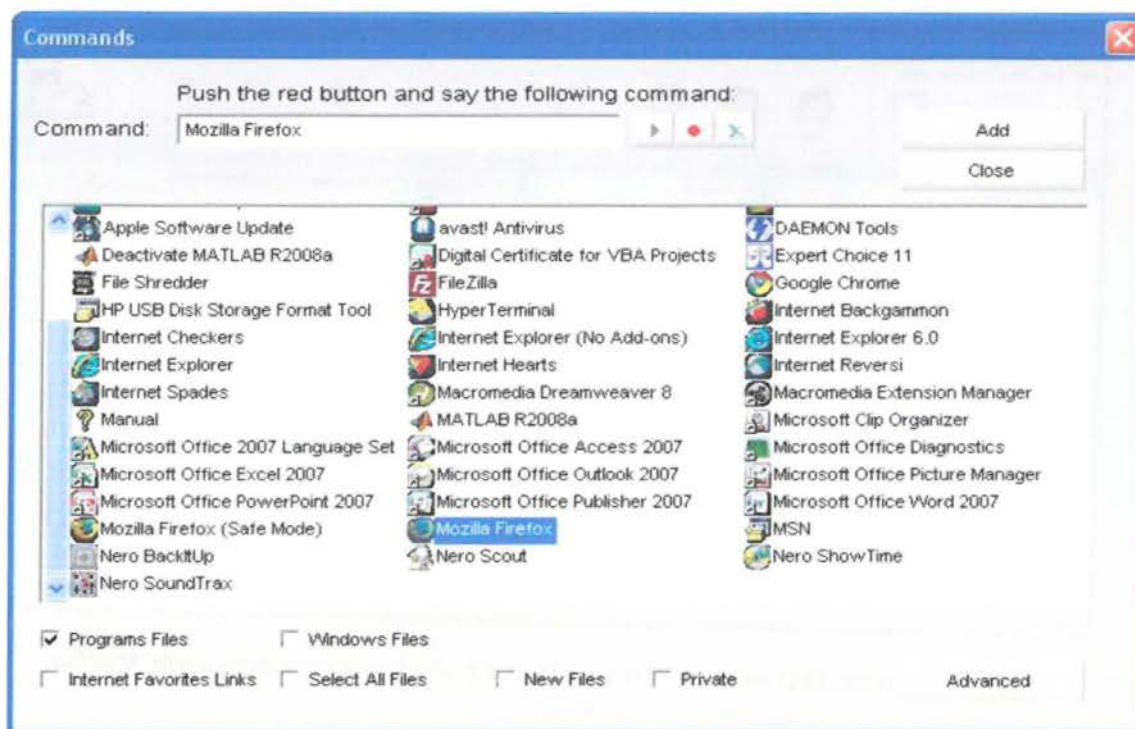
The image shows a dialog box titled "Users" with a close button (X) in the top right corner. It contains the following fields and controls:

- "Enter your name:" with a text input field containing "Spyros".
- "Enter a keyword:" with a text input field containing "Άνοιξε".
- Below the keyword field, the text "(Examples: Open, Browse, Play, etc.)" is displayed.
- "Record your keyword:" with a control panel containing a play button (▶), a red circular button, and a close button (X).
- Below the recording control, the text "(Press the red button and say the keyword after the beep.)" is displayed.
- At the bottom, there is a progress bar and two buttons: "Add" and "Cancel".

Παράθυρο προσθήκης νέου χρήστη

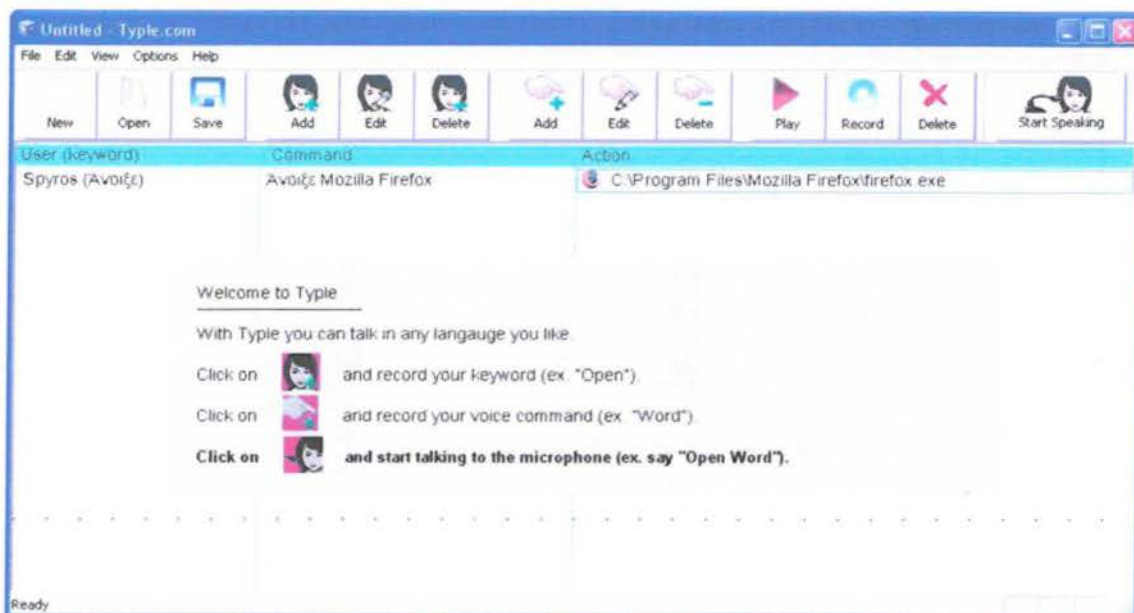
Ο χρήστης εισάγει το όνομά του και μία λέξη κλειδί για την απλή περιγραφή της κατηγορίας εντολών που θέλει να προσθέσει. Κάνει κλικ στο εικονίδιο με τον κύκλο ώστε να είναι σε θέση να καταγράψει φωνητικά την εντολή. Τέλος, πατώντας στο πλήκτρο Add καταχωρείται το όνομα του χρήστη μαζί με την κωδική του λέξη.

2. Το επόμενο βήμα είναι η προσθήκη εντολών μέσω του εργαλείου Add με το εικονίδιο του χεριού. Με αυτόν τον τρόπο εμφανίζεται στην οθόνη το παρακάτω παράθυρο.



Παράθυρο εντολών

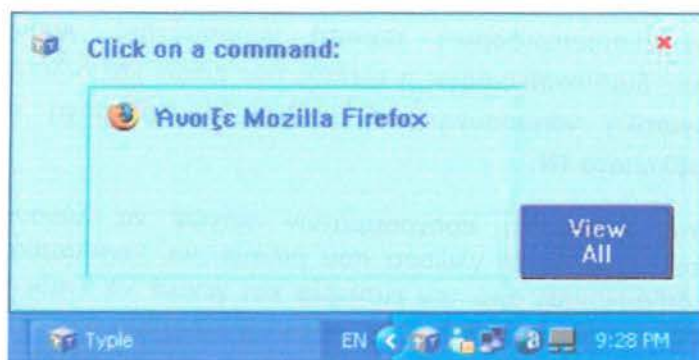
Αυτό περιέχει όλα τα προγράμματα που έχει εγκατεστημένα ο χρήστης στον υπολογιστή του. Επιλέγοντας μία κατηγορία από το κάτω μέρος του παραθύρου, μπορεί να δει τα προγράμματα που υπάγονται στην κατηγορία αυτή ή να επιλέξει την επιλογή 'Select All Files' για να εμφανιστούν όλα τα διαθέσιμα προγράμματα. Κάνοντας κλικ σε ένα από αυτά, καταγράφει μία ονομασία για την εντολή (π.χ. Mozilla Firefox) και στη συνέχεια καταγράφει ηχητικά την εντολή (π.χ. πάλι Mozilla Firefox). Κάνοντας κλικ στην επιλογή 'Add', αποθηκεύεται η εντολή που μόλις δημιούργησε ο χρήστης και είναι ορατή στην κεντρική οθόνη.



Ο χρήστης Spyros έχει καταγράψει εντολή για το άνοιγμα της εφαρμογής Mozilla Firefox.

Κλείνοντας το παράθυρο της εφαρμογής, υπάρχει το εικονίδιο της εφαρμογής στο κάτω δεξιό μέρος της οθόνης. Πρόκειται για έναν κύβο, του οποίου η πάνω πλευρά αναβοσβήνει με κόκκινο χρώμα. Αυτό αποτελεί ένδειξη ότι η εφαρμογή είναι ενεργοποιημένη και ο χρήστης έχει τη δυνατότητα να εκτελέσει τις εφαρμογές που θέλει, απλά εκφωνώντας τις αντίστοιχες εντολές που ο ίδιος έχει φτιάξει.

Με αυτό τον τρόπο, αν ο χρήστης προσφωνήσει τη φράση 'Mozilla Firefox', θα εκτελεστεί η αντίστοιχη εφαρμογή. Αν προσφωνήσει απλά τη λέξη κλειδί (π.χ. 'Ανοιξε'), θα εμφανιστεί πάνω από το εικονίδιο ένα μικρό παραθυράκι με τις διαθέσιμες εντολές της συγκεκριμένης κατηγορίας που περιγράφει η λέξη κλειδί.



Η κατηγορία της λέξης κλειδί 'Άνοιξε' με τις αντίστοιχες εντολές

Είναι πολύ σημαντικό το γεγονός ότι το πρόγραμμα αναγνωρίζει τη φωνή του χρήστη. Οι εντολές θα εκτελεστούν μόνο στην περίπτωση που προσφωνηθούν από το χρήστη που τις δημιούργησε.

3.ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ

Ο όρος τεχνητή νοημοσύνη (ΤΝ, εκ του Artificial Intelligence) αναφέρεται στον κλάδο της πληροφορικής ο οποίος ασχολείται με τη σχεδίαση και την υλοποίηση υπολογιστικών συστημάτων που μιμούνται στοιχεία της ανθρώπινης συμπεριφοράς τα οποία υπονοούν έστω και στοιχειώδη ευφυΐα: μάθηση, προσαρμοστικότητα, εξαγωγή συμπερασμάτων, κατανόηση από συμφραζόμενα, επίλυση προβλημάτων κλπ. Ο Τζον Μακάρθι όρισε τον τομέα αυτόν ως «επιστήμη και μεθοδολογία της δημιουργίας νοούντων μηχανών».

Η ΤΝ αποτελεί σημείο τομής μεταξύ πολλαπλών επιστημών όπως της πληροφορικής, της ψυχολογίας, της φιλοσοφίας, της νευρολογίας, της γλωσσολογίας και της επιστήμης μηχανικών, με στόχο τη σύνθεση ευφυούς συμπεριφοράς, με στοιχεία συλλογιστικής, μάθησης και προσαρμογής στο περιβάλλον, ενώ συνήθως εφαρμόζεται σε μηχανές ή υπολογιστές ειδικής κατασκευής. Διαιρείται στη συμβολική τεχνητή νοημοσύνη, η οποία επιχειρεί να εξομοιώσει την ανθρώπινη νοημοσύνη αλγοριθμικά χρησιμοποιώντας σύμβολα και λογικούς κανόνες υψηλού επιπέδου, και στην υποσυμβολική τεχνητή νοημοσύνη, η οποία προσπαθεί να αναπαράγει την ανθρώπινη ευφυΐα χρησιμοποιώντας στοιχειώδη αριθμητικά μοντέλα που συνθέτουν επαγωγικά νοήμονες συμπεριφορές με τη διαδοχική αυτοοργάνωση απλούστερων

δομικών συστατικών («συμπεριφορική τεχνητή νοημοσύνη»), προσομοιώνουν πραγματικές βιολογικές διαδικασίες όπως η εξέλιξη των ειδών και η λειτουργία του εγκεφάλου («υπολογιστική νοημοσύνη»), ή αποτελούν εφαρμογή στατιστικών μεθοδολογιών σε προβλήματα ΤΝ.

Η έρευνα στην ανάπτυξη προγραμμάτων ικανών να λύνουν σύνθετα προβλήματα, να καταλαβαίνουν τη γλώσσα που μιλάμε, να “αντιλαμβάνονται” τις οπτικές εικόνες, να διδάσκονται από την εμπειρία και γενικά να επιδεικνύουν μία νοήμονα συμπεριφορά, δημιούργησε ένα νέο τομέα στην επιστήμη της Πληροφορικής, την Τεχνητή Νοημοσύνη (ΤΝ) (Artificial Intelligence ή συντομότερα Α.Ι.). Η ΤΝ εστιάζεται στην ικανότητα του υπολογιστή να χειρίζεται μη-αριθμητικά σύμβολα, να εξάγει συμπεράσματα και να συνάγει νέα γεγονότα από σύνολα γνωστών γεγονότων. Ένας ικανοποιητικός ορισμός της ΤΝ είναι “η μελέτη των ιδεών που επιτρέπουν στους υπολογιστές να είναι ευφυείς”.

Μερικοί από τους κλάδους της ΤΝ είναι:

- Η επεξεργασία φυσικής γλώσσας, που σκοπός της είναι η επικοινωνία του ανθρώπου και του υπολογιστή σε φυσική γλώσσα.
- Τα Έμπειρα Συστήματα, (Expert Systems) που αποσκοπούν στην αποθήκευση στον υπολογιστή, της εμπειρίας ενός ειδικού σε κάποιο τομέα και στην εξασφάλιση της δυνατότητας του υπολογιστή να δρα πλέον ο υπολογιστής ως ειδικός.
- Τα παιχνίδια, με τα οποία μπορεί ο υπολογιστής να παίζει παιχνίδια όπως ένας άνθρωπος κάνοντας συλλογισμούς.
- Η Αναγνώριση Εικόνας, η οποία έχει στόχο να δίδει στον υπολογιστή τη δυνατότητα, μέσω καταλλήλων οπτικών συστημάτων, να βλέπει και να αναγνωρίζει αντικείμενα.
- Η Μηχανική Μάθηση ή Μάθηση της Μηχανής, που έχει σκοπό να δώσει στον υπολογιστή τη δυνατότητα να αυξάνει την απόδοση, αυτοβελτιώνοντας τη γνώση του σε ένα ορισμένο τομέα χωρίς να επεμβαίνει ο άνθρωπος.
- Τα Νευρωνικά Δίκτυα, μέσω των οποίων γίνεται προσομοίωση της λειτουργίας του ανθρώπινου εγκεφάλου σύμφωνα με τις θεωρίες της σύγχρονης Ψυχιατρικής, Ψυχολογίας και Νευροφυσιολογίας.
- Η Ρομποτική, η οποία εξασφαλίζει τη συνεργασία διαφόρων κλάδων της ΤΝ, σε συνδυασμό με ηλεκτρομηχανικές διατάξεις, για την εκτέλεση διαφόρων εξειδικευμένων εργασιών.
- Οι Γενετικοί Αλγόριθμοι (Genetic Algorithms) είναι προσανατολισμένοι στην αναζήτηση της βέλτιστης λύσης μέσα από ένα σύνολο αρχικών πιθανών λύσεων ενός προβλήματος.
- Οι Ευφυείς Πράκτορες (Intelligent Agents) αποτελούν έναν από τους νεώτερους κλάδους της ΤΝ. Ένας πράκτορας είναι μία οντότητα που έχει αντίληψη του περιβάλλοντος στο οποίο ευρίσκεται, επικοινωνεί με αυτό μέσω αι-

σθητήρων και έχει τη δυνατότητα να πραγματοποιεί συλλογισμούς και να επιδρά πάνω σε αυτό.

Σύμφωνα με τους Clancey and Shortliffe (1984), η Ιατρική Τεχνητή Νοημοσύνη (TN) ενδιαφέρεται πρώτιστα για την κατασκευή των προγραμμάτων TN που κάνουν τη διάγνωση και προτείνουν θεραπείες. Αντίθετα από ιατρικές εφαρμογές βασισμένες σε άλλες μεθόδους προγραμματισμού, όπως οι καθαρώς στατιστικές και πιθανολογικές μέθοδοι, τα ιατρικά προγράμματα TN βασίζονται σε συμβολικά πρότυπα (μοντέλα) οντοτήτων ασθενειών και στη σχέση τους με παράγοντες ασθενών και τις κλινικές υλοποιήσεις. Πολλά έχουν αλλάξει από τότε και σήμερα αυτός ο ορισμός θα θεωρείτο στενός στο πεδίο και το όραμα. Σήμερα, η σπουδαιότητα της διάγνωσης ως στόχου που απαιτεί την υποστήριξη της Πληροφορικής σε τυπικές κλινικές καταστάσεις καταλαμβάνει την ίδια θέση με άλλους κλινικούς στόχους.

Ο τομέας Πληροφορικής, στα πλαίσια της TN, που έχει ιδιαίτερα συμβάλει στην Ιατρική τα τελευταία τριάντα χρόνια με την υποβοήθηση στην παρακολούθηση, στη διάγνωση, στη λήψη αποφάσεων καθώς και στην αιτιολόγηση των προτάσεων, είναι τα Έμπειρα Συστήματα. Τα Έμπειρα Συστήματα (ΕΣ) μαζί με τους Ευφυείς Πράκτορες, αποτελούν σήμερα το κατ' εξοχήν εμπορικό προϊόν του τομέα της TN. Είναι προγράμματα τα οποία, στην προσπάθεια τους να αντιμετωπίσουν πολύπλοκα προβλήματα του πραγματικού κόσμου, κωδικοποιούν τη γνώση και τη συλλογιστική ενός Έμπειρου Ειδικού (domain expert) σε έναν εξειδικευμένο τομέα της επιστήμης ή της τεχνολογίας. Κατά συνέπεια, ως Έμπειρο Σύστημα, μπορεί να ορισθεί ένα πρόγραμμα υπολογιστή, που αναφέρεται σε ένα εξειδικευμένο τομέα της ανθρώπινης γνώσης και επιδεικνύει συμπεριφορά ανάλογη με εκείνη του ανθρώπου ειδικού στον εν λόγω τομέα.

Κύρια χαρακτηριστικά των ΕΣ μπορούν να θεωρηθούν τα ακόλουθα:

- Τα αποθέματα γνώσης που εμπεριέχουν.
- Ο απλός, ομοιόμορφος τρόπος με τον οποίο αναπαριστούν τη γνώση.
- Ο αυτόνομος Μηχανισμός Εξαγωγής Συμπερασμάτων που διαθέτουν (inference engine).
- Η διαφάνεια των λειτουργιών τους, η δυνατότητα δηλαδή να επεξηγούν τα συμπεράσματα στα οποία καταλήγουν.

Τα ΕΣ αποδεικνύονται χρήσιμα αφ' ενός μεν διότι οι άνθρωποι-ειδικοί συχνά υπόκεινται σε λάθη και μπορούν να βοηθηθούν σημαντικά χρησιμοποιώντας τα στη λήψη των αποφάσεων, αφ' ετέρου δε, διότι η εύρεση ανθρώπων-ειδικών δεν είναι πάντοτε μία εύκολη υπόθεση και γενικά πάντοτε αναζητούνται τρόποι για τη φύλαξη και τη διάχυση της γνώσης των ειδικών.

Η δομή της παρούσας εργασίας έχει ως εξής: Στο Κεφάλαιο 2 αναφέρονται περι-

ληπτικά τα συστατικά των ΕΣ, ενώ επιχειρείται ανασκόπηση των σημαντικότερων από αυτά που έχουν αναπτυχθεί για την υποβοήθηση τομέων της Ιατρικής. Στο Κε-

φάλαιο 3 παρουσιάζεται ο σχεδιασμός και η αρχιτεκτονική ενός υβριδικού Εμπείρου Συστήματος από την άποψη της ενοποίησης μεθοδολογιών συστημάτων βασισμένων στη γνώση, σχεσιακών Βάσεων Δεδομένων και ασαφών συνόλων επηρεαζόμενων από κριτήρια, ενώ δίνονται και παραδείγματα κανόνων που μπορούν να διαχειριστεί το σύστημα. Τέλος, στο Κεφάλαιο 4 παρουσιάζονται τα συμπεράσματα από τη χρήση ΕΣ στην Ιατρική και τα πλεονεκτήματα του προτεινόμενου υβριδικού Εμπείρου Συστήματος



3.1 ΕΜΠΕΙΡΑ ΣΥΣΤΗΜΑΤΑ ΣΤΗΝ ΙΑΤΡΙΚΗ

3.1.1 ΓΕΝΙΚΑ

Τα Έμπειρα Συστήματα (ΕΣ) αποτελούνται από τα παρακάτω συστατικά (Σχήμα 1), τα οποία αλληλεπιδρούν με στόχο την εξαγωγή συμπεράσματος ή συμπερασμάτων:

- ΒΑΣΗ ΓΝΩΣΗΣ
- Μνήμη Εργασίας (Working Memory)
- Συμπερασματικός Μηχανισμός
- Τμήμα σύλληψης & εσωτερικής αναπαράστασης Γνώσης
- Μηχανισμός επεξήγησης
- Σύστημα Διεπαφής Χρήστη
- ΧΡΗΣΤΗΣ

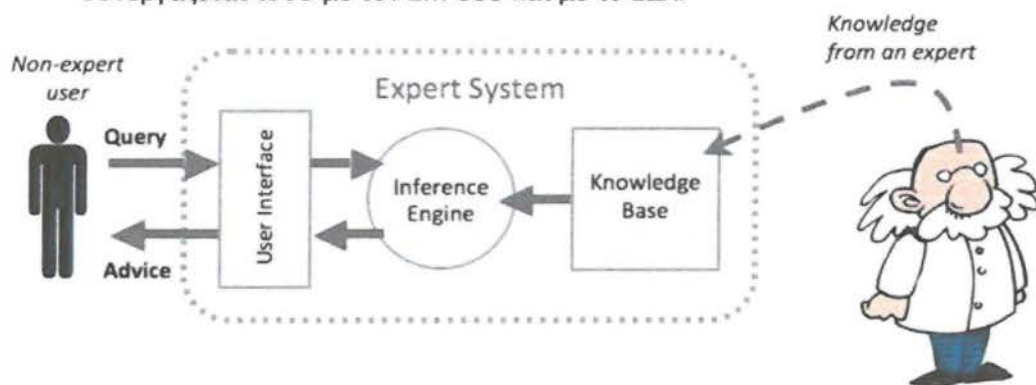
Σχήμα 1. Έμπειρο Σύστημα

Ως βασικά συστατικά θεωρούνται:

- Η Βάση Γνώσης (ΒΓ) (Knowledge Base), η οποία περιέχει κανόνες και γεγονότα που αναπαριστώνται με κατανοητό από τους ανθρώπους τρόπο.
- Ο Συμπερασματικός Μηχανισμός (ΣΜ) (Inference Engine), που είναι η καρδιά του ΕΣ, ο οποίος επεξεργάζεται τους κανόνες και τα γεγονότα της ΒΓ και φθάνει σε μία λύση.
- Το Σύστημα Διεπαφής Χρήστη (ΣΔΧ) (User Interface), ως μέσο επικοινωνίας μεταξύ του χρήστη και του ΕΣ.

Ενώ τα παρακάτω συστατικά είναι απαραίτητα για την εσωτερική λειτουργία ενός ΕΣ:

- Η Μνήμη Εργασίας στην οποία αποθηκεύονται ως νέα γεγονότα οι απαντήσεις του χρήστη που χρησιμοποιούνται στη συμπερασματική διαδικασία. Επίσης στη Μνήμη Εργασίας τοποθετούνται τα ενδιαμέσα και τα τελικά συμπεράσματα..
- Το Τμήμα Σύλληψης και Εσωτερικής Αναπαράστασης της Γνώσης το οποίο είναι υπεύθυνο για την εσωτερική αναπαράσταση της Γνώσης μέσα στη ΒΓ. Η ΒΓ γράφεται, τις περισσότερες φορές σε μία μορφή κειμένου ή ακόμη με κάποιο τυποποιημένο τρόπο που έχει σχεδιάσει ο κατασκευαστής.
Το τμήμα αυτό ελέγχει την ορθότητα της ΒΓ, τη μετατρέπει σε μία εσωτερική μορφή κατανοητή για το σύστημα, αλλά ταυτόχρονα η εσωτερική αυτή μορφή έχει τη δυνατότητα να μετατραπεί σε αυτή που είχε αρχικά δοθεί.
- Ο Μηχανισμός Επεξήγησης (Explanation Mechanism) είναι αυτός που δίδει στο χρήστη την απάντηση γιατί το σύστημα του απευθύνει μία συγκεκριμένη ερώτηση ή πώς έφθασε σε κάποιο συμπέρασμα. Ο μηχανισμός αυτός συνεργάζεται τόσο με τον ΣΜ όσο και με το ΣΔΧ.



3.1.2 ΧΡΗΣΗ ΕΜΠΕΙΡΩΝ ΣΥΣΤΗΜΑΤΩΝ ΣΤΗΝ ΙΑΤΡΙΚΗ

Τα Έμπειρα Συστήματα (ΕΣ) χρησιμοποιούνται στην ιατρική με ένα συμβουλευτικό τρόπο από ειδικούς του τομέα ώστε να ενισχύσουν τη διαδικασία επίτευξης ορθολογικών συμπερασμάτων. Είναι γνωστό, ότι η διάγνωση είναι μόνο ένα από πολλά προβλήματα στην κλινική Ιατρική. Σήμερα, τα ιατρικά Έμπειρα Συστήματα εφαρμόζονται κυρίως στους ακόλουθους τύπους κλινικών εργασιών:

- Παραγωγή προειδοποιήσεων και των υπενθυμίσεων. Σε καταστάσεις πραγματικού χρόνου, ένα ΕΣ συνδεδεμένο με ένα όργανο ελέγχου, προειδοποιεί για μεταβολές της κατάστασης του ασθενούς. Διαφορετικά, ανιχνεύει αποτε-

λέσματα εργαστηριακών δοκιμών και στέλνει υπενθυμίσεις ή προειδοποιήσεις.

- Διαγνωστική βοήθεια. Όταν μια περίπτωση ασθενούς είναι περίπλοκη (ή σπάνια), ή αυτός που κάνει τη διάγνωση είναι άπειρος, ένα ΕΣ βοηθά δίνοντας πιθανές διαγνώσεις βασισμένες στα στοιχεία του ασθενούς.

- Κριτική και σχεδιασμός θεραπείας. Τα ΕΣ χρησιμοποιούνται στην αναζήτηση ασυνεπειών και παραλείψεων σε ένα υφιστάμενο σχέδιο θεραπείας, ή στη σχηματοποίηση μια θεραπείας βασισμένης σε συγκεκριμένη κατάσταση του ασθενούς και σε υιοθετημένες οδηγίες θεραπείας.

- Εκπαίδευση. Τα ΕΣ χρησιμοποιούνται στην εκπαίδευση και εξάσκηση νοσοκομειακών ιατρών και σπουδαστών σε διάφορες ιατρικές εργασίες.

Παρά την αυξανόμενη χρήση των έμπειρων συστημάτων στην ιατρική, πρέπει να τονιστεί ότι η ανάπτυξη τέτοιων συστημάτων για τις ιατρικές εφαρμογές πρέπει να ξεπεράσει μερικά βασικά εμπόδια, όπως:

- Οι ιατρικές εργασίες είναι δύσκολες λόγω των ιδιαιτεροτήτων των ασθενών και της αβεβαιότητας των διαθέσιμων κλινικών στοιχείων.

- Το εύρος των αποδεκτών σφαλμάτων είναι πολύ μικρό λόγω των ηθικών ανησυχιών και των κινδύνων λανθασμένων ιατρικών πράξεων.

- Η χρηματοδότηση για τις κύριες δαπάνες είναι μικρή.

Από την άλλη πλευρά, οι παράγοντες που ευνοούν την αυξανόμενη διάδοση της τεχνολογίας των έμπειρων συστημάτων είναι:

- αποτελεσματικότητα διαχείρισης δαπανών και

- βελτιωμένη ποιότητα φροντίδας των ασθενών.

3.1.3 ΑΝΑΣΚΟΠΗΣΗ ΙΑΤΡΙΚΩΝ ΕΜΠΕΙΡΩΝ ΣΥΣΤΗΜΑΤΩΝ

Στην πρώτη δεκαετία της Τεχνητής Νοημοσύνης στην Ιατρική, τα περισσότερα ερευνητικά συστήματα αναπτύχθηκαν για να βοηθήσουν τους νοσοκομειακούς γιατρούς στο στάδιο της διάγνωσης, ιδιαίτερα με την πρόθεση ότι θα χρησιμοποιούταν με τη συγκατάθεση και τη βοήθεια του ασθενή. Τα περισσότερα από αυτά δεν αναπτύχθηκαν πέρα από το ερευνητικό εργαστήριο. Άλλα, εντούτοις, συνέχισαν να αναπτύσσονται και έχουν μετασηματιστεί στα εκπαιδευτικά συστήματα. Τα κυριότερα ΕΣ είναι τα παρακάτω:

MYCIN

Είναι, αυτή τη στιγμή, πιθανώς το πλέον γνωστό από όλα τα ιατρικά (και όχι μόνο) ΕΣ που έχουν αναπτυχθεί (Shortliffe and Buchanan, 1975) (Shortliffe, 1976). Και αυτό συμβαίνει, παρά το γεγονός ότι δεν έχει τεθεί ποτέ σε πραγματική πρακτική.

Αναπτύχθηκε στο Stanford Πανεπιστήμιο απλώς ως ερευνητική προσπάθεια με σκοπό να παρασχεθεί βοήθεια στους παθολόγους στη διάγνωση και στον καθορισμό αντιμικροβιακής θεραπείας των προσβολών ορισμένων μολυσματικών νοσημάτων (μηνιγγίτιδας και βακτηραιμίας). Η παράσταση των γνώσεων του συστήματος χρησιμοποιεί περίπου 500 κανόνες παραγωγής. Κάθε κανόνας έχει και ένα παράγοντα βεβαιότητας που καθορίζει την ισχύ της σχέσης συνθήκης και συμπεράσματος.

Σύμφωνα με τους ισχυρισμούς της ομάδας, η απόδοση του συστήματος είναι συγκρίσιμη με εκείνη των ειδικών. Ακολούθησε (1979) το EMycin, το οποίο θεωρείται ως το πρώτο περιβάλλον ανάπτυξης ΕΣ.

PUFF

Αναπτύχθηκε το 1979 (Yu et al, 1979), χρησιμοποιώντας το EMYCIN. Αποσκοπεί στη διάγνωση πνευμονολογικών νοσημάτων. Η διάγνωση στηρίζεται στο ιστορικό του ασθενή και σε μετρήσεις όγκου πνεύμονα, ροής αέρα και ικανότητας διάχυσης από τους πνεύμονες στο αίμα. Συγκρίνοντας τα στοιχεία αυτά με τα κανονικά το σύστημα προτείνει μια διάγνωση για τον τύπο και τη σοβαρότητα του πνευμονολογικού νοσήματος. Η παράσταση των γνώσεων του συστήματος χρησιμοποιεί περίπου 250 κανόνες παραγωγής. Το σύστημα μπορεί να προτείνει διαγνώσεις για άσθμα, βρογχίτιδα, εμφύσημα και άλλα νοσήματα.

CENTAUR

Το σύστημα CENTAUR έχει παρόμοιους σκοπούς με το PUFF αλλά χρησιμοποιεί πιο εξελιγμένη μεθοδολογία. Η παράσταση των γνώσεων του συστήματος χρησιμοποιεί εκτός από κανόνες παραγωγής και δομές πρότυπα. Τα πρότυπα αντιστοιχούν σε συγκεκριμένα νοσήματα και καθοδηγούν τη συμβουλευτική διαδικασία. Η ομάδα ισχυρίζεται ότι υπάρχει συμφωνία μέχρι 90% με ειδικό πνευμονολόγο.

INTERNIST

Αναπτύχθηκε στη δεκαετία του '80 στο Πανεπιστήμιο του Πίτσμπουργκ (Miller et al, 1982). Ο σκοπός του είναι να εκτελέσει διάγνωση της πλειοψηφίας των ασθενειών που συνδέονται με τον τομέα της παθολογίας. Είχε το στόχο να εξετάσει όλους τους πιθανούς συνδυασμούς ασθενειών που είναι δυνατόν να συμβούν σε κάποιον άνθρωπο. Αποτίμηση της συνεισφοράς του έγινε από τον Wolfram (1995).

ILLIAD

Ήταν υπό ανάπτυξη για αρκετά έτη (Lincoln et al, 1991), πρώτιστα για τη διάγνωση στην παθολογία, ενώ έφτασε στο σημείο να καλύπτει περίπου 1500 διαγνώσεις βασισμένες σε αρκετές χιλιάδες ευρήματα (Anderson et al, 1995). Η τωρινή χρήση του είναι ως εργαλείο διδασκαλίας για σπουδαστές Ιατρικής.

ToxoNet

Το ToxoNet είναι ένα ΕΣ σχεδιασμένο να δίνει στο νοσοκομειακό ιατρό πληροφορίες για πιθανή μόλυνση με Τοξόπλασμα (*Toxoplasma gondii*), αναλύοντας τα αποτελέσματα εξετάσεων αίματος εγκύων γυναικών. Επιπρόσθετα, βοηθά στη διαδικασία λήψης απόφασης σχετικά με την ευαισθησία χρήσης ορολογικών ελέγχων. Βασίστηκε στο Toxopert (Nagy et al, 1997).

DXplain

Αναπτύχθηκε στο Γενικό Νοσοκομείο της Μασαχουσέτης (Barnett et al, 1987) και χρησιμοποιείται για να βοηθήσει τη διαδικασία της διάγνωσης, λαμβάνοντας υπόψη ένα σύνολο κλινικών ευρημάτων συμπεριλαμβανομένων ενδείξεων, συμπτωμάτων και εργαστηριακών αποτελεσμάτων. Στη συνέχεια παράγει έναν ταξινομημένο κατάλογο διαγνώσεων. Το DXplain χρησιμοποιείται σε διάφορα νοσοκομεία και συνήθως για λόγους κλινικής εκπαίδευσης.

HELP

Το HELP (Health Evaluation through Logical Processing) είναι ένα Νοσοκομειακό Πληροφοριακό Σύστημα βασισμένο στη γνώση (Knowledge-based Hospital Information System) (Gardner et al, 1999). Παρέχει στους νοσοκομειακούς Ιατρούς προειδοποιήσεις και υπενθυμίσεις, ερμηνεία δεδομένων και διαγνώσεις ασθενών, καθώς και προτάσεις διαχείρισης ασθενών και κλινικά πρωτόκολλα. Το HELP λειτουργεί σε έξι μεγάλα νοσοκομεία στη Utah και σε άλλες πολιτείες της Αμερικής.

DoseChecker

Αναπτύχθηκε (McMullin et al, 1997) στο Νοσοκομείο Barnes του Μισούρι προκειμένου να βοηθηθούν οι φαρμακοποιοί στον έλεγχο των φαρμακευτικών παραγγελιών για ένα σύνολο φαρμάκων των οποίων η δόση πρέπει να υπολογιστεί προσεκτικά για ασθενείς με πιθανή νεφρική ανεπάρκεια.

QMR

Αναπτύχθηκε στο πανεπιστήμιο του Πίτσμπουργκ (First et al, 1985), με σκοπό την παροχή βοήθειας στους παθολόγους στη διάγνωση ασθενειών, βασιζόμενο στα συμπτώματα των ασθενών, τα συμπεράσματα των εξετάσεων και τα εργαστηριακά ευρήματα.

PEIRS

Το PEIRS (Pathology Expert Interpretative Reporting System) είναι ένα Έμπειρο Ερμηνευτικό Σύστημα Παθολογίας, το οποίο, αποτιμάται (Edwards et al, 1993) ότι ερμηνεύει περίπου 80-100 αναφορές τη ημέρα, με διαγνωστική ακρίβεια περίπου 95% στον Τομέα Χημικής Παθολογίας του Institute of Laboratory Medicine στο St. Vincent's Hospital, Sydney.

APACHE III

Σχεδιάστηκε να προβλέπει το κίνδυνο θανάτου στο νοσοκομείο. Συγκρίνει το ιατρικό προφίλ του ασθενούς με 18.000 περιπτώσεις που έχει στη Βάση Δεδομένων του και καταλήγει σε πρόγνωση με 95% ακρίβεια (κατά μέσον όρο). Υπάρχουν 16 νοσοκομεία στις ΗΠΑ που χρησιμοποιούν το ΕΣ APACHE III.

MDDB

Σχεδιάστηκε (Thomsen, 1997) και λειτουργεί επιχειρησιακά, με σκοπό να βοηθηθούν οι παθολόγοι στη διάγνωση των δυσμορφικών συνδρόμων. Μέχρι τώρα το σύστημα έχει χρησιμοποιηθεί σε περίπου 3.000 ασθενείς και όλη η γνώση για αυτούς τους ασθενείς ενσωματώθηκε στη Βάση Γνώσης.

HEPAXPERT I, II

Αναπτύχθηκε προκειμένου να ερμηνευθούν οι έλεγχοι για την ηπατίτιδα Α και Β και χρησιμοποιείται στο Εργαστήριο ηπατίτιδας της Ιατρικής Σχολής του Πανεπιστημίου της Βιέννης από το Σεπτέμβριο του 1989.

JEREMIAH

Σχεδιάστηκε (Brown et al, 1991) για να παρέχει στους οδοντιάτρους γνώση ορθοδοντικών τεχνικών για τις περιπτώσεις που είναι κατάλληλες για θεραπεία.

ATTENDING

Το σύστημα ATTENDING αποσκοπεί στην ανάλυση ιατρικών σχεδίων αναισθησίας. Το σύστημα έχει δύο "τρόπους" (modes) χρήσης: τον συμβουλευτικό και τον εκπαιδευτικό. Στον συμβουλευτικό τρόπο ο αναισθησιολόγος περιγράφει τα προβλήματα του ασθενή και προτείνει ένα αναισθησιολογικό σχέδιο που το σύστημα

πρέπει να σχολιάσει. Στον εκπαιδευτικό τρόπο το σύστημα περιγράφει έναν υποθετικό ασθενή στον εκπαιδευόμενο που πρέπει να προτείνει ένα σχέδιο αναισθησίας και που το οποίο το σύστημα το σχολιάζει στη συνέχεια.

EXPERT-D

Το σύστημα "EXPERT-D" αποσκοπεί στη διαφορική διάγνωση δερματολογικών νοσημάτων. Η ανάπτυξη του συστήματος έγινε με βάση το γενικό πλαίσιο παράστασης γνώσεων EXPERT. Η ομάδα ανάπτυξης ισχυρίζεται ότι το σύστημα έχει επιδείξει ακρίβεια μέχρι 84%.

CAA

Το σύστημα CAA αποσκοπεί στη διάγνωση αρρυθμιών με βάση το ηλεκτροκαρδιογράφημα (ΗΚΓ). Το σύστημα αποτελείται από μια πολυεπίπεδη βάση γνώσεων καρδιολογικής ηλεκτροφυσιολογίας και μορφολογίας του ΗΚΓ και μια δομή ελέγχου για την αναγνώριση σημάτων ΗΚΓ που χρησιμοποιεί αιτιακή γνώση, δεσμούς ομοιότητας και ιεραρχικές δομές. Η αιτιακή χρονική σχέση μεταξύ των εσωτερικών γεγονότων και των παρατηρουμένων ΗΚΓ σημάτων περιγράφεται με σημασιολογικά δίκτυα.

SYSTEM D (Dizziness)

Το σύστημα SYSTEM D αποσκοπεί στη διάγνωση νοσημάτων με κύριο σύμπτωμα τη ζάλη (dizziness). Πρόκειται για πολύπλοκο διαγνωστικό πρόβλημα λόγω του μεγάλου αριθμού των πιθανών αιτιών. Η βάση γνώσεων του συστήματος περιέχει πληροφορίες για περίπου 50 αιτίες ζάλης.

ONCOCIN

Το σύστημα ONCOCIN αποσκοπεί στην υποστήριξη της θεραπευτικής αγωγής καρκινοπαθών. Στην αρχική του μορφή λειτουργούσε αυτοδύναμα, αλλά στη συνέχεια, για λόγους καλύτερης αποδοχής από τους χρήστες, διαμορφώθηκε έτσι ώστε να επεμβαίνει μόνο όταν ανακαλύπτει διαφορές στο θεραπευτικό σχέδιο του θεράποντος και στο δικό του εξηγώντας τους λόγους των διαφορών. Η βάση γνώσεων του συστήματος στηρίζεται σε πρωτόκολλα αντικαρκινικής θεραπείας.

CASNET

Το σύστημα CASNET (Kulikowski and Weiss, 1982) αποσκοπεί στη διάγνωση οφθαλμολογικών νοσημάτων και ειδικότερα του γλαυκώματος. Η παράσταση ενός νοσήματος γίνεται με μια δυναμική διαδικασία χρησιμοποιώντας ένα δίκτυο από αιτιακά συνδεδεμένες παθοφυσιολογικές καταστάσεις. Εκτός από τη διάγνωση το σύστημα μπορεί να προσδιορίσει θεραπείες και να κάνει προγνώσεις. Ο αριθμός των καταστάσεων έχει φτάσει τις 150. Σύμφωνα με τους ισχυρισμούς της ομάδας το σύστημα αποδίδει σε επίπεδο συγκρίσιμο με ειδικούς.

ANTICIPATOR

Το σύστημα ANTICIPATOR αποσκοπεί στην παροχή συμβουλών για αντιβιοτική θεραπεία μολυσματικών νοσημάτων. Η βάση γνώσεων περιέχει περιγραφική και ευριστική γνώση. Η περιγραφική γνώση αποτελείται από ιεραρχίες αντιβιοτικών, βακτηριδίων και μολυσματικών νοσημάτων από πίνακες φάσματος αντίστασης, από φαρμακοκινητικούς πίνακες, από στατιστικούς αιτιολογικούς πίνακες και από κλινικές φαρμακολογικές πληροφορίες. Η ευριστική γνώση αποτελείται από λογικούς κανόνες παραγωγής και διαδικαστικές προσθήκες που περιγράφουν πώς πρέπει να

συγκεντρωθούν τα στοιχεία περιγραφής της κάθε περίπτωσης. Το σύστημα έχει υλοποιηθεί με γλώσσα Prolog/KR που είναι μια επέκταση της γλώσσας Prolog κατάλληλη για παράσταση γνώσεων. Το σύστημα χρησιμοποιείται από νοσοκομειακούς γιατρούς.

CADUCEUS

Το σύστημα CADUCEUS αποσκοπεί στη διάγνωση 500 περίπου νοσημάτων παθολογίας. Είναι διάδοχος του συστήματος INTERNIST.

MedFrame/CADIAG-IV

Το σύστημα MedFrame/CADIAG-IV είναι η τελευταία εξέλιξη του ερευνητικού έργου CADIAG που ξεκίνησε το 1968 βασισμένο στη Boolean λογική. Η παρούσα έκδοση είναι ένα διαγνωστικό και συμβουλευτικό σύστημα, βασισμένο σε ασαφείς κανόνες. Σχεδιάστηκε και υλοποιήθηκε από το Τμήμα Ιατρικής Πληροφορικής του Πανεπιστημίου της Βιέννης με σκοπό την υποστήριξη λήψης απόφασης στη διάγνωση και θεραπεία στον τομέα της Παθολογίας (Leitich et al, 1996). Λειτουργεί στο Γενικό Νοσοκομείο της Βιέννης, ενώ έχει ενσωματωθεί με το ιατρικό πληροφοριακό σύστημα του Νοσοκομείου εκμεταλλευόμενο τόσο τα διοικητικά όσο και τα ιατρικά δεδομένα της ΒΔ των ασθενών.

AI/RHEUM

Το σύστημα AI/RHEUM (Lindberg et al, 1980) αποσκοπεί στη διάγνωση ρευματολογικών νοσημάτων. Η διάγνωση στηρίζεται στην ανάλυση 875 ευρημάτων και επιλέγει πιο από τα 26 νοσήματα που γνωρίζει το σύστημα συμφωνεί με τα ευρήματα. Η βάση γνώσεων έχει αναπτυχθεί με βάση το γενικό πλαίσιο EXPERT και αποτελείται από 1000 περίπου κανόνες. Η ομάδα ισχυρίζεται ότι η ακρίβεια του συστήματος κυμαίνεται μεταξύ 87% και 100% ανάλογα με την κατηγορία νοσημάτων.

ISP

Το σύστημα ISP αποσκοπεί στην υποστήριξη της σχεδίασης μιας χειρουργικής επέμβασης. Μια σειρά τομογραφικών εικόνων χρησιμοποιείται για την εισαγωγή γεωμετρικών πληροφοριών. Ο χειρουργός επικοινωνεί με το σύστημα με γραφική οθόνη και καθορίζει τους χειρουργικούς χειρισμούς. Το σύστημα εκτελεί με προσομοίωση τους χειρισμούς στα πρότυπα των οργάνων του ασθενή και προβάλλει τα αποτελέσματα στη γραφική οθόνη. Ο χειρουργός επαναλαμβάνει τη διαδικασία μέχρι να ικανοποιηθεί από τα αποτελέσματα.

SEER

Το σύστημα SEER αποσκοπεί στην ανάλυση ηλεκτροεγκεφαλογραφημάτων. Η βάση γνώσεων αποτελείται από κανόνες που αναγνωρίζουν διάφορες "δραστηριότητες" στο ΗΚΓ όπως άλφα, βήτα, δέλτα και χαμηλού πλάτους. Αν τα αποτελέσματα της ανάλυσης δεν είναι ικανοποιητικά τότε ο χρήστης μπορεί να τροποποιήσει τους κανόνες της βάσης γνώσεων.

MUNIN

Το σύστημα MUNIN (Olesen et al, 1990) αναπτύχθηκε με στόχο τη διάγνωση νευρολογικών διαταραχών, βασιζόμενο σε δίκτυα βασισμένα στην θεωρία του Bayes (Bayesian networks) (Dempster, 1968).

HEADMED

Το σύστημα HEADMED αποσκοπεί στην παροχή ψυχοφαρμακολογικών συμβουλών. Τα στοιχεία εισόδου είναι η σοβαρότητα και η φύση των συμπτωμάτων των ασθενών, η ψυχιατρική ιστορία των ίδιων και των συγγενών τους, οι αντιδράσεις τους σε φαρμακευτική αγωγή και τα ψυχιατρικά τους προβλήματα.

VM

Το σύστημα VM (Ventilator Manager) (1977-81) ήταν ένα ΕΣ που ερμήνευε ποσοτικά δεδομένα που έρχονταν από την Μονάδα Εντατικής Θεραπείας και συμβούλευε τους Ιατρούς στην αντιμετώπιση χειρουργημένων ασθενών που χρειάζονταν μηχανική υποβοήθηση για να αναπνέουν. Αν και ήταν στηριγμένο στην αρχιτεκτονική του MYCIN, το VM ήταν σχεδιασμένο έτσι ώστε να ανταποκρίνεται σε δεδομένα που αλλάζουν με την πάροδο του χρόνου. Κατ' αυτόν τον τρόπο μπορούσε να παρακολουθεί την πρόοδο των ασθενών, να λαμβάνει υπόψη του το παρόν και το παρελθόν ενός ασθενή, και να προτείνει αλλαγές στην θεραπεία του. Το VM δοκιμάστηκε στην Χειρουργική Μονάδα Εντατικής Θεραπείας του νοσοκομείου Pacific Medical Center στο San Francisco.

MOLGEN

Το πρόγραμμα MOLGEN (1975-84) χρησιμοποιούσε μεθόδους της ΤΝ στην έρευνα της μοριακής βιολογίας. Αρχικά αποτέλεσε ένα ΕΣ που περιείχε την γνώση που είναι απαραίτητη για τον σχεδιασμό και την εξομοίωση πειραμάτων. Αργότερα εξελίχτηκε σε δύο ΕΣ που εφάρμοζαν διαφορετικές προσεγγίσεις στον σχεδιασμό πειραμάτων γενετικής. Το ένα ΕΣ χρησιμοποιούσε «σκελετικά σχέδια», που είναι αφηρημένες περιγραφές πειραμάτων, για συγκεκριμένους πειραματικούς στόχους και περιβάλλοντα. Το άλλο ΕΣ ήταν βασισμένο στον σχεδιασμό βάση περιορισμών, στον οποίο οι σχεδιαστικές επιλογές γίνονται στα πλαίσια μιας γενικότερης στρατηγικής. Αυτά τα δύο συστήματα αργότερα αποτέλεσαν το SPEX. Σήμερα το πρόγραμμα ονομάζεται MOLGEN-II.

Οι σύγχρονες διαγνωστικές μέθοδοι της ψυχιατρικής βασίζονται κατά κύριο λόγο στους κανόνες ακολουθιακής λογικής δενδροειδούς δομής. Είναι λοιπόν φυσικό τα ΕΣ να αποτελούν την προσφιλέστερη προσέγγιση στο πρόβλημα της ψυχιατρικής διάγνωσης μέσω υπολογιστών. Υπάρχουν πολλά τέτοια συστήματα στη βιβλιογραφία και μερικά από τα πιο επιτυχημένα είναι το DSM III / DIS και το PSE/CATEGO (Mauger et al, 1989). Το κλασικό και πολύ σημαντικό ΕΣ PSE/CATEGO εφαρμόστηκε στην βάση 796 κλινικών συνεντεύξεων. Παρόλο που τα αποτελέσματα που προέκυψαν δεν μπορούν απ' ευθείας να συγκριθούν με τα αντίστοιχα που πέτυχαν άλλες στατιστικές μέθοδοι βασισμένες στην αναγνώριση προτύπων, μας δείχνουν αρκετές σπουδαίες τάσεις γύρω από την απόδοση των ΕΣ. Τα Έμπειρα αυτά Συστήματα σχεδιάστηκαν λαμβάνοντας υπόψη πολύ διαφορετικούς πληθυσμούς ασθενών και συνεπώς είναι πολύ γενικά για να εφαρμοστούν με επιτυχία στις ανάγκες ενός συγκεκριμένου πληθυσμού. Τα ΕΣ CATEGO και DSM-III πέτυχαν 81% και 79% σε ακρίβεια κατηγοριοποίησης αντίστοιχα, για δείγματα σχιζοφρενών όπως επίσης 43% και 45% αντίστοιχα σε ακρίβεια αναγνώρισης (διάκρισης) για δείγματα ασθενών με κατάθλιψη.

Επίσης, ένα υβριδικό ΕΣ αναπτύχθηκε στην Κίνα (Lynn Ling, 1999) το οποίο

χρησιμοποιεί είτε επαγωγικές είτε συμπερασματικές μεθόδους με σκοπό τη διάγνωση και αποτίμηση ασθενών πιθανόν προσβεβλημένων από AIDS. Το σύστημα δοκιμάστηκε με πραγματικά δεδομένα και τα αποτελέσματα είναι ενθαρρυντικά.

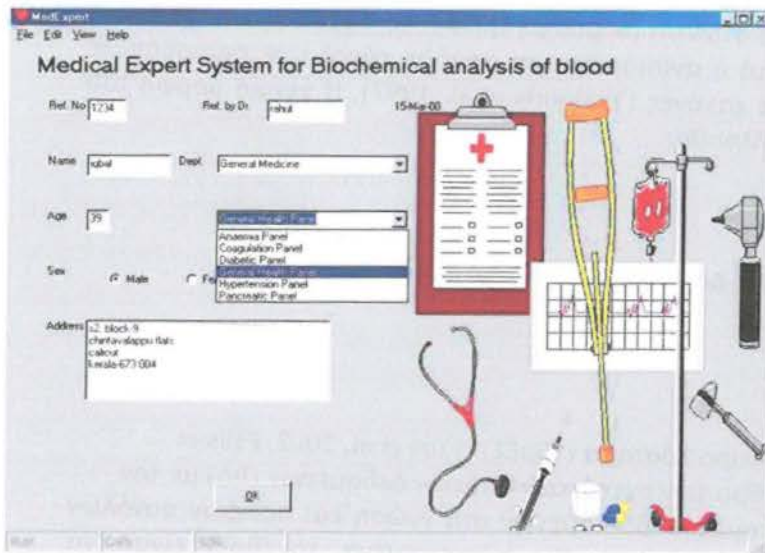
Από τις πρόσφατες εφαρμογές ξεχωρίζει η μελέτη εφαρμογής ΕΣ (Schechtman et al, 2005) για ειδικευόμενους παθολόγους που εξυπηρετούν κυρίως άπορο πληθυσμό στην κεντρική Βιρτζίνια, όπου 94 οι ιατροί εξετάζουν κατά μέσο όρο περίπου 28.000 επισκέψεις ασθενών ετησίως. Το ΕΣ συνταγών (Practice Partner® v7.5, Physician Micro Systems, Inc.) υλοποιήθηκε το Σεπτέμβριο του 2003 και επέτρεψε την ηλεκτρονική διαχείριση καταλόγων συνταγών φαρμάκων, την εκτύπωση και ανανέωση των συνταγών, του ελέγχου των φαρμάκων, καθώς και αλλεργικών αλληλεπιδράσεων μεταξύ φαρμάκων. Μετά από έρευνα εφαρμογής έξι μηνών 72% των εθελοντών ιατρών αξιολόγησε θετικά το σύστημα.

Πέρα από τα παραπάνω, πρέπει να αναφερθούν ΕΣ στα οποία αξιοποιούνται οι δυνατότητες της ασαφούς λογικής (Fuzzy Logic) (Lotfi Zadeh, 1965) και των ασαφών συνόλων (Fuzzy Sets) με σκοπό τη διαχείριση εννοιών και βελτιστοποίηση αποτιμήσεων που εμπεριέχουν την ασάφεια ως έναν βαθμό.

Ένα ΕΣ πραγματικού χρόνου, βασισμένο στη συλλογιστική ασαφούς λογικής, αναπτύχθηκε από τους Greenhow et al (1992) για τον έλεγχο του αναισθητικού, καθώς επίσης και κάποια ΕΣ για την Ψυχιατρική επιστήμη (Konacs and Juranovics, 1995), (Karlan et al, 1997). Οι Garibaldi and Ifeachor (2000) περιγράφουν ένα ασαφές ΕΣ για την ανάλυση του αίματος του ομφάλιου λώρου των νεογέννητων βρεφών μετά τον τοκετό. Το σύστημα έχει τη δυνατότητα παροχής πολύτιμων πληροφοριών για την υγεία του νεογέννητου και παρέχει οδηγίες για νεογνική φροντίδα. Και σε αυτό το ΕΣ χρησιμοποιήθηκε η Ασαφής Λογική λόγω των σφαλμάτων των δεδομένων τα οποία κατέστησαν άλλες τεχνικές αδύναμες να διερμηνεύσουν τα αποτελέσματα με ακρίβεια. Στην Παθολογία, χρησιμοποιήθηκε ένα ΕΣ βασισμένο στην ασαφή λογική, για την επιλογή κατάλληλης θεραπευτικής στρατηγικής (van Ginneken and Smeulders 1991).

Στην εργασία των (Mahfouf M. et al, 2001) γίνεται αποτίμηση των συστημάτων παρακολούθησης και ελέγχου στην Ιατρική με χρήση μεθόδων ασαφούς λογικής, ενώ στην έρευνα χρήσης ασαφούς λογικής στην Ιατρική και στην υγιεινή (Maysam F. et al, 2001) γίνεται εκτενέστατη ανασκόπηση και κατηγοριοποίηση ανά ιατρικό τομέα, όλων των παρεμφερών συστημάτων.

Τέλος, πρέπει να αναφερθεί η επισήμανση των Crowley and Gryzbicki (2006) για τη συνεισφορά των ευφυών ιατρικών εκπαιδευτικών συστημάτων (Intelligent medical training systems - IMTS) στην Ιατρική, καθώςσον εκτιμάται (Kohn et al, 2000) ότι με- ταξύ 44.000 και 98.000 Αμερικανοί πεθαίνουν κάθε χρόνο εξαιτίας λανθασμένης ιατρικής φροντίδας.



3.2 ΥΒΡΙΔΙΚΟ ΕΜΠΕΙΡΟ ΣΥΣΤΗΜΑ

3.2.1 ΓΕΝΙΚΑ

Τα προαναφερθέντα Έμπειρα Συστήματα, ειδικά τα παλαιότερα, είχαν ένα βασικό μειονέκτημα: τον τρόπο αποθήκευσης της Βάσης Γνώσης (ΒΓ), καθόσον αυτή δεν ακολουθούσε καμία μορφή τυποποίησης και τηρείτο σε μορφή μη εκμεταλλεύσιμη από άλλα συστήματα. Αυτό είχε ως επακόλουθο, την απαρχής δόμηση της ΒΓ από το Μηχανικό Γνώσης (Knowledge Engineer) μαζί με τον ειδικό εμπειρογνώμονα (Expert) σε περίπτωση αλλαγής του ΕΣ.

Το θεμελιώδες συστατικό της ΒΓ σε ένα Rule Based Expert System αποτελεί ο Παραγωγικός Κανόνας ο οποίος είναι μια υποθετική δήλωση δύο μερών. Το πρώτο μέρος, που αποτελείται από έναν ή περισσότερους 'ΕΑΝ' όρους (ισχυρισμούς), προσδιορίζει τις συνθήκες (προϋποθέσεις) του κανόνα. Το δεύτερο μέρος, που αποτελείται από ένα ή περισσότερους 'ΤΟΤΕ' όρους (δηλώσεις), προσδιορίζει τα συμπεράσματα του κανόνα και τελικά, το πρώτο μέρος εφαρμόζει με το δεύτερο. Οι

όροι και των δύο μερών, αναπαριστώνται με μορφή τριπλέτας 'Αντικείμενο – Ιδιότητα – Τιμή' (ΑΙΤ), δεδομένου ότι αυτή η αναπαράσταση αρμόζει εύκολα σε οποιοδήποτε Έμπειρο Σύστημα βασιζόμενο σε κανόνες (Yialouris et al, 1997). Η γενική μορφή του Παραγωγικού Κανόνα είναι η ακόλουθη:

EAN < Ισχυρισμός_Ι1 >
ΤΟΤΕ < Δήλωση_Δ1 >
Η / ΚΑΙ < Ισχυρισμός_Ι2 > Η / ΚΑΙ < Δήλωση_Δ2 >
Η / ΚΑΙ... Η / ΚΑΙ...
Η / ΚΑΙ < Ισχυρισμός_Ιn >
Η / ΚΑΙ < Δήλωση_Δn >

Στο προτεινόμενο Υβριδικό Έμπειρο Σύστημα (ΥβρΕΣ) (Filis et al, 2002, Filis et al, 2003) αξιοποιείται το υπόβαθρο των σχεσιακών Βάσεων Δεδομένων (ΒΔ) με την ενοποίηση μεθοδολογιών συστημάτων βασισμένων στη γνώση και ασαφών συνόλων δυνάμενων να επηρεάζονται από κριτήρια. Πιο συγκεκριμένα, το ΥβρΕΣ είναι ένα Έμπειρο Σύστημα βασισμένο σε κανόνες (Rule Based Expert System) το οποίο λειτουργεί αποθηκεύοντας τους κανόνες του σε μορφή τριπλέτας 'Αντικείμενο – Ιδιότητα – Τιμή' (Α-Ι-Τ) σε μία σχεσιακή Βάση Δεδομένων. Κάθε κανόνας ισχύει για ένα συγκεκριμένο χρονικό διάστημα οπότε το όλο μοντέλο είναι έγκυρο για αυτά τα χρονικά διαστήματα χρησιμοποιώντας έγκυρες εκδόσεις γνώσης. Οι τιμές των κανόνων καθορίζονται από ασαφή σύνολα και χρησιμοποιούνται ως γλωσσικές μεταβλητές (linguistic variables). Επίσης, εφόσον πολλά ΕΣ αξιοποιούν κάποιο είδος ασάφειας, το ΥβρΕΣ έχει τη δυνατότητα τα κριτήρια να επηρεάζουν την ασάφεια των συνθηκών και των συμπερασμάτων. Το σύστημα συμπεραίνει χρησιμοποιώντας το συλλογισμό είτε της Ορθόδρομης είτε της Οπισθόδρομης Αλυσίδας. Χρησιμοποιώντας μια απλή τεχνική, η Μνήμη Εργασίας του συστήματος χωρίζεται σε Μνήμη Εργασίας Συνθηκών (Conditions' Working Memory) και Μνήμη Εργασίας Κανόνων (Rules' Working Memory) οι οποίες χρησιμοποιούνται σύμφωνα με τις ανάγκες του συστήματος. Ένα κρίσιμο σημείο είναι η χρήση της έγκυρης χρονικά γνώσης (Valid Time Knowledge), (Lorentzos et al, 1999) ένα πολύ χρήσιμο στοιχείο στις μεταβαλλόμενες Βάσεις Γνώσης για τις οποίες θέλουμε να παρακολουθήσουμε τις οποιεσδήποτε μεταβολές τους.

3.2.2 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ ΥΒΡΙΔΙΚΟΥ ΕΜΠΕΙΡΟΥ ΣΥΣΤΗΜΑΤΟΣ

Στο μοντέλο μας, η ΒΓ τηρείται σε σχεσιακή ΒΔ (Σχήμα 2), με τις κατάλληλες σχέσεις μεταξύ των πινάκων.

Αναλυτικότερα, οι συνθήκες 'Ισχυρισμός_Ι1, Ισχυρισμός_Ι2, ..., Ισχυρισμός_Ιn' του 'ΕΑΝ' μέρους του κανόνα, τηρούνται στον πίνακα 'Condition' και αναπαριστώνται με μορφή τριπλέτας 'Αντικείμενο – Ιδιότητα – Τιμή' (Α-Ι-Τ). Η 'Τιμή' (Value) κάθε συνθήκης μπορεί να είναι είτε μια αυτόνομη τιμή, είτε το συμπέρασμα ενός άλλου υφιστάμενου κανόνα. Τα συμπεράσματα 'Δήλωση_Δ1, Δήλωση_Δ2, ..., Δήλωση_Δn' του 'ΤΟΤΕ' μέρους του κανόνα, τηρούνται στον πίνακα 'Conclusion' και αυτά,

παρόμοια με τους ισχυρισμούς, μπορούν να είναι είτε μια αυτόνομη τιμή, είτε μία ενιαία συνθήκη, κωδικοποιημένη σε μορφή τριπλέτας 'A-I-T'. Είμαστε σε θέση να χρησιμοποιήσουμε μία ενιαία συνθήκη ως συμπέρασμα για να υποστηρίξουμε το Συμπερασματικό Μηχανισμό Ορθόδρομης Αλυσίδας. Επιπλέον, εάν υπάρχει ανάγκη για μια αντενέργεια ως συνεπαγωγή σε ένα συγκεκριμένο συμπέρασμα, επεκτείνουμε

Σχήμα 2. Αρχιτεκτονική του Υβριδικού Εμπειρού Συστήματος το μοντέλο του Παραγωγικού Κανόνα (εάν ... τότε ...), χρησιμοποιώντας τον πίνακα 'Reaction'.

Δεδομένου ότι ένας κανόνας αποτελείται από συνθήκη (ή συνθήκες), συμπέρασμα (ή συμπεράσματα) και ενδεχομένως την αντενέργεια (ή αντενέργειες), οι τρεις ανωτέρω πίνακες συσχετίζονται με τον πίνακα 'Rule', με ενδιάμεσους πίνακες με 'πολλοί σε πολλά' σχέσεις μεταξύ τους. Η εγκυρότητα του χρόνου (Validity of Time) εκτελείται σε επίπεδο κανόνα, ώστε χρησιμοποιώντας τα κατάλληλα πεδία, ένας κανόνας είναι έγκυρος για ένα συγκεκριμένο χρονικό διάστημα.

Υπάρχουν επίσης πεδία στους συσχετιζόμενους πίνακες για να αντιπροσωπεύσουν τα αναγκαία χαρακτηριστικά κάθε κανόνα. Πιο συγκεκριμένα, στο επίπεδο της συνθήκης, προκειμένου να εκφραστεί η αβεβαιότητα της γνώσης που περιλαμβάνεται στη Βάση Γνώσης (Yen, 1999) έχουμε συμπεριλάβει τον Παράγοντα Έλλειψης (Missing Factor), τον Παράγοντα Βαρύτητας (Weight Factor) και τον Παράγοντα Βεβαιότητας του Χρήστη (User's Certainty Factor) (Vouros, 2000), καθώς επίσης και την ένδειξη θέσης κάθε συνθήκης μέσα σε κάθε κανόνα. Στους πίνακες συμπεράσματος και αντενέργειας (Conclusion and Reaction tables), έχουμε συμπεριλάβει τον Παράγοντα Αυτοπεποίθησης (Confidence Factor) και την ένδειξη θέσης μέσα σε κάθε κανόνα, όπως στις συνθήκες. Αναφορικά με τον παράγοντα βεβαιότητας, αυτός είχε αρχικά χρησιμοποιηθεί στο ιατρικό Έμπειρο Σύστημα MYCIN, προκειμένου να υποστηριχθεί η αβεβαιότητα στην αιτιολόγηση. Πιο συγκεκριμένα, στο MYCIN, ο παράγοντας βεβαιότητας αποφασίζει το βαθμό πεποίθησης ενός κανόνα κατά τη διάρκεια της φάσης επαλήθευσής του και η τιμή του περιορίζεται στο διάστημα [-1, 1].

Το μοντέλο που έχει περιγραφεί μέχρι τώρα, είναι ένα απλό Έμπειρο Σύστημα βασιζόμενο σε κανόνες, το οποίο αναπαρίσταται σε σχεσιακή Βάση Δεδομένων. Σε πιο σύνθετη μορφή Εμπειρών Συστημάτων, οι τιμές που υπάρχουν στις συνθήκες, μπορεί να είναι ασαφείς. Παραδείγματος χάριν στη συνθήκη: 'η δεξαμενή έχει στάθμη ύδατος που είναι χαμηλή', η τιμή 'χαμηλή' μπορεί να αντιπροσωπευθεί είτε από ένα ασαφές σύνολο με διακριτές τιμές, είτε από μια Συνάρτηση Ιδιότητας Μέλους που καθορίζει ασαφώς την τιμή 'χαμηλή' για τη συγκεκριμένη που είναι μέρος ενός ορισμένου κανόνα.

Στο μοντέλο μας, δεν χρησιμοποιήσαμε το γενικό ασαφές σχεσιακό πρότυπο δεδομένων, αλλά οι τιμές κάθε τριπλέτας 'A-I-T' (συνθήκης) χρησιμοποιούνται ως χωριστά ασαφή σύνολα συσχετισμένα με μία ή περισσότερες συνθήκες. Επομένως, προκειμένου να αντιπροσωπευθεί η ασάφεια των τιμών, χρησιμοποιούμε ένα πίνακα για να αποθηκεύσουμε τα διακριτά ασαφή σύνολα των τιμών. Από τους δύο εναλλακτικούς τρόπους που χρησιμοποιούνται για να αντιπροσωπεύσουν μια Συνάρτηση Ιδιότητας Μέλους (συνεχής και διακριτή), έχουμε χρησιμοποιήσει τη διακριτή.

Το επόμενο βήμα που εισάγουμε, είναι ο τρόπος χειρισμού ασαφών τιμών κάτω από διάφορες καταστάσεις (κριτήρια). Στο προηγούμενο παράδειγμα που ορίσαμε τιμή 'χαμηλή', αυτή αφορούσε μία δεξαμενή, ενώ προφανώς η τιμή 'χαμηλή' είναι τελείως διαφορετική για ένα ποτήρι νερού. Προκειμένου να αντυπροσωπευθεί η μετατόπιση των ασαφών συνόλων κάτω από τα διάφορα κριτήρια, χρησιμοποιούμε ένα πίνακα για να αποθηκεύσουμε τα κριτήρια και έναν άλλο ενός για να τα συσχετίσουμε με τα διακριτά ασαφή σύνολα.

3.2.3 ΠΑΡΑΔΕΙΓΜΑΤΑ ΚΑΝΟΝΩΝ

Ως παράδειγμα θα χρησιμοποιηθεί η διερεύνηση του σχετικά απλού στη διερεύνηση συμπτώματος του αισθήματος καψίματος της γλώσσας (γλωσσικού καύσου).

Συγκεκριμένα, παρουσιάζεται ο τρόπος μερικής διερεύνησης του γλωσσικού καύσου μέσα από ένα σύνολο κανόνων, βασισμένων στην μεθοδολογία του ΥβρεΣ.

Οι κανόνες στην απλή τους μορφή είναι οι εξής:

1ος Κανόνας

ΕΑΝ υπάρχει αίσθημα καψίματος της γλώσσας ΚΑΙ Δεν υπάρχει αλλοίωση βλεννογόνου

ΤΟΤΕ Πιθανά αίτια: Υποκείμενη νόσος Ή Συστηματικοί βλαπτικοί παράγοντες

Ή Ψυχοσωματικοί παράγοντες Ή Κακοήθεια

ΑΝΤΙΔΡΑΣΗ: Παραπομπή Παθολόγο Ή Παραπομπή οδοντίατρο Ή Παραπομπή ψυχολόγο/ίατρο Ή Εξετάσεις στόματος.

2ος Κανόνας

ΕΑΝ η εξέταση υπερήχων στόματος είναι ενδεικτική κακοήθειας ΚΑΙ η βιοψία βλεννογόνου του στόματος έχει θετική ιστολογική για κακοήθεια

ΤΟΤΕ Πιθανά αίτια: CA

ΑΝΤΙΔΡΑΣΗ: Εξέταση με CT – MRI / Εγχείρηση / Χημειοθεραπεία / Ακτινοθεραπεία.

3ος Κανόνας (συμπληρωματικός)

ΕΑΝ η εξέταση υπερήχων στόματος έχει ευρήματα

ΤΟΤΕ είναι ενδεικτική κακοήθειας.

4ος Κανόνας (συμπληρωματικός)

ΕΑΝ η βιοψία βλεννογόνου στόματος έχει θετική ιστολογική

ΤΟΤΕ είναι ενδεικτική κακοήθειας.

Σημειώνεται ότι οι Κανόνες υπ' αριθμ. 3 και 4 είναι συμπληρωματικοί και χαρακτηρίζουν ένδειξη κακοήθειας αναλόγως των αποτελεσμάτων των εξετάσεων. Επίσης, οι τιμές των συνθηκών (ευρήματα, θετική ιστολογική) των εν λόγω κανόνων, ως γλωσσικές μεταβλητές (linguistic variables) εμπεριέχουν ασάφεια η οποία χειρίζεται από το προτεινόμενο μοντέλο μέσω του αντίστοιχου πίνακα (FuzzyValue) (Σχήμα 2).

Όλοι οι παραπάνω απλοί κανόνες απεικονίζονται στο σχεσιακό μοντέλο κανόνων του ΥβρεΣ ως εξής:

ΣΥΝΘΗΚΗ (μορφή A-I-T)

ΣΥΜΠΕΡΑΣΜΑ Πιθανά αίτια: Συστηματικοί βλαπτικοί παράγοντες Ή
ΕΑΝ η γλώσσα έχει αίσθημα που είναι κάψιμο Υποκείμενη νόσος Ή
ΚΑΙ η βλεννογόνος έχει αλλοίωση που είναι αρνητική
Ψυχοσωματικοί παράγοντες Ή
Κακοήθεια

ΕΑΝ η εξέταση υπερήχων Πιθανά αίτια: CA
στόματος έχει ένδειξη που
είναι κακοήθειας
ΚΑΙ η βιοψία βλεννογόνου
στόματος έχει ιστολογική
που είναι θετική για κακοή-
θεια

ΕΑΝ η εξέταση υπερήχων Ενδεικτική κακοήθειας
στόματος έχει ευρήματα

ΑΝΤΙΔΡΑΣΗ

Παραπομπή Παθολόγο
Παραπομπή οδοντίατρο
Παραπομπή
ψυχολό-
γο/ίατρο

Εξετάσεις στόματος
Εξέταση με CT – MRI

Εγχείρηση

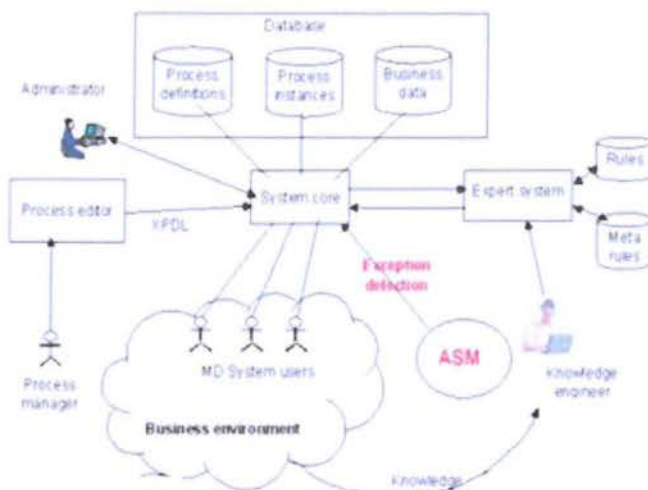
Χημειοθεραπεία

Ακτινοθεραπεία

που είναι θετικά

ΕΑΝ η βιοψία βλεννογόνου Ενδεικτική κακοήθειας
στόματος έχει ιστολογική που είναι θετική

Είναι προφανές ότι το παράδειγμα είναι υπεραπλουστευμένο και η απεικόνιση
των κανόνων στο ΥβρεΣ ιδιόρρυθμη για τον Ιατρό, αλλά αντανάκλα τη δυνατότητα
του συστήματος να χειριστεί προβλήματα διάγνωσης και συμβουλευτικής, ενώ
ταυτόχρονα καταδεικνύει τις δυσκολίες εκμείωσης και αποθήκευσης της Γνώσης με
τρόπο κοινά αποδεκτό.



2.2.4 ΣΥΜΠΕΡΑΣΜΑΤΑ

Σήμερα ένα περιορισμένο ποσοστό ιατρικής έρευνας επί των Εμπειρών Συστημάτων εκτελείται συνήθως από τους Ιατρούς ή τους Ακαδημαϊκούς. Επειδή τα ιατρικά Έμπειρα Συστήματα είναι μια πραγματική δύναμη και όχι μόνο μια ακαδημαϊκή αντίληψη (ως επίσης ούτε μια ακαδημαϊκή παραίτηση), πρέπει να σεβαστούμε την προσχώρηση της τεχνολογίας ΕΣ στην υγειονομική περίθαλψη γενικά. Πρέπει να εστιάσουμε στην υποστήριξη της διοικητικής διαδικασίας του ασθενή στην ολότητά της, παρά να εστιάσουμε μεμονωμένα στη διάγνωση. Περαιτέρω, οι εφαρμογές πρέπει να υποστηρίξουν ένα ευρύ φάσμα των κλινικών χρηστών, από τους Ιατρούς και τις Νοσηλεύτριες έως τους διαιτολόγους, τους φαρμακοποιούς, κ.λπ.

- Ο ρόλος των ιατρικών ΕΣ στην κλινική φροντίδα είναι πολύ σαφής. Υπάρχουν δε τομείς εφαρμογής που είναι κατάλληλοι για την άμεση εφαρμογή ή την περαιτέρω ανάπτυξη και την εκμετάλλευση. Αυτοί οι κλινικοί τομείς εφαρμογής περιλαμβάνουν:
- Εργαστηριακά συστήματα. Τα κλινικά εργαστήρια έχουν αποδειχθεί μια εύφορη περιοχή για τη χρήση των εμπειρων συστημάτων.
- Συμβουλευτικά συστήματα φαρμάκων. Υπάρχει δυνατότητα να σχεδιαστούν ΕΣ που θα βοηθήσουν τους νοσοκομειακούς γιατρούς στη συνταγογραφία και την επιλογή των οικονομικότερων θεραπειών.
- Ερμηνεία σημάτων. Η ανάπτυξη συστημάτων εγρήγορσης για αποστολή κλινικών σημάτων σε πραγματικό χρόνο σε περιοχές όπως η Μονάδα Εντατικής Θεραπείας, θα προσφέρει βοήθεια στην κλινική επαγρύπνηση.
- Εκπαίδευση. Η ανάγκη συνεχούς εκπαίδευσης για τους ασχολούμενους με την Υγεία, καθώς επίσης και για τους ασθενείς προσφέρεται μέσω της αυτοματοποιημένης βοήθειας. Το ευτυχές είναι ότι για τους περισσότερους νοσοκομειακούς γιατρούς, η εκπαίδευση είναι μια συνεχής επιδίωξη και διαδικασία. Τα ΕΣ μπορούν να υποστηρίξουν αυτή τη

διαδικασία.

Απ' ότι προκύπτει από την ανασκόπηση των υφισταμένων ιατρικών ΕΣ, αυτά έχουν επεκταθεί στους περισσότερους τομείς της Υγείας και συνεισφέρουν το καθένα στον τομέα του, ανάλογα με την ποσότητα και την ποιότητα γνώσης που έχουν αποθηκευμένη. Παρόλα ταύτα, δεν πρέπει να παραβλέψουμε τον αντίλογο, όπου σύμφωνα με τους Wyatt and Spiegelhalter (1990), μόνο 10% της πληθώρας των ιατρικών συστημάτων βασιζόμενων στη γνώση, έχουν δοκιμαστεί σε εργαστηριακές συνθήκες, ενώ ακόμα λιγότερα έχουν εκτεθεί σε κλινικές δοκιμές. Η αιτία αυτού του γεγονότος εντοπίζεται στη τη φύση της διαγνωστικής συλλογιστικής διαδικασίας στην ιατρική, η οποία σύμφωνα με τον Porle (1982) 'Το ερώτημα από τι είναι αυτό το οποίο διαχωρίζει την εμπειρογνωμοσύνη από έναν ταλαντούχο νοσοκομειακό γιατρό έχει ερευνηθεί

με διάφορες τεχνικές. Μια σχεδόν καθολική (παγκόσμια) διαπίστωση είναι ότι ο θεραπευτής απαντά στις ασάφειες των κλινικών δεδομένων συλλαμβάνοντας ένα ή περισσότερα διαγνωστικά σημεία τα οποία παίζουν σημαντικό ρόλο στην επακόλουθη απόφαση επιπρόσθετων δεδομένων και στην διάταξη επιλογών που λαμβάνονται υπόψη στην διεργασία τελική διαγνωστική απόφαση. Ένα διακριτικό γνώρισμα ενός ειδικού είναι η ικανότητα του να διαμορφώσει συγκεκριμένα κατάλληλα διαφορικές διαγνωστικές λειτουργίες στη βάση κάποιων δυσδιάκριτων ασαφειών στο φάκελο του ασθενούς. Το προτεινόμενο ΥβρΕΣ είναι κατάλληλο για τις περισσότερες από τις περιπτώσεις όπου άνθρωποι και υπολογιστές χρειάζεται να επικοινωνούν χρησιμοποιώντας

γλωσσικούς όρους προκειμένου να ληφθεί μια απόφαση. Εξαρτάται από το Μηχανικό Γνώσης να προσκτήσει τη γνώση από τον ειδικό εμπειρογνώμονα, να τη σχηματοποιήσει κατάλληλα και να την αποθηκεύσει στη ΒΓ ώστε να είναι εύκολα εκμεταλλεύσιμη.

4.ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΤΕΧΝΟΛΟΓΙΑΣ ΣΤΗΝ ΧΕΙΡΟΥΡΓΙΚΗ ΡΟΜΠΟΤΙΚΗ

Κατά την επέμβαση με τη χρήση ρομπότ, ο χειρουργός βρίσκεται μπροστά σε μια χειρουργική κονσόλα – Η-Υ, όπου βλέπει σε μια οθόνη το χειρουργικό πεδίο. Από εκεί δίνει εντολές, κινώντας ειδικούς μοχλούς, οι οποίες μεταφέρονται ψηφιακά ταυτόχρονα και με θαυμαστή ακρίβεια στους χειρουργικούς βραχίονες ενός ρομπότ, που εκτελεί τις κινήσεις μέσα στο χειρουργικό πεδίο. Η καινοτομία αυτή έχει πολλά πλεονεκτήματα τόσο για τον χειρουργό, όσο και για τον ασθενή, που την κάνουν προσιτή. Ένας ακόμα τρόπος με τον οποίο η τεχνολογία έχει συμβάλει στην πρόοδο της ιατρικής, είναι πως οι άνθρωποι που, εξαιτίας ατυχημάτων, πολέμου ή διαβήτη, έχουν χάσει μέλη του σώματός τους, μπορούν πλέον να χειρουργηθούν και να τους προστεθούν τεχνητά μέλη. Τα τεχνητά μέλη χωρίζονται σε προθέσεις κάτω και άνω άκρων και ειδικά πέλματα. Επίσης μια ειδική τεχνική που έχει αρχίσει να εφαρμόζεται, επιτρέπει στον ασθενή να κινεί το τεχνητό μέλος με τη σκέψη του, όπως και να νιώθει. Ο τρίτος τρόπος που μελετάμε είναι η εφεύρεση των μικροσκοπίων, τα οποία μας

επιτρέπουν την μεγέθυνση αντικειμένων που δεν μπορούμε παρατηρήσουμε με γυμνό μάτι. Όπως φαίνεται, η ιατρική έχει εξελιχθεί ιδιαίτερα με την είσοδο των ρομποτικών συστημάτων σε αυτή τον 21ο αιώνα. επιθυμούσε να προμηθευτεί ένα ποιο εξελιγμένο μοντέλο μικροσκοπίου από αυτό που έχει.

Η ρομποτική ιατρική αναπτύχθηκε όταν η NASA και ο στρατός ήθελαν να βρουν ένα τρόπο να χειρουργούνται οι αστροναύτες και οι στρατιώτες από γιατρούς χωρίς αυτοί να χρειάζεται να είναι μαζί με τους αστροναύτες η στο πεδίο της μάχης.

Κατά τη διάρκεια των τελευταίων 100 χρόνων, σημειώθηκαν οι περισσότερες ανακαλύψεις από κάθε άλλον αιώνα, στην ιστορία της ιατρικής.

Τον 21ου αιώνα στο χώρο της χειρουργικής εισάγεται η ρομποτική χειρουργική, δηλαδή ρομποτικών συστημάτων, καθοδηγούμενων από ηλεκτρονικούς υπολογιστές, που χειρίζονται από ανθρώπους.

Μερικά σημαντικά επιστημονικά επιτεύγματα της ιατρικής τεχνολογίας :

- ένας άρρωστος μπορεί να υποβληθεί ταυτόχρονα σε πολλαπλές επιτυχείς μεταμοσχεύσεις ζωτικών οργάνων
- βελτίωση των υπερήχων με την είσοδο της ενδοκοιλιακής και τρισδιάστατης υπερηχογραφίας,
- αξονική και μαγνητική τομογραφία,
- οι εφαρμογές των μικροχειρουργικών επεμβάσεων, που επιτρέπουν την επανασύνδεση ακόμα και των πλέον μικροσκοπικών αγγείων και νεύρων,
- οι τελειοποιήσεις των τεχνητών οργάνων,
- Λαπαροσκοπική χειρουργική: Ο χειρουργός κατά τη Λαπαροσκοπική Χειρουργική χειρίζεται τους ιστούς, όχι υπό άμεση όραση, αλλά μέσω μιας οθόνης και με τη βοήθεια ειδικών εργαλείων.
- Η Ρομποτική ιατρική μπορεί ένας χειρουργός να χειρουργήσει από απόσταση χιλιομέτρων.
- Εφευρέθηκε το ηλεκτρονικό μικροσκόπιο, το οποίο ψάχνει μόνο του τα νεοπλασματικά κύτταρα και να τα εντοπίζει.
- Στην αρθροπλαστική, ειδικά μηχανήματα αντιγράφουν με ακρίβεια την άρθρωση, ώστε το μόσχευμα να ταιριάζει απόλυτα.

4.1 ΜΗΧΑΝΗΜΑΤΑ ΡΟΜΠΟΤΙΚΗΣ ΧΕΙΡΟΥΡΓΙΚΗΣ

Η εκσυγχρονισμένη σημερινή τεχνολογία, μας επιτρέπει να την χρησιμοποιήσουμε για την υγεία του ανθρώπου. Έχει δημιουργηθεί μια ποικιλία από ρομποτικά μηχανήματα, τα οποία έχουν σκοπό την αποκατάσταση τραυμάτων και βλαβών του ανθρωπίνου οργανισμού. Πιο συγκεκριμένα, δημιουργήθηκαν συστήματα, που εξειδικεύονται στην χειρουργική. Τα ρομποτικά συστήματα που χρησιμοποιούνται στις χειρουργικές επεμβάσεις διαίρονται σε δύο κατηγορίες: τα παθητικά, τα οποία απαιτούν τον φυσικό χειρισμό από το χειρουργό, και τα ενεργά, στα οποία το

ρομποτικό σύστημα ελέγχεται από έναν Η/Υ και δεν απαιτεί ανθρώπινη παρέμβαση, όμως η παρουσία του χειρουργού είναι σημαντική. Τα πιο σημαντικά ρομποτικά συστήματα υποβοήθησης χειρουργικών επεμβάσεων τα οποία χρησιμοποιούνται διεθνώς είναι:

- τα ρομποτικά συστήματα ενδοσκόπησης,
- τα συστήματα ελέγχου και τηλεσυνεργασίας,
- τα συστήματα ρομποτικής ακτινοχειρουργικής
- τα ρομποτικά συστήματα τύπου master- slave



4.1.1 ΤΟ ΧΕΙΡΟΥΡΓΙΚΟ ΣΥΣΤΗΜΑ DAVINCI

Το χειρουργικό σύστημα daVinci είναι το πρώτο σύστημα ρομποτικής χειρουργικής που εγκρίθηκε από τον Αμερικανικό Οργανισμό Φαρμάκων και Υλικών (FDA) για την πραγματοποίηση επεμβάσεων. Ο σχεδιασμός του ξεκίνησε το 1995 και από το 2000 χρησιμοποιούνται παγκοσμίως περισσότερα από 400 ρομποτικά συστήματα daVinci εκ των οποίων πάνω από 80 στην Ευρώπη.

Αποτελείται από τρία αρχικά εξαρτήματα: την κονσόλα εξέτασης και ελέγχου του χειρουργού, το χειρουργικό «καροτσάκι», το οποίο στεγάζει τις ρομποτικές μονάδες βραχίονες στους οποίους τοποθετούνται και ελίσσονται τα αποσπώσιμα χειρουργικά όργανα και τον ενδοσκοπικό πύργο που περιλαμβάνει δύο video cameras, σύστημα αυτόματης ρύθμισης εικόνας, video monitor υψηλής ευκρίνειας, την κάμερα ελέγχου και τις μονάδες τηλεοπτικής επεξεργασίας.

Τα εργαλεία εισάγονται μέσα από τομές λίγων χιλιοστών. Ο χειρουργός κατευθύνει και συντονίζει το όλο σύστημα έχοντας στην κονσόλα τρισδιάστατη εικόνα κινεί με ειδικά χειριστήρια τους τέσσερις βραχίονες του ρομπότ. Κάθε βραχίονας είναι συνδεδεμένος με ένα χειρουργικό εργαλείο και στον κεντρικό βραχίονα υπάρχει κάμερα. Ο χειρουργός μπορεί να λυγίζει και να γυρίζει τους βραχίονες, όπως τον καρπό του ανθρώπινου χεριού. Οι καρποί των ρομποτικών εργαλείων μιμούνται τις κινήσεις που γίνονται από τον παθολόγο, ο οποίος κάθεται στην κονσόλα. Κάθε κίνηση αναπαράγεται με απόλυτη ακρίβεια και

σταθερότητα στο χειρουργικό πεδίο από τους χειρουργικούς βραχίονες του ρομπότ. Ο χειρουργός μέσω ειδικών φακών αντιλαμβάνεται το χειρουργικό πεδίο και συνομιλεί και συνεργάζεται με το ρομπότ και την υπόλοιπη χειρουργική ομάδα που είναι

παρούσα και ο ίδιος πρέπει να ειδικά εκπαιδευμένος στη χρήση του ρομποτικού συστήματος.



4.1.2 ΣΥΣΤΗΜΑΤΑ ΥΠΟΒΟΗΘΗΣΗΣ ΧΕΙΡΟΥΡΓΙΚΩΝ ΕΠΕΜΒΑΣΕΩΝ

Το PUMA 200, που έχει σκοπό να βοηθά το χειρουργό κατά τη διάρκεια επεμβάσεων νευροχειρουργικής και η κίνησή του επιτυγχάνεται χάρη στους DC σερβοκινητήρες.

Το Probot είναι σχεδιασμένο για την αφαίρεση του προστάτη. Η λειτουργία του επιτρέπει στο χειρουργό να εντοπίσει τον όγκο στον προστάτη και έπειτα να αφαιρέσει το συγκεκριμένο τμήμα χωρίς επιπλέον παρέμβαση χειρουργού.

Το NeuroMate αποτελείται από ένα, βασισμένο σε Η/Υ, σύστημα σχεδιασμού, καθώς και έναν ρομποτικό βραχίονα και επιτρέπει την απεικόνιση πολλαπλών τροχιών και ακριβή σχεδιασμό, βασισμένο σε εικόνες.

Το ROBODOC είναι το πρώτο ρομποτικό σύστημα που χρησιμοποιήθηκε σε χειρουργικές επεμβάσεις ορθοπαιδικής. Με την βοήθεια του, οι χειρουργοί σχεδιάζουν προεγχειρητικά, σε ένα τριών διαστάσεων εικονικό περιβάλλον, τις επεμβάσεις και κατόπιν τις εκτελούν στον ασθενή με τον τρόπο που σχεδιάστηκαν.

Το Minerva σχεδιάστηκε στο Πολυτεχνείο της Λωζάννης στην Ελβετία το 1991, με σκοπό την βοήθεια επεμβάσεων νευροχειρουργικής.

Το Acrobot είναι ένα ημιενεργό ρομποτικό σύστημα που έχει χρήση στην ορθοπαιδική χειρουργική.

Το neuroArm έχει εφαρμογές στη νευροχειρουργική και είναι ένα καθοδηγούμενο από εικόνα και ελεγχόμενο από Η/Υ, ρομποτικό σύστημα. Είναι ειδικά σχεδιασμένο τόσο για εφαρμογές βιοψίας και στερεοταξίας όσο και για μικροχειρουργικές επεμβάσεις.



4.1.3 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΓΙΑ ΤΟΝ ΑΣΘΕΝΗ ΚΑΙ ΤΟΝ ΧΕΙΡΟΥΡΓΟ

Τα πλεονεκτήματα που παρέχει η ρομποτική χειρουργική τόσο στον ασθενή είναι πολλά και σημαντικά. Καταρχάς μειώνει στο ελάχιστο την απώλεια αίματος από τον ασθενή κατά την διάρκεια της επέμβασης αλλά και του εξασφαλίζει λιγότερο πόνο. Επίσης η αναισθησία διαρκεί λιγότερο χρόνο ενώ η ανάρρωση είναι ταχύτερη. Ακόμα, το κόστος νοσηλείας είναι μικρότερο που είναι ιδιαίτερα σημαντικό στις δύσκολες μέρες που βιώνουμε. Περαιτέρω, τα αισθητικά αποτελέσματα αποδεικνύονται καλύτερα καθώς δεν φαίνονται. Όσον αφορά τα κέρδη που αποκομίζει ο χειρουργός, καταρχάς υπάρχει μεγαλύτερη ακρίβεια στις κινήσεις του και έτσι, μπορεί να πραγματοποιεί δύσκολες χειρουργικές επεμβάσεις με μεγαλύτερη σιγουριά. Επίσης, η χρήση αυτού του εξοπλισμού, του παρέχει μεγαλύτερη άνεση κατά την διάρκεια της επέμβασης καθώς έχει την δυνατότητα να πραγματοποιεί τις επεμβάσεις καθιστός. Επιπλέον έχει την δυνατότητα να προετοιμάσει την επέμβαση στον υπολογιστή χρησιμοποιώντας τις εικόνες των εσωτερικών οργάνων τις οποίες μπορεί κατά την διάρκεια της επέμβασης να ανακαλέσει στην οθόνη ανά πάσα στιγμή.

4.2 ΡΟΜΠΟΤΙΚΗ ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ

Τα παιδιά όταν σχεδιάζουν, κατασκευάζουν και προγραμματίζουν ρομπότ έχουν την ευκαιρία να μάθουν παίζοντας και να αναπτύξουν δεξιότητες. Η ρομποτική αφενός, είναι μία διασκεδαστική και ενδιαφέρουσα δραστηριότητα που δίνει τη δυνατότητα στο μαθητή να εμπλακεί με τη δράση, αφετέρου μπορεί να χρησιμοποιηθεί σε όλες τις βαθμίδες εκπαίδευσης για τη διδασκαλία διαφόρων εννοιών, κυρίως, από τις Φυσικές Επιστήμες και άλλα γνωστικά αντικείμενα.

- Φυσική (μελέτη της κίνησης, μελέτη της επίδρασης της τριβής, μελέτη της σχέσης των δυνάμεων, μεταφορά ενέργειας κ.α)
- Μαθηματικά και Γεωμετρία (αναλογίες, μέτρηση αποστάσεων, κατανόηση βασικών γεωμετρικών ιδιοτήτων όπως η περίμετρος κ.α)
- Μηχανική (κατασκευή, έλεγχος και αξιολόγηση μηχανικών λύσεων κ.α)
- Τεχνολογία (τεχνολογικός αλφαριθμητισμός κ.α)
- Ιστορία (πχ. με την κατασκευή ενός ρομπότ καταπέλτη - του Αρχιμήδη - τα παιδιά έχουν την ευκαιρία να γνωρίσουν την ανάπτυξη της τεχνολογίας εκείνης της εποχής καθώς και το έργο και την προσωπικότητα του Αρχιμήδη κ.α)
- Ο συνδυασμός εννοιών από διαφορετικές, γνωστικές περιοχές (τεχνολογία, τέχνη, περιβάλλον, κοινωνία, μαθηματικά, φυσικές επιστήμες) με διαθεματικά project (συνθετικές εργασίες)

Η εκπαιδευτική Ρομποτική έχει **θετικές επιπτώσεις** εκτός από το γνωστικό τομέα και **στο συναισθηματικό** (αυτοεκτίμηση, αυτοπεποίθηση) και **κοινωνικό** (κοινωνικοποίηση, απομυθοποίηση).

Επιπλέον, με τη βοήθεια της ρομποτικής στη διδασκαλία του ο εκπαιδευτικός μπορεί να επικεντρωθεί στην **ανάπτυξη και άλλων κρίσιμων δεξιοτήτων** του 21ου αιώνα:

- ομαδική εργασία
- επίλυση προβλημάτων (ανάλυση, σχεδίαση, υλοποίηση, δοκιμή και πειραματισμός, αξιολόγηση)
- καινοτομία
- διαχείριση έργου (διαχείριση χρόνου, κατανομή έργου και πόρων κ.α)
- προγραμματισμός
- δεξιότητες επικοινωνίας
- πολύτιμες νοητικές δεξιότητες (αναλυτική και συνθετική σκέψη, δημιουργικότητα, κριτική σκέψη κ.α)
κ.λπ.

Το όραμα της ρομποτικής είναι όλοι οι μαθητές να αναπτύξουν αυτές τις δεξιότητες, οι οποίες στα πλαίσια της παγκοσμιοποίησης αποτελούν επιτακτική ανάγκη για την προετοιμασία πολιτών του κόσμου που θα μπορούν να συνεισφέρουν θετικά σε παγκόσμια κλίμακα.

4.2.1 Η ΡΟΜΠΟΤΙΚΗ ΑΛΛΑΖΕΙ ΤΟΝ ΠΑΡΑΔΟΣΙΑΚΟ ΤΡΟΠΟ ΔΙΔΑΣΚΑΛΙΑΣ

- Η εκπαιδευτική ρομποτική συνδυάζει τη μάθηση με το παιχνίδι και έτσι μετατρέπει την εκπαίδευση σε μία διασκεδαστική δραστηριότητα - είναι γνωστό άλλωστε πως η μάθηση επιτυγχάνεται ευκολότερα, ταχύτερα και ουσιαστικότερα όταν συνδυάζεται με το παιχνίδι. Η πτυχή του παιχνιδιού που εμπεριέχουν τα προγραμματιζόμενα ρομπότ αποτελεί ένα σημαντικό παράγοντα θετικού κινήτρου και παρώθησης, κυρίως στην πρωτοβάθμια εκπαίδευση.
- Ευνοεί την ανάπτυξη ερευνητικού ενδιαφέροντος. Η εκπαιδευτική ρομποτική δίνει τη δυνατότητα στα παιδιά να δράσουν ως επιστήμονες - εφευρέτες και να ανακαλύψουν δικές τους καινοτόμες ιδέες και λύσεις.
- Εμπλέκει ενεργά τους μαθητές στη μάθησή τους με την επίλυση αυθεντικών προβλημάτων.
- Υποστηρίζει τη διερευνητική μάθηση και ενισχύει τη διερευνητική στάση των μαθητών.
- Δίνει κίνητρα στους μαθητές να μελετήσουν την επιστήμη και την τεχνολογία.
- Η ρομποτική, εμπλέκει τους μαθητές σε καταστάσεις που απαιτούν από αυτούς να εφαρμόσουν τα μαθηματικά και την επιστήμη και όχι απλά να τα μελετήσουν. Διότι η κατανόηση είναι κάτι περισσότερο από μάθηση, είναι μάθηση και γνώση πώς να εφαρμόσεις αυτό που γνωρίζεις στο πλαίσιο αυτό.
- Η Εκπαιδευτική ρομποτική παρέχει στους μαθητές ευκαιρίες επίλυσης προβλημάτων με προσωπικό νόημα για τους ίδιους μέσω χειρισμού και

κατασκευής πραγματικών ή ιδεατών αντικειμένων. Η γνώση που προκύπτει από προβληματικές καταστάσεις δίνει την ευκαιρία στους μαθητές να αναπτύξουν μία ισχυρή εννοιολογική βάση για την ανακατασκευή των γνώσεών τους σε μεταγενέστερο χρόνο.

- Επιτρέπει την ελεύθερη έκφραση και την ανάπτυξη της δημιουργικότητας και φαντασίας.
- Μέσα από την κατασκευή θέτει πραγματικά προβλήματα και παρέχει άμεση ανατροφοδότηση.
- Επιτρέπει την πρόσκτηση γνώσεων και δεξιοτήτων που συνδέονται με πολλά γνωστικά αντικείμενα (και συνεπώς την προώθηση της διεπιστημονικής και διαθεματικής προσέγγισης).
- Δίνει τη δυνατότητα για πιθανή διαισθητική συνειδητοποίηση σύνθετων φαινομένων, όπως η σχέση ανάμεσα στην ταχύτητα, το χρόνο και τη μετακίνηση.
- Μπορεί ακόμα να αξιοποιηθεί σαν «σκαλωσιά» για την κατάκτηση συνθετότερων εννοιών όπως των δυναμικών συστημάτων της Βιολογίας κλπ.
- Στα πλαίσια διαθεματικών εκπαιδευτικών δραστηριοτήτων, μέσω αυτής, είναι δυνατόν να αναπτυχθούν κίνητρα για μάθηση και σε άλλα μαθήματα (Μαθηματικά, Φυσική κ.λπ).
- Στηρίζεται στη συνεργασία και στην αλληλεπίδραση ατόμων κι ομάδων και στην προώθηση της σκέψης μέσω γνωστικών και κοινωνικογνωστικών συγκρούσεων.
- Καλλιεργείται η επικοινωνία και η λεκτική έκφραση ιδεών μέσα από τη λειτουργία της ομάδας όταν οι μαθητές αναγκάζονται να εξηγήσουν ιδέες και σκέψεις τους.
- Υποστηρίζει τη βιωματική μάθηση.
- Ως παιδαγωγική προσέγγιση εγγράφεται στο πλαίσιο του κλασικού εποικοδομιστικού (constructivism) (Piaget, 1972) και ειδικότερα του κατασκευαστικού εποικοδομιστικού (constructionism) (Papert, 1993). Η εποικοδομητική (constructivist) αντίληψη για το φαινόμενο της μάθησης υποστηρίζει ότι το μαθησιακό περιβάλλον θα πρέπει να παρέχει αυθεντικές δραστηριότητες ενταγμένες σε διαδικασίες επίλυσης ανοιχτών προβλημάτων από τον πραγματικό κόσμο, να ενθαρρύνει την έκφραση και την προσωπική εμπλοκή ή στη μαθησιακή διαδικασία και να υποστηρίζει την κοινωνική αλληλεπίδραση. Επιπλέον ο “κατασκευαστικός” εποικοδομητισμός (constructionism) υποστηρίζει ότι οι μαθητευόμενοι οικοδομούν πιο αποτελεσματικά τη γνώση όταν εμπλέκονται ενεργά στη σχεδίαση και κατασκευή (χειρωνακτική και ψηφιακή) πραγματικών αντικειμένων που έχουν νόημα για τους ίδιους είτε αυτά είναι

- κάστρα από άμμο, είτε κατασκευές LEGO και προγράμματα υπολογιστών (Papert, 1991).
- Η μάθηση διαδικασιών ανάλυσης, σχεδιασμού δράσεων και στη συνέχεια η υλοποίησή τους (μέσω μιας μηχανής) συνιστά μια νοητική δεξιότητα υψηλού επιπέδου - που εντάσσεται στη μεγάλη κατηγορία έργων που οι ψυχολόγοι ονομάζουν επίλυση προβλημάτων.
 - Διευκολύνει την εκμάθηση του προγραμματισμού
 - Ο προγραμματισμός ρομποτικών κατασκευών δημιουργεί ένα εντελώς νέο περιβάλλον εργασίας για τους μαθητές με τα εξής χαρακτηριστικά:
 - (α) Είναι έντονα παρακινητικό, και συνεπώς παράγοντας υψίστης σημασίας για τη διδακτική
 - (β) Ο προγραμματισμός της συμπεριφοράς των ρομπότ προκύπτει από μεταφορά υπαρχόντων και ήδη γνωστών συμπεριφορών από τους ζώντες οργανισμούς
 - (γ) Ευνοεί τη στρατηγική δοκιμής – πλάνης, που είναι στρατηγική οικεία στους μαθητές του δημοτικού
 - (δ) Αναδεικνύει παραδεκτές προσεγγίσεις και λύσεις και όχι μία και μοναδική σωστή λύση αφού μία συμπεριφορά μπορεί να αποδοθεί με πολλούς τρόπους
 - (ε) Υποστηρίζει μεταγνωστικές διεργασίες μάθησης, δεδομένου ότι η προγραμματιστική δραστηριότητα οδηγεί στη συγκρότηση, την ανάλυση και την εξωτερικευση νοητικών διεργασιών. Αυτή η προσπάθεια έχει μεταγνωστικό χαρακτήρα αφού μας αναγκάζει να σκεφτόμαστε πάνω στον τρόπο που σκεφτόμαστε και ενεργούμε

4.2.2 ΜΕΙΟΝΕΚΤΗΜΑΤΑ

Η χρήση φυσικών

μοντέλων κατά τη διδασκαλία των αρχών του προγραμματισμού μπορεί να δημιουργήσει διδακτικά προβλήματα, τα σημαντικότερα από τα οποία οφείλονται σε φυσικούς περιορισμούς και τεχνικούς περιορισμούς του υλικού - περιορισμοί που μπορεί να δημιουργήσει η τριβή στην κίνηση ενός αντικειμένου κ.α, χρονικούς περιορισμούς (χρόνος εκμάθησης του συστήματος) και περιορισμούς κόστους επένδυσης (μεγάλη οικονομική επιβάρυνση για την απόκτηση και συντήρησή - επικαιροποίηση απαιτούμενου αριθμού συστημάτων).



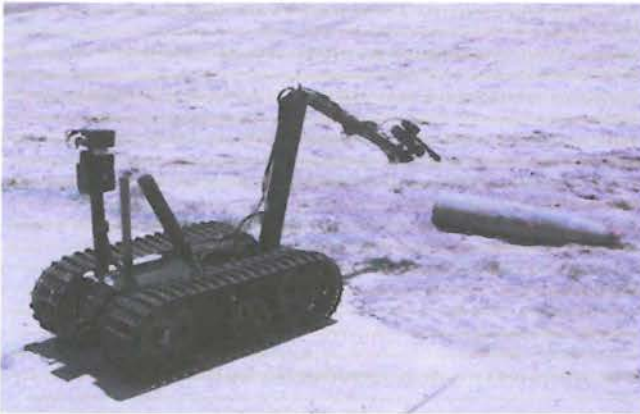
4.3 Η ΡΟΜΠΟΤΙΚΗ ΣΤΟΝ ΣΤΡΑΤΟ

Οι στρατιώτες ρομπότ λειτουργούν αυτόνομα ή με συσκευές τηλεχειρισμού που είναι ειδικά σχεδιασμένες για στρατιωτικές εφαρμογές. Αυτά τα συστήματα αυτού του αντικειμένου ερευνώνται από διάφορους στρατούς. Τα στρατιωτικά ρομπότ πρωτοεμφανίστηκαν στον Δεύτερο Παγκόσμιο και στον Ψυχρό Πόλεμο, με τη μορφή της γερμανικής Goliath. Το MQ-1 Predator drone ήταν όταν οι αξιωματικοί της CIA άρχισαν να βλέπουν την πρώτη πρακτική επιστροφή της δεκαετίας τους ώστε να χρησιμοποιούν τα εναέρια-ρομπότ. Η χρήση των ρομπότ σε πολεμικές επιχειρήσεις, αν και παραδοσιακά θυμίζει σενάριο από ταινία επιστημονικής φαντασίας, διερευνάται ως πιθανό μελλοντικό μέσο διεξαγωγής πολέμων. Ήδη αρκετοί στρατιώτες ρομπότ έχουν αναπτυχθεί από διάφορους στρατούς. Μερικοί πιστεύουν ότι το μέλλον του σύγχρονου πολέμου θα είναι με αυτοματοποιημένα συστήματα όπλων. Ο αμερικανικός στρατός επενδύει σημαντικά στην έρευνα και στην ανάπτυξη προς δοκιμές και στην ανάπτυξη όλο ένα και πιο αυτοματοποιημένων συστημάτων. Το πιο σημαντικό σύστημα που χρησιμοποιείται σήμερα είναι το μη επανδρωμένο εναέριο όχημα (IAI Pioneer & RQ-1 Predator), το οποίο μπορεί να οπλιστεί με βλήματα που πέφτουν από τον αέρα και καταλήγουν στο έδαφος και λειτουργεί με τηλεχειρισμό από ένα κέντρο διοίκησης στην αναγνώριση των ρόλων. Η DARPA έχει φιλοξενήσει διαγωνισμούς το 2004 & 2005 για την εμπλοκή ιδιωτικών εταιρειών και πανεπιστημίων ώστε να αναπτύξουν μη επανδρωμένα οχήματα εδάφους για να πλοηγηθούν σε ανώμαλο έδαφος στην έρημο Mojave για ένα τελικό έπαθλο των 2 εκατομμυρίων δολαρίων. Το

πεδίο του πυροβολικού έχει δει επίσης κάποια πολλά υποσχόμενη έρευνα με ένα πειραματικό σύστημα όπλων που ονομάζεται "Dragon Fire II", το οποίο αυτοματοποιεί τους υπολογισμούς και την επαναφόρτωση βαλλιστικών όπλων που απαιτούνται για ακριβείς πρόβλεψη πυροβολισμού, έχοντας απόκριση 12 δευτερόλεπτα του χρόνου σε αιτήματα υποστήριξης πυροβολικού. Ωστόσο, τα όπλα του πολέμου έχουν ένα όριο για να γίνουν πλήρως αυτόνομα: εξακολουθούν να υπάρχουν σημεία στα οποία απαιτείται επέμβαση του ανθρώπου για την εξασφάλιση ότι οι στόχοι δεν είναι μέσα σε ζώνες περιορισμένης πρόσβασης πυροβολισμού, όπως ορίζεται από της Συμβάσεις της Γενεύης για τον νόμο του πολέμου. Έχουν γίνει κάποιες εξελίξεις προς την κατεύθυνση της ανάπτυξης αυτόνομων μαχητικών και βομβαρδιστικών αεροπλάνων. Η χρήση των αυτόνομων μαχητικών και βομβαρδιστικών είναι για την καταστροφή των στόχων του εχθρού και είναι ιδιαίτερα ελπιδοφόρα. Λόγω της έλλειψης εκπαίδευσης που απαιτείται για τους πιλότους των ρομποτικών, τα αυτόνομα αεροπλάνα είναι σε θέση να εκτελέσουν τους ελιγμούς που δεν θα μπορούσαν διαφορετικά με τους ανθρώπινους πιλότους (λόγω του υψηλού ποσού της G-Force), τα αεροπλάνα είναι σχεδιασμένα να μην απαιτούν ένα σύστημα υποστήριξης της ζωής, και μια απώλεια ενός αεροπλάνου δεν σημαίνει την απώλεια ενός πιλότου. Ωστόσο, το μεγαλύτερο μειονέκτημα για την ρομποτική είναι η αδυναμία τους να μην μπορούν να φιλοξενήσουν μη κανονικές συνθήκες. Οι εξελίξεις στην τεχνητή νοημοσύνη στο εγγύς μέλλον μπορεί να βοηθήσει για την επανόρθωση αυτής. Η US Mechatronics έχει δημιουργήσει ένα όπλο φρουρό το οποίο λειτουργεί αυτόματα και και τώρα αναπτύσσεται κι' άλλο για εμπορική και στρατιωτική χρήση. Το MIDARS, είναι ένα τετράτροχο ρομπότ εξοπλισμένο με πολλές κάμερες, ραντάρ, και, ενδεχομένως, ένα πυροβόλο όπλο, που εκτελεί αυτόματα τυχαίες ή προγραμματισμένες περιπολίες γύρω από μια στρατιωτική βάση ή άλλη εγκατάσταση της κυβέρνησης. Ειδοποιεί έναν ανθρώπινο παρατηρητή όταν ανιχνεύει μια κίνηση σε μία μη εξουσιοδοτημένη περιοχή. Ο χειριστής μπορεί να δώσει εντολή στη συνέχεια το ρομπότ να αγνοήσει το γεγονός, ή να πάρει ο ίδιος το τηλεχειριστήριο και να ασχοληθεί με τον εισβολέα, ή καλύτερα να κοιτάξει την κάμερα έκτακτης ανάγκης. Το ρομπότ σαρώνει τακτικά τις ετικέτες αναγνώρισης ραδιοσυχνότητας (RFID) που διατίθενται στην αποθήκη αποθεμάτων καθώς περνάει από εξώ και αναφέρει τυχόν στοιχεία που λείπουν. Autonomous Rotorcraft Sniper System είναι ένα πειραματικό ρομποτικό σύστημα όπλων που αναπτύσσεται από τον αμερικανικό στρατό από το 2005. Αποτελείται από ένα τηλεχειριζόμενο όπλο ελεύθερου σκοπευτή που συνδέεται με ένα μη επανδρωμένο αυτόνομο ελικόπτερο. Προορίζεται για χρήση σε αστική μάχη ή για διάφορες άλλες αποστολές που απαιτούν ελεύθερους σκοπευτές. Οι δοκιμές της πτήσης έχουν προγραμματιστεί να ξεκινήσουν το καλοκαίρι του 2009. Το «Mobile Autonomous Robot Software» είναι ένα ερευνητικό πρόγραμμα το οποίο ξεκίνησε το

Δεκέμβριο του 2003 από το Πεντάγωνο οι οποίοι αγόρασαν 15 Segways σε μια προσπάθεια για την ανάπτυξη πιο προηγμένων στρατιωτικών ρομπότ. Το πρόγραμμα ήταν μέρος ενός προγράμματος του πενταγώνου αξίας 26 εκατομμυρίων δολαρίων και ήταν για την ανάπτυξη λογισμικού για αυτόνομα συστήματα. Ο Ταγματάρχης Kenneth Rose ο οποίος είναι εκπαιδευτής του Στρατού των ΗΠΑ και της Doctrine Command περιγράφει μερικά από τα πλεονεκτήματα της ρομποτικής τεχνολογίας στον πόλεμο: «Οι μηχανές δεν κουράζονται, δεν κλείνουν τα μάτια τους, δεν κρύβονται κάτω από τα δέντρα, όταν βρέχει και δεν μιλάνε με τους φίλους τους. Μία ανθρώπινη προσοχή στη λεπτομέρεια την ώρα της σκοπιάς μειώνεται δραματικά στα πρώτα 30 λεπτά. Οι μηχανές δεν ξέρουν πως να φοβούνται. "Επίσης δίνεται μεγάλη και αυξανόμενη προσοχή στο πως να γίνουν τα ρομπότ πιο αυτόνομα και στο τέλος να επιτρέποντάς τα να λειτουργούν από μόνα τους για μεγάλα χρονικά διαστήματα, πιθανώς πίσω από τις εχθρικές γραμμές. Για αυτές τις λειτουργίες, τα συστήματα όπως το Energetically Autonomous Tactical Robot δοκιμάζονται, με στόχο να αποκτήσουν τη δική του ενέργεια από την αναζήτηση τροφής για φυτική ύλη.

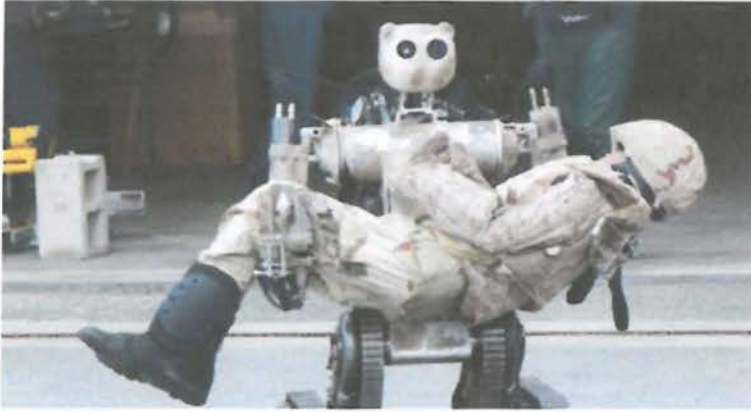
Το 2009, επιστήμονες και τεχνικοί εμπειρογνώμονες συμμετείχαν σε διάσκεψη για να συζητήσουν τις επιπτώσεις του υποθετικού ενδεχόμενου τα ρομπότ και οι υπολογιστές να μπορούσαν να γίνουν αυτόνομα και να είναι σε θέση να παίρνουν τις δικές τους αποφάσεις. Συζητήσαν τη δυνατότητα και το βαθμό στον οποίο οι υπολογιστές και τα ρομπότ μπορούν να είναι σε θέση να αποκτήσουν οποιοδήποτε επίπεδο αυτονομίας, και σε ποιο βαθμό θα μπορούσαν να χρησιμοποιήσουν τέτοιες ικανότητες ώστε να θέσει οποιαδήποτε απειλή ή κίνδυνο. Σημείωσαν ότι κάποια ρομπότ έχουν αποκτήσει διάφορες μορφές όπως η ήμι-αυτονομία, συμπεριλαμβανομένης της ικανότητας να είναι σε θέση να βρίσκουν τις πηγές ενέργειας από μόνα τους και να είναι σε θέση να επιλέγουν ανεξάρτητα στόχους για την επίθεσή τους με όπλα. Σημείωσαν επίσης ότι ορισμένοι ιοί υπολογιστών μπορούν να αποφύγουν την κατάργησή και έχει την επιτύχει «cockroach intelligence». Σημείωσαν ότι η αυτογνωσία, όπως απεικονίζεται σε ταινία επιστημονικής φαντασίας είναι μάλλον απίθανο, αλλά επίσης υπήρχαν και άλλοι πιθανοί κίνδυνοι και παγίδες. Μερικοί ειδικοί και ακαδημαϊκοί έχουν αμφισβητήσει τη χρήση των ρομπότ για την καταπολέμηση των στρατιωτικών, ιδίως όταν σε τέτοια ρομπότ έχει δοθεί κάποιος βαθμός αυτόνομης λειτουργίας. Το Πολεμικό Ναυτικό των ΗΠΑ έχει χρηματοδοτήσει μια έκθεση που δείχνει ότι τα στρατιωτικά ρομπότ γίνονται πιο σύνθετα, πρέπει να υπάρξει μεγαλύτερη προσοχή στις επιπτώσεις των ικανοτήτων τους να παίρνουν αυτόνομες αποφάσεις.



Μία αμερικανική στρατιωτική υπηρεσία (**DARPA**), που έχει και στο παρελθόν επικεντρωθεί στην οδήγηση οχημάτων από ρομπότ, έχει στρέψει την προσοχή της σε ανθρωποειδή ρομπότ που θα μπορούσαν να οδηγούν αυτόματα στο πεδίο μάχης. Με ανακοίνωσή της προτείνει στο στρατό να βάλει στη θέση του οδηγού ρομπότ που μπορεί να κατευθύνει ένα όχημα, να ξεκλειδώνει μια πόρτα, να μπορεί να ανέβει μια σκάλα και να εκτελεί πολύπλοκες επισκευές. Τα ρομπότ αυτά θα χρησιμοποιούν το δικό τους «μυαλό» στην εκτέλεση των εργασιών και μία χαλαρή επίβλεψη από τον άνθρωπο είναι αρκετή. Τα στοιχεία αυτά προέκυψαν από την ομιλία του Gill Pratt σε συνέδριο της **DARPA** (U.S. Defense Advanced Research Projects Agency) για ένα νέο πολλά υποσχόμενο ρομπότ. Η DARPA είχε και στο παρελθόν επικεντρωθεί στα αυτοκίνητα - ρομπότ. Ωστόσο, η στρατιωτική υπηρεσία έχει δείξει έντονο ενδιαφέρον για ρομπότ με πόδια και χρηματοδοτεί προγράμματα, όπως, για παράδειγμα, τετράποδο ρομπότ -μουλάρι «**Boston Dynamics**» το οποίο έχει σχεδιαστεί για να φέρει τα όπλα μάχης Αμερικανών πεζοναυτών. Τώρα πλέον, η DARPA φαίνεται ότι είναι έτοιμη να ξεκινήσει ένα νέο «**Grand Challenge**» ρομπότ, που ωθεί δηλαδή τις ρομποτικές ικανότητες σε νέα επίπεδα. Το ρομπότ πρέπει να έχει το συντονισμό για να οδηγεί όπως ένα... τρακτέρ.

Τέτοιου είδους ρομποτικές κατασκευές θα μπορούσαν να φανούν χρήσιμα και σε μελλοντικές πυρηνικές καταστροφές, όπως στην περίπτωση της Φουκουσίμα της Ιαπωνίας.

Οι προδιαγραφές του νέου ανδροειδούς «ταιριάζει» και στο **Ναυτικό των ΗΠΑ** ως ρομπότ - πυροσβέστης που θα υπακούει σε ανθρώπινες φωνές και εντολές κίνησης.



Πειράματα σε αυτόν τον τομέα πραγματοποιούνταν και στην ΕΣΣΔ ως παράδειγμα, μπορούμε να αναφέρουμε τα τηλεκατευθυνόμενα μέσω ραδιοσυχνότητων άρματα μάχης, προορισμένα για την ανίχνευση ναρκοπεδίων, ως φλογοβόλα, για την εγκατάσταση προπετάσματος καπνού. Όμως, σε τεχνολογικό επίπεδο ήταν δύσκολο να γίνει αξιόπιστο ένα τέτοιο σύστημα. Με την ανάπτυξη των τεχνικών μέσων και της τεχνολογίας ο αριθμός των τηλεκατευθυνόμενων μέσων αυξανόταν, μέχρι που στο τέλος του προηγούμενου αιώνα πέρασαν στην εξειδικευμένη ποιότητα τα ρομπότ του στρατού ξηράς, της αεροπορίας και του ναυτικού όλων των ειδών διαδραματίζουν ολοένα και μεγαλύτερο ρόλο στη διεξαγωγή πολεμικών επιχειρήσεων.

Τα μη επανδρωμένα αεροσκάφη χρησιμοποιούνται για κατασκοπεία, και υπόδειξη στόχων, για καθοδήγηση κατευθυνόμενων πυραύλων. Τα επίγεια ρομπότ χρησιμοποιούνται για αποναρκοθέτηση, για άρση εμποδίων· πολλά από αυτά είναι οπλισμένα και μπορούν να χρησιμοποιηθούν σε στρατιωτικές επιχειρήσεις και σε συνθήκες υψηλού κινδύνου, για παράδειγμα, εντός πόλης.

Τα ρομπότ αρχίζουν να χρησιμοποιούνται και στον πολιτικό τομέα έτσι, η εταιρεία Oshkosh Truck αναπτύσσει «μη επανδρωμένα» φορτηγά, ενώ η εταιρεία Boston Dynamics έχει δημιουργήσει ένα ρομπότ-αχθοφόρο, με μορφή που θυμίζει μεγάλο σκύλο (μάλιστα το έχουν ονομάσει ως Big Dog), ικανό να μεταφέρει φορτίο μέχρι 75 κιλά.



4.4 ΡΟΜΠΟΤΙΚΗ ΣΤΟ ΔΙΑΣΤΗΜΑ

Μία από τις σημαντικότερες εφαρμογές της ρομποτικής επιστήμης είναι στον τομέα της εξερεύνησης και μελέτης του διαστήματος. Τα ρομπότ, βλέπετε, είναι πιο ανθεκτικά από ό,τι ο άνθρωπος στις δύσκολες συνθήκες του διαστήματος, και έτσι μπορούν να ταξιδέψουν στα πιο απομακρυσμένα μέρη του γαλαξία μας!

Η συμβολή των ρομπότ στην εξερεύνηση του σύμπαντος είναι τεράστια. Ακριβώς επειδή τα ρομπότ δεν είναι ζωντανοί οργανισμοί, αλλά μηχανήματα, οι επιστήμονες μπορούν να τα σχεδιάσουν έτσι, ώστε να μπορούν να λειτουργήσουν σε αντίξοες συνθήκες. Για παράδειγμα, μπορούν να κάνουν τις εργασίες για τις οποίες έχουν προγραμματιστεί ακόμα και σε συνθήκες πολύ υψηλών ή πολύ χαμηλών θερμοκρασιών.

Ένας επιπλέον παράγοντας που οι επιστήμονες του διαστήματος χρησιμοποιούν τα ρομπότ, είναι ότι πολλές από τις αποστολές που οργανώνουν διαρκούν για ολόκληρα χρόνια! Έτσι καταλαβαίνετε, ότι ακόμα κι αν υπήρχε κάποιος τόσο παθιασμένος αστροναύτης που να άντεχε για χρόνια μακριά από τη Γη, σίγουρα δεν θα χωρούσαν στο διαστημόπλοιο αρκετά τρόφιμα ή αρκετό οξυγόνο για να επιβιώσει!

Όπως καταλαβαίνετε, η δημιουργία ενός ρομπότ είναι μία διαδικασία που... κοστίζει αρκετά χρήματα! Κι όμως, το να στείλει κανείς ένα ρομπότ στο διάστημα είναι πολύ πιο οικονομικό από το να στείλει έναν άνθρωπο. Κι αυτό γιατί τα διαστημόπλοια που θα το μεταφέρουν δεν χρειάζεται να είναι τόσο μεγάλα. Παράλληλα, δεν χρειάζεται να είναι εξοπλισμένα με τροφοδοσία οξυγόνου και όλων των άλλων απαραίτητων στοιχείων για την επιβίωση ενός ανθρώπου, όπως είναι το φαγητό ή οι εγκαταστάσεις υγιεινής.

Υπάρχει όμως κι άλλος ένας λόγος που προτιμούμε να στέλνουμε ρομπότ στο διάστημα και ίσως είναι ο σημαντικότερος: τα επανδρωμένα διαστημόπλοια (αυτά δηλαδή που μεταφέρουν ανθρώπους) πρέπει να είναι απολύτως ασφαλή, ώστε να μην κινδυνεύσουν οι επιβάτες τους. Αυτά τα συστήματα ασφαλείας είναι εξαιρετικά ακριβά, ενώ υπάρχει πάντα και ο κίνδυνος να μη λειτουργήσουν σωστά. Χάρη στα

ρομπότ λοιπόν, μπορούμε να γνωρίσουμε το διάστημα χωρίς να κινδυνεύσουν ανθρώπινες ζωές!

Διάσημες αποστολές στις οποίες χρησιμοποιήθηκαν ρομπότ:

Βόγιατζερ (1977 – σήμερα)

Διαστημικό λεωφορείο Endeavour (2010)

Discovery (σειρά αποστολών από το 1984 έως το 2011)



5 ΕΥΦΥΕΙΣ ΕΙΚΟΝΙΚΑ ΠΕΡΙΒΑΛΛΟΝΤΑ

Τα εικονικά περιβάλλοντα είναι μια διαρκώς αναπτυσσόμενη περιοχή με πολλές και ενδιαφέρουσες εφαρμογές σε διάφορους τομείς, όπως η εκπαίδευση η διασκέδαση και η προσομοίωση. Ένα σημαντικό τμήμα των εικονικών περιβαλλόντων που ενισχύει την αυτονομία και την αλληλεπιδραστικότητά τους είναι οι εικονικοί πράκτορες, δηλαδή οι αυτόνομες οντότητες του περιβάλλοντος που είναι σε θέση να αντιλαμβάνονται τον κόσμο γύρω τους και να εκτελούν ενέργειες ώστε να επιτύχουν τους στόχους τους.

Οι εικονικοί πράκτορες χωρίζονται σε φυσικούς και γνωσιακούς. Στους πρώτους δίνεται έμφαση στην κίνηση και αλληλεπίδρασή τους με τα αντικείμενα του περιβάλλοντος, ενώ στους δεύτερους σε περισσότερο νοητικές ικανότητες, όπως η δυνατότητα μάθησης, επικοινωνίας, εξαγωγής συμπερασμάτων, κλπ. Κατά συνέπεια, οι φυσικοί πράκτορες αξιοποιούν κυρίως την έρευνα στους χώρους της ρομποτικής,

συνθετικής κίνησης και μοντελοποίησης φυσικών νόμων, ενώ οι γνωσιακοί στηρίζονται σε αρχιτεκτονικές ευφυών πρακτόρων από τον χώρο της τεχνητής νοημοσύνης.

Ένα βασικό μειονέκτημα της τρέχουσας έρευνας στους εικονικούς πράκτορες είναι η έλλειψη γενικών εργαλείων ανάπτυξης τέτοιων οντοτήτων, καθώς οι υπάρχουσες μηχανές τρισδιάστατων γραφικών φαίνεται να παρουσιάζουν σημαντικές ελλείψεις σε ό,τι αφορά την λειτουργικότητα των πρακτόρων. Το πρόβλημα αυτό εντείνεται ακόμη περισσότερο με την απουσία κοινών μεθοδολογιών και αρχιτεκτονικών για τον σχεδιασμό και την ανάπτυξη εικονικών πρακτόρων, γεγονός που καθιστά δύσκολη την σύγκριση μεταξύ διαφορετικών προσεγγίσεων, αλλά και την άμεση επαναχρησιμοποίηση τμημάτων κώδικα της υπάρχουσας έρευνας. Τέλος, ένα ακόμη θεμελιώδες πρόβλημα που εμφανίζεται στην περιοχή των εικονικών πρακτόρων είναι η δυσκολία διασύνδεσης μεταξύ φυσικών και γνωσιακών ικανοτήτων ενός εικονικού πράκτορα, λόγω του διαφορετικού επιπέδου αφαίρεσης στον τρόπο αναπαράστασης της γνώσης στις δύο αυτές περιοχές.

Τα παραπάνω προβλήματα αποτελούν το αντικείμενο μελέτης της τρέχουσας διατριβής, που έχει ως θέμα την εισαγωγή ευφυών πρακτόρων σε εικονικά περιβάλλοντα, δηλαδή την δυνατότητα σχεδιασμού και ανάπτυξης τρισδιάστατων συνθετικών χαρακτήρων με ενισχυμένα γνωσιακά χαρακτηριστικά. Πιο συγκεκριμένα:

- Προτείνεται ένα γενικό μοντέλο για ευφυείς εικονικούς πράκτορες το οποίο διατηρεί τους κοινώς αποδεκτούς στους χώρους των γραφικών και της εικονικής πραγματικότητας τρόπους αναπαράστασης του κόσμου και εκτέλεσης ενεργειών, αλλά αξιοποιεί ταυτόχρονα και την υπάρχουσα έρευνα στους ευφυείς πράκτορες, καθώς στηρίζεται στην αρχιτεκτονική Αίσθησης - Ελέγχου - Ενέργειας (Sense - Control - Act) για την λειτουργία των πρακτόρων.

5.1 ΕΙΣΑΓΩΓΗ

Τα εικονικά περιβάλλοντα, που αναφέρονται στη βιβλιογραφία και ως συστήματα εικονικών περιβαλλόντων (virtual environment systems) ή και ως εικονικοί κόσμοι (virtual worlds), είναι σε γενικές γραμμές τα συστήματα τα οποία στο παρελθόν αναφέρονταν με το γενικότερο όρο εικονική πραγματικότητα. Δυστυχώς, μέχρι σήμερα δεν έχουν υπάρξει γενικώς αποδεκτοί ορισμοί για τους παραπάνω όρους. Οι Loffler και Anderson (1994) ορίζουν την εικονική πραγματικότητα ως εξής:

«Εικονική πραγματικότητα είναι ένα τρισδιάστατο περιβάλλον προσομοίωσης σε υπολογιστή του οποίου η απεικόνιση γίνεται σε πραγματικό χρόνο και εξαρτάται από τη συμπεριφορά του χρήστη» Τέλος, ο Gigante (1993) ορίζει την εμπειρία της αλληλεπίδρασης με ένα εικονικό περιβάλλον:

«Η εικονική πραγματικότητα είναι μια εμβυθισμένη, πολυ-αισθητική εμπειρία. Χαρακτηρίζεται από την ψευδαίσθηση της συμμετοχής σε ένα συνθετικό περιβάλλον και όχι απλώς από την εξωτερική παρατήρηση ενός τέτοιου περιβάλλοντος».

Με βάση τους παραπάνω ορισμούς τα βασικά χαρακτηριστικά των εικονικών περιβαλλόντων φαίνεται να είναι τα τρισδιάστατα γραφικά και ένα μοντέλο περιβάλλοντος, που αναπαριστά μια τοποθεσία από την πραγματική ζωή ή κάποια τεχνητή δομή. Ο λόγος για τον οποίο έχει επικρατήσει τα τελευταία χρόνια ο όρος «εικονικό περιβάλλον» έναντι του όρου «εικονική πραγματικότητα» είναι, γιατί στα συστήματα αυτά αφενός δε γίνεται προσπάθεια μοντελοποίησης ολόκληρου του σύμπαντος αλλά ενός περιορισμένου περιβάλλοντος ανάλογα με την εφαρμογή και αφετέρου γιατί δεν είναι υποχρεωτικό η αναπαράσταση να αφορά κάποια ρεαλιστική δομή, αλλά μπορεί κάλλιστα να είναι και κάποιος φανταστικός χώρος με ιδιόμορφους νόμους.

Ένας χρήστης εννοιολογικά «κατοικεί» στο περιβάλλον έχοντας μια τρέχουσα θέση σε αυτό και, κατά συνέπεια, μια περιορισμένη άποψη του χώρου. Έχει την ικανότητα να ταξιδεύει σε αυτό και να αλληλεπιδρά με τα αντικείμενά που τον περιβάλλουν. Τόσο η αντίληψη του περιβάλλοντος όσο και η

αλληλεπίδραση του χρήστη με αυτό μπορούν να μοντελοποιηθούν με βάση την πραγματικότητα. Στην περίπτωση της αντίληψης, ένα τέτοιο παράδειγμα είναι η χρήση πολυαισθητήριων ερεθισμάτων που μιμούνται τα ερεθίσματα του πραγματικού κόσμου ενισχύοντας έτσι την αληθοφάνεια του εικονικού περιβάλλοντος. Αντίστοιχο παράδειγμα αλληλεπίδρασης είναι το βάδισμα του χρήστη πάνω σε κυλιόμενο επίπεδο που προκαλεί την πλοήγησή του στο εικονικό περιβάλλον.

5.1.1 ΚΑΤΗΓΟΡΙΕΣ

Τα εικονικά περιβάλλοντα συνήθως συνδέονται με εξειδικευμένο υλικό που παρέχει πολύπλοκα συστήματα διεπαφής και προσφέρει νέες δυνατότητες αλληλεπίδρασης του χρήστη. Ανάλογα με τη σχέση τους με τον πραγματικό κόσμο και τις συσκευές που χρησιμοποιούνται τα εικονικά περιβάλλοντα χωρίζονται στις εξής κατηγορίες:

- Περιβάλλοντα εμβύθισης (immersive environments): οι χρήστες είναι εφοδιασμένοι με οθόνη προσαρμοσμένη στο κεφάλι (head-mounted display) και δε δέχονται δεδομένα από το φυσικό κόσμο (Slater and Usoh, 1995).
- Περιβάλλοντα οθόνης (desktop environments): Η αναπαράσταση γίνεται σε οθόνη υπολογιστή και έτσι ο χρήστης εξακολουθεί να έχει αντίληψη του φυσικού κόσμου.
- Περιβάλλοντα προβολής (projected environments): Το εικονικό περιβάλλον προβάλλεται σε ένα φυσικό χώρο, όπως ένα δωμάτιο (Wloka, 1996; Hirose, 1996) ή μια επιφάνεια εργασίας (Kruger et al., 1995).
- Ενισχυμένα περιβάλλοντα (augmented environments): Τα εικονικά αντικείμενα προβάλλονται πάνω στον πραγματικό κόσμο, πιθανώς με τη χρήση οθονών προσαρμοσμένων στο κεφάλι που επιτρέπουν στο χρήστη να βλέπει και μέσα από αυτές (see through) (Adam, 1993).

Οι παραπάνω τρόποι διεπαφής με εικονικά περιβάλλοντα έχουν

μια σειρά πλεονεκτημάτων αλλά και μειονεκτημάτων. Τα περιβάλλοντα εμπύθισης έχουν το μεγάλο προσόν ότι προσφέρουν στο χρήστη την απόλυτη εμπειρία του περιβάλλοντος. Παρόλα αυτά, απομονώνουν το χρήστη, χρησιμοποιούν πολύ εξειδικευμένα και ακριβή τεχνολογία και έχουν συνδεθεί με διάφορα προβλήματα υγείας (Travis et al., 1994). Τα περιβάλλοντα οθόνης αποφεύγουν αυτά τα προβλήματα, αλλά δεν παρέχουν εμπειρία πλήρους εμπύθισης. Αντίθετα, βασίζονται στην «ψυχολογική εμπύθιση» (Robertson et al., 1993), δηλαδή σε αρκετές περιπτώσεις καταφέρνουν να επιστήσουν όλη την προσοχή του χρήστη στο περιβάλλον. Τα περιβάλλοντα προβολής περιλαμβάνουν ένα πιο φυσικό σκηνικό και τη δυνατότητα να μοιραστεί η εμπειρία και με άλλους χρήστες (Kruger et al., 1995), απαιτούν όμως κι αυτά τη χρήση εξειδικευμένης τεχνολογίας. Τέλος, τα μεικτά περιβάλλοντα προσθέτουν στον πραγματικό κόσμο επιπλέον εικονικό περιεχόμενο και πληροφορίες (Adam, 1993), αλλά απαιτούν μεγάλη ακρίβεια στο συσχετισμό μεταξύ της πραγματικής εικόνας και της προβαλλόμενης εικονικής πληροφορίας (Wloka, 1996).

5.12 ΤΑΞΙΝΟΜΗΣΗ

Ο Zeltzer (1992) έχει προτείνει μια τρισδιάστατη ταξινόμηση των εικονικών περιβαλλόντων, γνωστή ως κύβος Αυτονομίας - Αλληλεπίδρασης - Παρουσίας (Autonomy - Interaction - Presence, AIP cube). Κάθε άξονας του κύβου μετράει τα επίπεδα αυτονομίας, αλληλεπίδρασης και παρουσίας σε κλίμακα από μηδέν έως ένα (εικόνα 2.1). Τα επίπεδα μετρώνται ως εξής:

- Αυτονομία: η ικανότητα των αντικειμένων του εικονικού περιβάλλοντος να δρουν από μόνα τους. Τα αντικείμενα μπορούν να λειτουργούν ξεχωριστά

υπό τον έλεγχο ενός προγράμματος υπολογιστή είτε ακολουθώντας κάποιο προκαθορισμένο σενάριο, είτε παρουσιάζοντας προσαρμοστική συμπεριφορά.

- Αλληλεπίδραση: το ποσοστό στο οποίο το σύστημα παρέχει έλεγχο των αντικειμένων του εικονικού περιβάλλοντος σε πραγματικό χρόνο και με αληθοφανή τρόπο.
- Παρουσία: το εύρος και η πιστότητα των αισθητήριων καναλιών του χρήστη που ελέγχει το περιβάλλον

Σύμφωνα με τον Zeltzer «δεν είναι δυνατόν να προσομοιωθεί ο φυσικός κόσμος στην πλήρη λεπτομέρεια και πολυπλοκότητά του». Τα σημερινά εικονικά περιβάλλοντα παρουσιάζουν υψηλό βαθμό αλληλεπίδρασης και παρουσίας, χάρη στο εξειδικευμένο υλικό που υπάρχει για τους σκοπούς αυτούς, στερούνται όμως αυτονομίας, και για τον λόγο αυτό η παρούσα διατριβή εστιάζεται κυρίως σε αυτή τη διάσταση.

52 ΑΝΑΠΤΥΞΗ ΕΝΟΣ ΕΙΚΟΝΙΚΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ

Η δημιουργία ενός εικονικού περιβάλλοντος περιλαμβάνει την κατασκευή του μοντέλου του περιβάλλοντος και το σχεδιασμό των αλληλεπιδράσεων με το χρήστη. Όσον αφορά στο μοντέλο του περιβάλλοντος μια κατηγοριοποίηση των επιμέρους τμημάτων του είναι η εξής (Ellis, 1993; Thalmann, 1994):

- ο χώρος σκηνικού ή η γεωμετρία, που παραμένει αμετάβλητη,
- οι χρήστες, που μπορούν να εκτελούν ενέργειες και ελέγχουν το δικό τους οπτικό πεδίο,
- οι πράκτορες ή εικονικοί ηθοποιοί, οι οποίοι έχουν ευφυΐα και μπορούν να ενεργούν ανεξάρτητα από τους χρήστες, και
- τα αντικείμενα που υπάρχουν στο χώρο. Αυτά διαφοροποιούνται στο επίπεδο αλληλεπίδρασης με το χρήστη και στις δυνατότητες αλλαγής της κατάστασής τους.

Σε ό,τι αφορά την αλληλεπίδραση του χρήστη με το περιβάλλον οι βασικοί τύποι είναι (Bordegoni, 1993; Herndon et al., 1994):

- πλοήγηση και έλεγχος οπτικού πεδίου
- αλληλεπιδράσεις με αντικείμενα, όπως πιάσιμο, περιστροφή, μετακίνηση και χρήση αντικειμένων.

5.21 ΜΟΝΤΕΛΟ ΠΕΡΙΒΑΛΛΟΝΤΟΣ

Ο χώρος σκηνικού του περιβάλλοντος μπορεί να κατασκευαστεί με βάση κάποια περιοχή του πραγματικού κόσμου, όπως ένα συγκεκριμένο κτήριο, ή ακόμα και να βασιστεί σε κάποια φανταστική ή αφηρημένη δομή ανάλογα με το είδος του περιβάλλοντος. Σε ό,τι αφορά τη γραφική αναπαράσταση, εκτός από την προβολή του χώρου ως συνόλου από συμπαγή αντικείμενα, που είναι πιο ρεαλιστική, σε πολλές περιπτώσεις αποφέρει πλεονεκτήματα και η χρήση εναλλακτικών τεχνικών, όπως η τεχνική προβολής ακμών (wireframe) (Osborn and Agogino, 1992) που ελαχιστοποιεί τα κρυμμένα αντικείμενα (Stytz et al., 1995), ή η προβολή ημιδιαφανών αντικειμένων. Για παράδειγμα, σε χειρουργικές εφαρμογές είναι πιθανό να υπάρχουν ημιδιαφανείς απόψεις, καθώς και απόψεις με ή χωρίς την απεικόνιση της κυκλοφορίας του αίματος (Kruger et al., 1995). Τέλος, ένα εξίσου σημαντικό ζήτημα που μπορεί να βοηθήσει στην καλύτερη αντίληψη του περιβάλλοντος από το χρήστη είναι ο κατάλληλος φωτισμός του σκηνικού (Stoper and Cohen, 1993). Μια από τις πιο σημαντικές δομές στο οπτικό πεδίο του χρήστη είναι η αναπαράσταση του ίδιου του σώματός του (Slater and Usoh, 1995), δηλαδή των τμημάτων αυτού που είναι ορατά με προοπτική πρώτου προσώπου. Αυτό είναι ακόμα πιο σημαντικό σε εφαρμογές με πολλαπλούς χρήστες, όπου οι ενσαρκώσεις των χρηστών μπορούν να ενισχύσουν την επικοινωνία (Carin et al., 1995) για παράδειγμα παρέχοντας πληροφορίες για τη θέση και την ιδιότητα του καθενός (Benford et al., 1995). Το γενικό χαρακτηριστικό που θα πρέπει να έχουν οι ενσαρκώσεις είναι να είναι συνεπείς και προβλέψιμες (Slater and Usoh, 1994), δηλαδή να υπάρχει μια καλή αντιστοιχία μεταξύ του εικονικού σώματος και της στάσης που έχει το πραγματικό σώμα του χρήστη, έτσι ώστε οι εκτιμήσεις των υπολοίπων χρηστών για τις ενέργειες ενός χρήστη με βάση την ενσάρκωσή του να είναι όσο το δυνατόν πιο κοντά στην πραγματικότητα.

Τα βασικά θέματα που έχουν προκύψει στις αναπαραστάσεις εικονικών περιβαλλόντων είναι ο ρεαλισμός, η αίσθηση του βάθους και η χρήση πολλαπλών αισθήσεων. Σε ό,τι αφορά το ρεαλισμό, τα αντικείμενα μπορούν να είναι από ακριβή αντίγραφα ως αφαιρετικά μοντέλα της πραγματικότητας ή ακόμα και να ενισχύουν την πραγματικότητα με παραπάνω πληροφορίες (Bolzonì, 1994). Η αναπαράσταση των αντικειμένων δε χρειάζεται να είναι πάντοτε φωτορεαλιστική. Ο βαθμός πιστότητας που απαιτείται θα εξαρτηθεί από το πεδίο εφαρμογής και πιθανώς από την εμπειρία του χρήστη με το αντίστοιχο περιβάλλον (Herdson et al., 1994). Ο ρεαλισμός του σκηνικού μπορεί να ενισχυθεί με την μοντελοποίηση περισσότερων λεπτομερειών, όπως σκιές, τραχιές επιφάνειες, λεκέδες, εξογκώματα και φυσικές διαφοροποιήσεις (Fischer et al., 1996). Υπάρχουν βέβαια και περιπτώσεις που λιγότερο ρεαλιστικές αναπαραστάσεις είναι πιο αποδοτικές, όπως για παράδειγμα στην απεικόνιση ιατρικών δεδομένων, όπου η εξομάλυνση των επιφανειών βοηθάει στην καλύτερη κατανόηση, ενώ κάποιες δομές μπορεί να περιβάλλονται με παχιά περιγράμματα για μεγαλύτερη ευκρίνεια (McConathy and Doyle, 1993). Τα μοντέλα με υψηλό επίπεδο λεπτομέρειας μπορούν να περιοριστούν στα αντικείμενα που είναι πιο σημαντικά, σε αυτά που είναι πιο κοντά, ή μόνο σε αυτά που βρίσκονται στο κέντρο του πεδίου προβολής (Watson et al., 1995; Brelstaff, 1995). Σε πιο μεγάλα σκηνικά μπορεί να δοθεί έμφαση στις περιοχές που χρειάζεται πολύπλοκη μετακίνηση και αλληλεπίδραση (McGovern, 1993). Εναλλακτικά, μπορεί να αξιοποιηθεί η πληροφορία για το σημείο στο οποίο κοιτάζει το μάτι του χρήστη (Arthur et al., 1993), ο οποίος μπορεί να επιλέγει ο ίδιος τις περιοχές που θα αναπαρασταθούν με μεγαλύτερη λεπτομέρεια (Yamaashi et al., 1996), ή μπορούν να γίνουν δοκιμές χρηστών, για να καθοριστούν οι απαιτήσεις του τελικού επιπέδου λεπτομέρειας (Pratt et al., 1997).

Η μεταφορά πληροφορίας βάθους είναι πολύ σημαντική για την αναπαράσταση τρισδιάστατων αντικειμένων. Έχουν ελεγχθεί διάφοροι

παράγοντες και έχει υπάρξει σημαντική πρόοδος στη σύγκριση των πλεονεκτημάτων που προκύπτουν από την αξιοποίηση των διαφορετικών ενδείξεων βάθους. Μπορούν να χρησιμοποιηθούν ενδείξεις βάθους εικόνας, όπως προοπτική, επικάλυψη αντικειμένων και φωτοσκίαση (Pimentel and Teixeira, 1993; Wanger et al. 1992). Η κίνηση των αντικειμένων μπορεί να είναι χρήσιμη για την αποκάλυψη της τρισδιάστατης μορφής (Profitt and Kaiser, 1993), ενώ και οι σκιές σε πραγματικό χρόνο αποτελούν μια ακόμη ένδειξη για τις σχετικές θέσεις των αντικειμένων (Arthur et al., 1993; Buck et al., 1996). Έχει μελετηθεί ότι η στερεοσκοπική όραση (δηλαδή η προσαρμογή της απεικόνισης σε κάθε μάτι ξεχωριστά) μπορεί να βοηθήσει στην αντίληψη της δομής και της θέσης (Russell and Miles, 1993; Volbracht et al., 1997), ειδικά σε περιπτώσεις πολύπλοκου σκηنيού και περιορισμένης ορατότητας των αντικειμένων (Kim et al., 1993), καθώς και για υψηλής ακρίβειας αίσθηση του χώρου από το χρήστη (Wann and Mon-Williams, 1996). Εκτός από το οπτικό κανάλι έχουν χρησιμοποιηθεί και άλλοι τρόποι ανάδρασης σε εικονικά περιβάλλοντα. Στη βιβλιογραφία έχουν περιγραφεί δυνατότητες χρήσης ήχου και αφής. Ο ήχος μπορεί να χρησιμοποιηθεί ως μουσική στο περιθώριο, για επικοινωνία ή και για προειδοποιητικά σήματα (Gilkey and Weisenberger, 1995). Ακόμη, μπορούν να υπάρχουν συγκεκριμένοι ήχοι που συνδέονται με ενέργειες και γεγονότα του κόσμου (Pressing, 1997). Ο ήχος μπορεί να πληροφορήσει για διαδικασίες που δεν είναι ορατές στο χρήστη στο συγκεκριμένο χρονικό σημείο (Astheimer et al., 1994), καθώς και να παρέχει ένα δεύτερο κανάλι πληροφοριών (Brown et al., 1989). Για παράδειγμα, ο ήχος χρησιμοποιήθηκε για να παρέχει ανάδραση σε ο,τι αφορά τις ενεργειακές τιμές παράλληλα με την οπτική επαφή σε ένα σύστημα χειρισμού μοριακών δομών (Cruz-Neira, 1996). Ακόμη χρησιμοποιήθηκε για να δημιουργήσει ατμόσφαιρα στο εικονικό περιβάλλον Diamond Park (Waters et al., 1997) για παράδειγμα χρησιμοποιήθηκαν ήχοι του αέρα και των πουλιών σε σκηνικά ανοιχτού χώρου.

Η αλληλεπίδραση αφής μπορεί να χρησιμοποιηθεί για να μεταδώσει ιδιότητες των αντικειμένων κατά τη διάρκεια της χρήσης τους από το χρήστη (Bergamasco, 1994), όπως το σχήμα, η ελαστικότητα, η υφή και η ταχύτητα (Johnson and Cutt, 1992), ή μπορεί να χρησιμοποιηθεί για να δώσει ανάδραση σε σκοτεινά ή γενικώς δυσδιάκριτα σκηνικά (Richard et al., 1996). Για παράδειγμα, η ανάδραση αφής είναι σημαντική για χειρουργικές εφαρμογές, ώστε να υπάρχει ρεαλιστική αίσθηση του δέρματος, του μαλακού ιστού και του κόκαλου (Kruger et

al., 1995). Σε περιπτώσεις που χρειάζεται μεγάλη ακρίβεια στην αλληλεπίδραση του χρήστη με το περιβάλλον, μπορούν να χρησιμοποιηθούν ταυτόχρονα παραπάνω από ένα κανάλια αισθήσεων, όπως συνδυασμός ανάδρασης αφής και ήχου, ώστε να υπάρχουν περισσότερες ενδείξεις για την τρέχουσα κατάσταση του εικονικού περιβάλλοντος.

52.2 ΑΛΛΗΛΕΠΙΔΡΑΣΕΙΣ ΜΕ ΤΟΝ ΧΡΗΣΤΗ

Ένας πολύ σημαντικός τύπος αλληλεπίδρασης σε εικονικά περιβάλλοντα είναι η πλοήγηση του χρήστη και έχουν προταθεί διάφορες τεχνικές για την υλοποίησή του. Η πλοήγηση μπορεί να γίνει με επιλογή σημείου στο χώρο και αυτόματη μετακίνηση σε αυτό (point-and-fly), με μετακίνηση και περιστροφή του ματιού μέσα στη σκηνή (eyeball in hand) ή με διατήρηση σταθερού οπτικού πεδίου και μετακίνηση ολόκληρου του σκηνικού (scene in hand) (Astheimer et al., 1994). Πιο απλές τεχνικές χρησιμοποιώντας ένα κοινό ποντίκι είναι έναρξη / τερματισμός της μετακίνησης με ένα πλήκτρο και διεύθυνση της κίνησης με βάση την κίνηση του ποντικιού (Strommen,) ή χρήση διαφορετικών πλήκτρων του ποντικιού για κίνηση προς τα εμπρός, προς τα πίσω και για σταμάτημα της κίνησης (Bliss et al., 1997). Εναλλακτικά μπορεί να χρησιμοποιηθεί χειρονομία «πιασίματος» και «τραβήγματος» του κόσμου (Mapes and Moshell, 1995), εικονικό χειριστήριο (Mohageg et al., 1996) ή φυσικό βάδισμα πάνω στο σκηνικό. Το βάδισμα μπορεί να βελτιώσει την αίσθηση παρουσίας στο εικονικό περιβάλλον, αλλά οι τεχνικές επιλογής οποιουδήποτε σημείου στο χώρο και μετάβασης σε αυτό μπορούν να οδηγήσουν σε ευκολότερη πλοήγηση (Slater et al., 1995). Το οπτικό πεδίο μπορεί επίσης να «κλειδωθεί» σε συγκεκριμένα αντικείμενα ή ενέργειες στο περιβάλλον (Bolter et al., 1995; Stytz et al., 1995), πράγμα που βοηθάει στη διατήρηση της εστίασης πάνω σε χώρους ενδιαφέροντος (Eyles, 1993). Για παράδειγμα, σε ένα σύστημα οπτικοποίησης του διαστήματος μπορούν να υπάρχουν προεπιλεγμένα οπτικά πεδία σε αστέρια, στη γη ή να ακολουθείται η κίνηση του διαστημικού σκάφους.

Τα θέματα που έχουν προκύψει κατά το σχεδιασμό της πλοήγησης είναι η ταχύτητά της και οι βαθμοί ελευθερίας της. Η τεχνική πλοήγησης μπορεί να περιλαμβάνει μέχρι και έξι βαθμούς ελευθερίας, αλλά ο έλεγχος και των έξι μπορεί να είναι προβληματικός (Drucker and Zeltzer, 1994). Η πλοήγηση μπορεί να περιοριστεί σε δύο διαστάσεις, αποκλείοντας την ικανότητα πετάγματος (Bliss et al., 1997), ή μπορούν να χρησιμοποιηθούν διαφορετικοί τύποι κινήσεων,

όπως μόνο μετακίνησης (όπου ο προσανατολισμός παραμένει σταθερός), μόνο προσανατολισμού και συνδυασμένος (Deering, 1995). Για παράδειγμα, στο πρόγραμμα Berlin 2010 (Vorsteher, 1996) οι τύποι πλοήγησης περιλαμβάνουν ελεύθερο πέταγμα στο σκηνικό ή πεζή μετακίνηση ή μετακίνηση σε τρένο. Η ταχύτητα μετακίνησης θα πρέπει να είναι κατάλληλη για το μέγεθος του εκάστοτε σκηνικού (Mohageg et al., 1996). Οι ταχύτερες μετακινήσεις μειώνουν την προσπάθεια και επιτυγχάνουν γρηγορότερη άφιξη στο στόχο, αλλά και οι πιο αργές δίνουν μεγαλύτερη ακρίβεια στην προσέγγιση ενός στόχου (Johnsgard, 1994). Η μετακίνηση με σταθερή ταχύτητα μπορεί να οδηγήσει σε προβλήματα, όπως η επιλογή της θέσης που πρέπει να σταματήσει κάποιος. Αντίθετα, με τη χρήση μεταβλητής ταχύτητας, ο τρόπος μετακίνησης γίνεται πιο φυσικός (Rushton and Wann, 1993) και δίνει στους χρήστες έλεγχο πάνω στην ταχύτητά της (Mohageg et al., 1996).

Ο άλλος πολύ βασικός τύπος αλληλεπίδρασης στα εικονικά περιβάλλοντα είναι η αλληλεπίδραση με αντικείμενα. Οι αλληλεπιδράσεις με αντικείμενα μπορούν να επιτευχθούν χρησιμοποιώντας τεχνικές, όπως η αποστολή ακτίνων (raycasting) για επιλογή αντικειμένων, ο άμεσος χειρισμός για τοποθέτηση αντικειμένων ή οι χειρονομίες (Herndon et al., 1994), δηλαδή η κίνηση των χεριών με τρόπο όμοιο με αυτόν που χρησιμοποιείται στην αλληλεπίδραση με αντίστοιχο αντικείμενο του πραγματικού κόσμου. Οι χειρονομίες είναι μια φυσική τεχνική αλληλεπίδρασης (Astheimer et al., 1994) και είναι ιδιαίτερα αποδοτική, ειδικά εφόσον υπάρχουν κάποιοι περιορισμοί στις επιτρεπόμενες χειρονομίες (Wexelblat, 1995). Παρόλα αυτά, για την επιλογή ενός αντικειμένου το ποντίκι είναι συνήθως πιο γρήγορο και πιο ακριβές από ένα γάντι δεδομένων (Johnsgard, 1994). Μπορούν επίσης να χρησιμοποιηθούν και τα δύο χέρια για το χειρισμό αντικειμένων (Mapes and Moshell, 1995) για παράδειγμα μεγέθυνση και συρρίκνωση αντικειμένων με δύο χέρια (Mercurio and Erickson, 1990). Εναλλακτικά, μπορεί να χρησιμοποιηθεί η φωνή για αλληλεπίδραση. Αυτή μπορεί είτε να δίνει εντολές σε πράκτορες ή γενικά στο περιβάλλον, είτε να χρησιμοποιηθεί για διαλόγους με αντικείμενα ενδιαφέροντος (McGlashan and Axling, 1996). Παρόλο που το λεξιλόγιο θα πρέπει να είναι περιορισμένο, η φωνή επιτρέπει την αποστολή εντολών χωρίς να απασχολεί τα χέρια και τα μάτια και έτσι μπορούν να εκτελούνται ακόμα και ενέργειες που αφορούν αντικείμενα που δεν είναι στο οπτικό πεδίο του χρήστη (McGlashan and Axling, 1996). Η χρήση συνδυασμού φωνής και χειρονομιών είναι μια καλή τεχνική που μπορεί να αυξήσει την ικανότητα του συστήματος να αναγνωρίζει τις ενέργειες του χρήστη (Hauptmann

and McAvinney, 1993).

Οι αλληλεπιδράσεις με τα αντικείμενα μπορούν να προγραμματιστούν σε διαφορετικούς βαθμούς ρεαλισμού ως απλοποιημένες ενέργειες του πραγματικού κόσμου, ως πολύπλοκες ενέργειες που αναπαριστούν με πιστότητα τον τρόπο αλληλεπίδρασης ή και ως ενέργειες που δεν είναι δυνατό να συμβούν στην πραγματικότητα (Slater and Usoh, 1994). Ο απαιτούμενος βαθμός ρεαλισμού εξαρτάται από το πεδίο εφαρμογής (Herndon et al., 1994). Για παράδειγμα, σε περιπτώσεις εργονομικής ανάλυσης το πιάσιμο των αντικειμένων μπορεί να χρειαστεί να μοντελοποιηθεί όσο το δυνατόν πιο κοντά στην πραγματικότητα, ενώ σε άλλες περιπτώσεις μπορεί να είναι αρκετό το άγγιγμά του (Dai et al, 1996). Τέλος, σε περιβάλλοντα που δεν έχουν στενή σχέση με τον πραγματικό κόσμο, μπορούν να χρησιμοποιηθούν και μεταφορές, οι οποίες μπορούν να βοηθήσουν σημαντικά το χρήστη. Είναι παρόλα αυτά σημαντικό οι μεταφορές αυτές να είναι σχεδιασμένες με τέτοιο τρόπο, που να γίνονται εύκολα κατανοητές από το χρήστη.

4.3 ΤΕΧΝΟΛΟΓΙΕΣ

Κάθε εικονικό περιβάλλον αποτελείται από το υλικό που θα χρησιμοποιήσει ο χρήστης και από το λογισμικό που εκτελείται σε κάποιον υπολογιστή. Το υλικό χρησιμοποιείται αφενός για να καταγράψει την τρέχουσα κατάσταση του χρήστη, όπως θέση στο χώρο, κινήσεις χεριών και κεφαλιού, ομιλία, κλπ., και αφετέρου για να μεταδώσει τα κατάλληλα ερεθίσματα στις αισθήσεις του, δηλαδή στην όραση, ακοή και αφή του. Αντίθετα το λογισμικό περιλαμβάνει το ίδιο το περιβάλλον αλλά και τα τμήματα διεπαφής του υλικού με το περιβάλλον. Στην εικόνα 4 παρουσιάζεται ένα γενικό σύστημα εικονικού περιβάλλοντος, όπου η εικόνα και ο ήχος προσαρμόζονται στην κίνηση του κεφαλιού, και υπάρχει αλληλεπίδραση με το περιβάλλον με τη χρήση γαντιού δεδομένων και με φωνητικές εντολές.

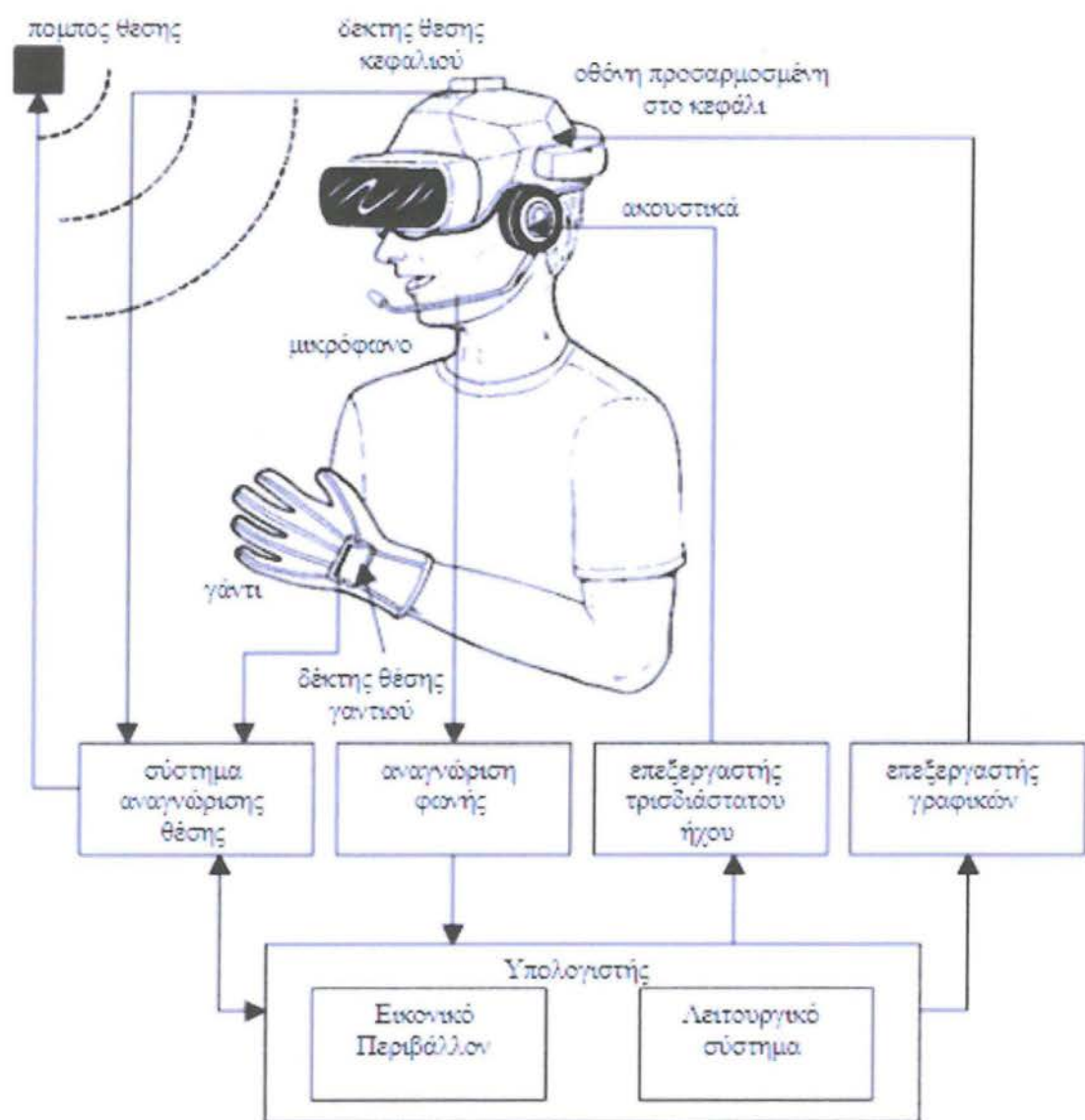
5.3.1 ΥΛΙΚΟ

Οι συσκευές που χρησιμοποιούνται ως σήμερα σε εικονικά περιβάλλοντα είναι οι παρακάτω:

- Συσκευές εισόδου: Υπάρχει ένας αριθμός συσκευών που χρησιμοποιούνται για τη μετακίνηση του χρήστη στο περιβάλλον σε όλες τις διαστάσεις με

διάφορους βαθμούς ελευθερίας. Τέτοια συσκευές είναι το χειριστήριο (joystick), η σφαίρα πλοήγησης (space ball) (εικόνα 2.3:δ) και το τρισδιάστατο ποντίκι. Ακόμα, μπορούν να χρησιμοποιηθούν γάντια δεδομένων (data gloves) (εικόνα 2.3:γ) που εντοπίζουν τη θέση και την κίνηση του χεριού και των δακτύλων για αλληλεπίδραση με εικονικά αντικείμενα. Συσκευές εντοπισμού θέσης (position trackers) ή καταγραφής video μπορούν να χρησιμοποιηθούν για να εντοπιστούν κινήσεις του κεφαλιού ή του σώματος, ώστε να βοηθήσουν τη μετακίνηση του χρήστη στο περιβάλλον ή την αλληλεπίδραση με αντικείμενα.

- Συσκευές οπτικοποίησης: Οθόνες προσαρμοσμένες στο κεφάλι (head-mounted displays) που περιλαμβάνουν ανεξάρτητες οθόνες για κάθε μάτι, μπορούν να χρησιμοποιηθούν για να παρέχουν εμβυθισμένη, στερεοσκοπική εμφάνιση (εικόνα 2.3:α,β). Εναλλακτικά, μπορούν να χρησιμοποιηθούν συσκευές προβολής video που προβάλλουν το περιβάλλον σε φυσικό μέγεθος πάνω σε τοίχους δωματίων ή επιφάνειες τραπεζιών.
- Συσκευές τρισδιάστατου ήχου: οι συσκευές αυτές μπορούν να παράγουν πολυφωνικούς ήχους με τρόπο τέτοιο, ώστε να γίνεται αντιληπτή από το χρήστη η σχετική θέση στο χώρο της κάθε πηγής ήχου.
- Συσκευές ανάδρασης: Υπάρχουν συσκευές ανάδρασης δυνάμεων και ανάδρασης αφής, που μπορούν να χρησιμοποιηθούν για να προσομοιώσουν την αίσθηση που έχει το δέρμα, όταν αγγίζει διαφόρων ειδών υλικά, ή για να εξομοιώσουν τη βαρύτητα ή την αντίσταση στην κίνηση.



Εικόνα 4: Ένα γενικό σύστημα εικονικού περιβάλλοντος και τα επιμέρους τμήματά του



Εικόνα 4.1: Συσκευές εικονικής πραγματικότητας.
 α) Οθόνη προσαρμοσμένη στο κεφάλι,
 β) Γυαλιά ενισχυμένης πραγματικότητας,
 γ) γάντια δεδομένων,
 δ) σφαίρα πλοήγησης

5.32 ΛΟΓΙΣΜΙΚΟ

Το λογισμικό διακρίνεται σε δύο βασικές κατηγορίες, το λογισμικό ανάπτυξης, δηλαδή αυτό που χρησιμοποιείται για το σχεδιασμό του εικονικού περιβάλλοντος και το λογισμικό εκτέλεσης, δηλαδή αυτό που εκτελείται κατά τη διάρκεια της αλληλεπίδρασης του χρήστη με το περιβάλλον.

Κατά την διάρκεια της ανάπτυξης του περιβάλλοντος χρησιμοποιούνται οι παρακάτω τύποι προγραμμάτων:

- πρόγραμμα τρισδιάστατης μοντελοποίησης ή τρισδιάστατης ψηφιοποίησης από

δεδομένα του πραγματικού κόσμου για τη δημιουργία των τρισδιάστατων μοντέλων,

- εργαλεία βελτιστοποίησης γεωμετρίας για μείωση του όγκου των δεδομένων και καλύτερη απόδοση του περιβάλλοντος,

- εργαλεία επεξεργασίας εικόνας για την κατασκευή των υφών που θα προστεθούν στα μοντέλα γενικού τύπου γλώσσα προγραμματισμού ή ειδική script γλώσσα για εικονικά περιβάλλοντα για τον προγραμματισμό των κανόνων του κόσμου και της κίνησης των αντικειμένων.

Οι διαδικασίες που εκτελούνται κατά τη διάρκεια εκτέλεσης ενός εικονικού περιβάλλοντος είναι οι παρακάτω:

- Μηχανή τρισδιάστατης οπτικοποίησης και παραγωγής ήχου σε πραγματικό χρόνο: είναι η διαδικασία που δημιουργεί την προοπτική πρώτου προσώπου του ρήστη και αναπαριστά το περιβάλλον σε πραγματικό χρόνο. Σε περίπτωση που το περιβάλλον περιέχει και τρισδιάστατο ήχο αναλαμβάνει και την προσαρμογή του ήχου ανάλογα με την τρέχουσα θέση του χρήστη.

- Διαδικασία χειρισμού συσκευών εισόδου / εξόδου του χρήστη: η διαδικασία αυτή λαμβάνει δεδομένα από τις συσκευές εισόδου (π.χ. γάντι δεδομένων, μικρόφωνο, θέση κεφαλιού) και τα μεταφράζει σε δεδομένα του εικονικού περιβάλλοντος. Αντίστοιχα, μετά από κάθε οπτικοποίηση ή παραγωγή τρισδιάστατου ήχου μεταφέρει την εικόνα και τον ήχο στις αντίστοιχες συσκευές εξόδου.

- Διαδικασία προσομοίωσης: είναι ουσιαστικά ο «πυρήνας» του εικονικού περιβάλλοντος, καθώς είναι η διαδικασία που εφαρμόζει τους νόμους του περιβάλλοντος (που μπορεί να αφορούν στην κίνηση ή σε άλλες επιτρεπτές ενέργειες) στα αντικείμενα και υπολογίζει τις νέες τους ιδιότητες.

- Βάση δεδομένων γεωμετρίας του περιβάλλοντος: η γεωμετρία του περιβάλλοντος αποθηκεύεται και ανακτάται με τέτοια μορφή, ώστε το περιβάλλον να έχει όσο το

δυνατόν μεγαλύτερη απόδοση. Η διαδικασία αυτή αναλαμβάνει να μην εμφανίσει τα αντικείμενα που δεν είναι στο οπτικό πεδίο του χρήστη ή να αντικαταστήσει τα αντικείμενα που βρίσκονται μακριά από το χρήστη με αντίστοιχα μειωμένης λεπτομέρειας.

Σήμερα υπάρχει ήδη ένα πρότυπο τυποποιημένο από το διεθνή οργανισμό τυποποίησης για την αναπαράσταση τρισδιάστατων δεδομένων στο Internet, η VRML97. Το πρότυπο είναι ιδιαίτερα πολύπλοκο, με αποτέλεσμα πολλές από τις εφαρμογές που ισχυρίζονται ότι το χρησιμοποιούν να μην παρέχουν πλήρη υποστήριξη. Ενώ υπάρχει υποστήριξη του μοντέλου γεωμετρίας της VRML από τα δημοφιλή πακέτα τρισδιάστατης μοντελοποίησης, έχουν υπάρξει λίγες υλοποιήσεις του πλήρους μοντέλου συμβάντων (events). Για τις ανάγκες του Παγκόσμιου Ιστού έχουν γραφτεί ορισμένα δημοφιλή προγράμματα που εγκαθίστανται στο πρόγραμμα πλοήγησης (browser), όπως τα Cosmo Player, Blaxxun Contact και ParallelGraphics Cortona. Με την χρήση της γλώσσας VRML σε συνδυασμό με κάποιο από τα παραπάνω προγράμματα μπορούν να κατασκευαστούν αξιολογικά μικρής κλίμακας εικονικά περιβάλλοντα οθόνης, όμως η απόδοση τέτοιων συστημάτων δεν είναι ιδιαίτερα ικανοποιητική.

Ως συνέχεια της VRML97 ετοιμάζεται σήμερα ένα νέο πρότυπο, το X3D. Το πρότυπο αυτό σε γενικές γραμμές είναι η VRML97, που έχει μετατραπεί σε μορφή XML, παρέχοντας όμως την αναγκαία προς τα πίσω συμβατότητα, ώστε να υποστηρίζονται οι υπάρχουσες εφαρμογές.

Ένας αριθμός από εικονικούς κόσμους έχει αναπτυχθεί με τη χρήση μηχανών παιχνιδιών για PC (Cavazza et al, 2001), όπως οι τρισδιάστατες μηχανές των δημοφιλών παιχνιδιών Doom, Quake και Unreal. Ορισμένες τέτοιες μηχανές, όπως αυτή του Unreal, συνοδεύονται από ισχυρές γλώσσες σεναρίων (script languages) για την καλύτερη παραμετροποίηση του περιβάλλοντος, αλλά τείνουν να λειτουργούν καλύτερα σε εικονικά περιβάλλοντα που έχουν παρόμοια δομή με αυτά για τα οποία σχεδιάστηκαν οι μηχανές. Ακόμη, έχουν την τάση να θυσιάζουν την υποστήριξη γενικού τύπου τρισδιάστατων μοντέλων για χάρη της απόδοσης και των οπτικών εφφέ, και κατά συνέπεια είναι πολύ καλές για τους χώρους για τους οποίους έχουν σχεδιαστεί, αλλά δεν αποτελούν λύση που καλύπτει όλες τις περιπτώσεις.

Τέλος, τη μεγαλύτερη ευκολία και άνεση στη δημιουργία ενός εικονικού περιβάλλοντος τη δίνουν οι διάφορες βιβλιοθήκες τρισδιάστατων γραφικών, οι οποίες περιλαμβάνουν έτοιμες ρουτίνες για απεικόνιση και κίνηση τρισδιάστατων δεδομένων και διευκολύνουν τον προγραμματιστή στην ανάπτυξη τέτοιων εφαρμογών. Υπάρχουν βιβλιοθήκες χαμηλού επιπέδου, όπως η OpenGL και το Direct3D, και βιβλιοθήκες υψηλότερου επιπέδου, που χρησιμοποιούν τις παραπάνω, για να παρέχουν επιπλέον λειτουργικότητα. Παραδείγματα τέτοιων βιβλιοθηκών είναι η Java3D, που ενισχύει τα προγράμματα Java με δυνατότητες αναπαράστασης και χειρισμού τρισδιάστατων δεδομένων, το OpenInventor, ένα ολοκληρωμένο εργαλείο για την κατασκευή αλληλεπιδραστικών τρισδιάστατων εφαρμογών, και το Performer, που είναι εξειδικευμένο για συστήματα πολλαπλών επεξεργαστών.

5.4 ΕΦΑΡΜΟΓΕΣ

Τα εικονικά περιβάλλοντα έχουν χρησιμοποιηθεί με επιτυχία σε πολλά και διαφορετικά πεδία εφαρμογών:

Στις τέχνες και τη διασκέδαση τα εικονικά περιβάλλοντα μπορούν να παρέχουν μια πιο ενδιαφέρουσα και αλληλεπιδραστική εμπειρία για το χρήστη, όπως για παράδειγμα τα παιχνίδια εικονικής πραγματικότητας ή το τρισδιάστατο θέατρο.

Εφαρμογές μάρκετινγκ όπως εικονικά ταξίδια με ξεναγό, πριν από την κράτηση διακοπών ή εικονικός εκθεσιακός χώρος ελκύουν την προσοχή πιθανών πελατών και τους επιτρέπουν να δοκιμάσουν προϊόντα με νέους, πιο ρεαλιστικούς τρόπους.

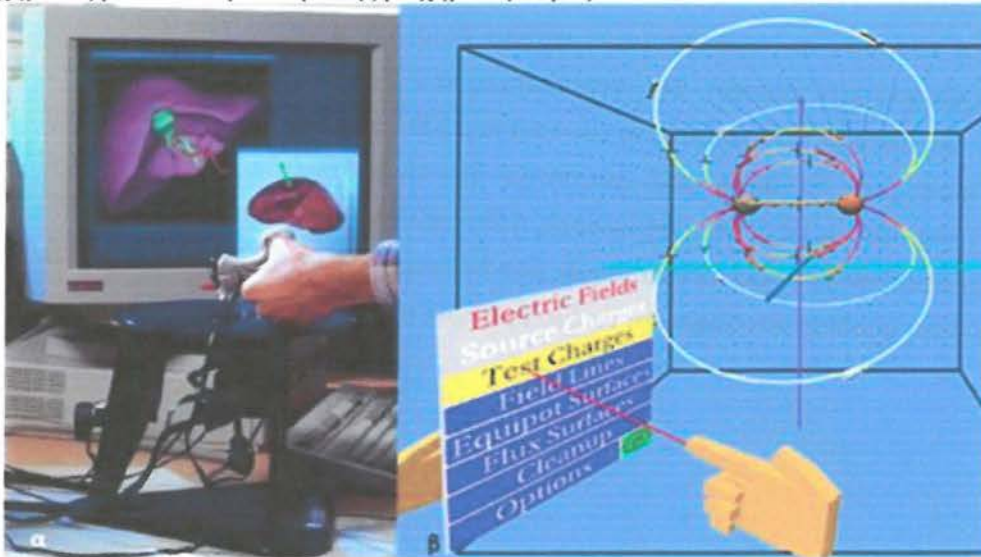
Στον τομέα του τηλεχειρισμού (teleoperation) τα εικονικά περιβάλλοντα παρέχουν ένα ρεαλιστικό σύστημα διεπαφής μέσα από το οποίο επικίνδυνες ή απομακρυσμένες ενέργειες του πραγματικού κόσμου μπορούν να λάβουν χώρα έμμεσα με την τηλεκατεύθυνση ρομπότ, όπως για παράδειγμα η επισκευή πυρηνικών αντιδραστήρων ή συντήρηση δικτύου υπολογιστών.

Η Τηλεπαρουσία (telepresence) και τα συνεργατικά εικονικά περιβάλλοντα, όπως η τηλεχειρουργική (Docimo et al., 1997) ή οι εικονικές πόλεις (Loeffler and Anderson,

1994) επιτρέπουν σε ανθρώπους που βρίσκονται σε μεγάλη απόσταση μεταξύ τους να επικοινωνήσουν ή να εργαστούν μαζί σε ένα ρεαλιστικό περιβάλλον.

Εφαρμογές σχεδιασμού και αξιολόγησης επιτρέπουν την οπτικοποίηση ενός μοντέλου από το εσωτερικό του, καθώς και τον χειρισμό του. Για παράδειγμα, εικονικά μοντέλα για σχεδιασμό ρούχων ή έλεγχος ασφάλειας από φωτιά με μετακίνηση σε εικονικά μοντέλα κτηρίων.

Σε ό,τι αφορά την εκπαίδευση, τα εικονικά περιβάλλοντα επιτρέπουν σε ένα χρήστη να εξερευνήσει και να μάθει για μια θεματική περιοχή με άμεσο πειραματισμό, ειδικά σε περιπτώσεις που το αντικείμενο δεν είναι προσβάσιμο στην πραγματικότητα (Bullinger, 1996). Τέτοια παραδείγματα είναι η εκμάθηση φυσικής μέσα από το πρόγραμμα NewtonWorlds (Dede et al., 1994) (εικόνα 2.4:β) ή η γνώση της χημείας μέσω αλληλεπίδρασης με χημικές δομές.



Εικόνα 4.2: Εφαρμογές εικονικής πραγματικότητας α) στην ιατρική, β) στην εκπαίδευση

Εικονικά περιβάλλοντα εξάσκησης, όπως προσομοιωτές πτήσης για την εξάσκηση των πιλότων ή εικονικές χειρουργικές επεμβάσεις για την εξάσκηση των χειρουργών (εικόνα 4.2:α), παρέχουν ρεαλιστικά περιβάλλοντα, στα οποία μπορεί κανείς να εξασκηθεί σε διάφορες ικανότητες, κυρίως σε εκείνες της αντίληψης του χώρου (Regian and Shebilske, 1992).

Παρόμοια, εφαρμογές θεραπείας, όπως αυτή της ακροφοβίας (Hodges et al., 1995), ή εφαρμογές ρόλων για ασθενείς σε στάδιο ψυχοθεραπείας, παρέχουν ρεαλιστικά περιβάλλοντα στα οποία οι ασθενείς βοηθούνται στο να ξεπεράσουν το πρόβλημά τους μέσω έκθεσης σε καταστάσεις αυξανόμενης δυσκολίας.

Τέλος, τα εικονικά περιβάλλοντα οπτικοποίησης πληροφορίας παρέχουν ένα αλληλεπιδραστικό τρισδιάστατο οπτικό μοντέλο ενός συνόλου πληροφοριών, που βοηθάει στην εξερεύνηση και ερμηνεία των δεδομένων, όπως για παράδειγμα το σύστημα διαχείρισης πληροφοριών NASA VIEW (Fisher, 1990) ή οι εικονικές βιβλιοθήκες.

Όπως φαίνεται και από τα παραπάνω, τα εικονικά περιβάλλοντα έχουν εφαρμοστεί με επιτυχία σε πληθώρα εφαρμογών. Οι εφαρμογές αυτές μοιράζονται ένα κοινό θέμα, που είναι η εμπειρία, εξερεύνηση, και διαχείριση κόσμων γεννημένων σε υπολογιστή. Οι κόσμοι αυτοί αναπαριστούν κάποια δομή ή τοποθεσία είτε του πραγματικού κόσμου είτε τεχνητή. Στις περιπτώσεις που μοντελοποιούνται καταστάσεις του πραγματικού κόσμου, τα εικονικά περιβάλλοντα μπορούν να είναι ιδιαίτερα χρήσιμα, όταν το πραγματικό αντίστοιχο δεν είναι δυνατόν να προσεγγιστεί για λόγους μεγέθους, απόστασης, ή επικινδυνότητας. Τέλος τα εικονικά περιβάλλοντα δεν είναι χρήσιμα σε εφαρμογές που δεν είναι αναγκαίο για το χρήστη να μετακινείται στον τρισδιάστατο χώρο και να έχει εμπειρία από διαφορετικές απόψεις.

5.5 ΕΙΚΟΝΙΚΟΙ ΠΡΑΚΤΟΡΕΣ

5.5.1 ΕΙΣΑΓΩΓΗ

Η χρήση των εικονικών περιβαλλόντων ως αλληλεπίδραση με το χρήστη δε θα πρέπει να περιορίζεται απλώς στην πλοήγηση σε ένα όμορφο τρισδιάστατο σκηνικό. Όσο ελκυστικός κι αν είναι ένας συνθετικός κόσμος, εάν δεν παρουσιάζει μεταβολές, μειώνεται το ενδιαφέρον του χρήστη κατά την εμπύθισή του στο περιβάλλον. Για παράδειγμα, οι πόλεις χωρίς ανθρώπους δεν μπορούν να κάνουν τους χρήστες να έχουν αληθινή αίσθηση παρουσίας σε αυτές. Η συνθετική κίνηση μπορεί να βοηθήσει στη δυναμικότητα του περιβάλλοντος, αλλά το γεγονός, ότι όλες οι μεταβολές είναι προαποφασισμένες, είναι αντίθετο με την ελευθερία αλληλεπίδρασης που θα έπρεπε να έχει ένα εικονικό περιβάλλον, και μπορεί να κρατήσει το ενδιαφέρον του χρήστη για περιορισμένο χρόνο. Οι εικονικοί κόσμοι μπορούν να γίνουν πιο ελκυστικοί, εφόσον υπάρχει το στοιχείο της αυτονομίας και λαμβάνουν χώρα σε αυτούς ενέργειες που δεν προκαλούνται από το χρήστη. Οι αυτόνομες οντότητες σε ένα εικονικό περιβάλλον ονομάζονται εικονικοί πράκτορες (virtual agents). Ένας εικονικός πράκτορας που δεν ακολουθεί προκαθορισμένα βήματα, αλλά μπορεί να παίρνει αποφάσεις από μόνος του και χρησιμοποιεί τεχνικές από το χώρο της Τεχνητής Νοημοσύνης, για να επιτυγχάνει τους στόχους του, ονομάζεται ευφυής εικονικός πράκτορας (intelligent virtual agent), ενώ ευφυή εικονικά περιβάλλοντα (intelligent virtual environments) ονομάζονται τα εικονικά περιβάλλοντα που περιέχουν ευφυείς εικονικούς πράκτορες (Aylett and Luck, 2000; Aylett and Cavazza, 2001).

Οι εικονικοί πράκτορες έχουν συνήθως τη μορφή συνθετικών χαρακτήρων και αλληλεπιδρούν με το περιβάλλον μέσω αισθητήρων (sensors) και επιδραστών (effectors) λαμβάνοντας αποφάσεις με βάση κάποιο μοντέλο συμπεριφοράς. Οι χαρακτήρες δε χρειάζεται να είναι κατ' ανάγκη ανθρωπίνοι σε μορφή: μπορούν να είναι αφηρημένοι (Sims, 1995), μηχανικοί (Prophet, 1996), μπορούν να είναι ζώα, όπως πουλιά (Reynolds, 1987) ή ψάρια (Terzopoulos et al, 1994; Martihno et al, 2000) ή μπορούν να είναι φανταστικοί, όπως τα Teletubbies (Aylett et al, 1999). Οι ανθρώπινες μορφές (Badler et al, 1993b) είναι, φυσικά, οι πιο συνηθισμένες.

Εκτός από τους εικονικούς πράκτορες υπάρχουν και άλλα, παραπλήσια είδη συνθετικών χαρακτήρων σε εικονικά περιβάλλοντα, στα οποία γίνεται αρκετές φορές

αναφορά στο παρόν κεφάλαιο, καθώς η έρευνα είναι σε πολλά σημεία κοινή. Τα είδη αυτά είναι:

- **εικονικός ηθοποιός (virtual actor)** (Shawver, 1997; Wavish and Connah, 1997): Διαφέρει από έναν εικονικό πράκτορα στο ότι δεν έχει αυτονομία. Όλες οι ενέργειές του είναι προκαθορισμένες από κάποιο σενάριο ή βασίζονται σε σχετικές εντολές του χρήστη (Cavazza et al., 1998).
- **ενσάρκωση (avatar)** (Damer, 1998): εξυπηρετεί ως εικονικός εκπρόσωπος ενός πραγματικού ανθρώπου και οι ενέργειές του ελέγχονται (σε πραγματικό χρόνο) από το άτομο αυτό. Συνεπώς, εκτός από έλλειψη αυτονομίας, χαρακτηρίζεται και από έλλειψη αισθητήρων, καθώς η αντίληψη του χώρου γίνεται από τον ίδιο το χρήστη.
- **εικονικός άνθρωπος (virtual human)** (Magnetat Thalmann and Thalmann, 1998; Gratch et al., 2000): είναι μια ειδική περίπτωση εικονικού πράκτορα στην οποία έχει δοθεί έμφαση στη λεπτομερή μοντελοποίηση και προσομοίωση της ανθρώπινης κίνησης και αλληλεπίδρασης. Συνήθως οι χαρακτήρες αυτοί έχουν ιδιαίτερα περίπλοκους αισθητήρες και επιδραστές αλλά χαμηλό επίπεδο αυτονομίας.

5.5.1 ΤΑΞΙΝΟΜΗΣΗ ΕΙΚΟΝΙΚΩΝ ΠΡΑΚΤΟΡΩΝ

Μια πρώτη προσπάθεια ταξινόμησης της υπάρχουσας έρευνας στους εικονικούς πράκτορες καταλήγει σε ένα φάσμα πρακτόρων. Στη μία άκρη του φάσματος τοποθετούνται οι φυσικοί πράκτορες (physical agents), δηλαδή αυτοί στους οποίους η έμφαση έχει δοθεί στην αληθοφανή φυσική συμπεριφορά. Στο χώρο αυτό η έρευνα επικεντρώνεται σε εικονικούς ανθρώπους ή και ζώα και τα θέματά της περιλαμβάνουν τη ρεαλιστική κίνηση και τη φυσική αλληλεπίδρασή τους με το περιβάλλον, την ομιλία, τις χειρονομίες και τις εκφράσεις προσώπου. Τέτοιοι πράκτορες συνήθως αλληλεπιδρούν με το περιβάλλον μέσω εικονικών αισθητήρων και επιδραστών που δε λειτουργούν σε συμβολικό επίπεδο, αλλά προσομοιώνουν το φυσικό (Badler et al., 1993a).

Στην άλλη άκρη του φάσματος αυτού τοποθετούνται οι γνωσιακοί πράκτορες (cognitive agents), δηλαδή αυτοί στους οποίους η έμφαση έχει δοθεί στην ανθρώπινη γνωσιακή συμπεριφορά, καθώς και στη γνωσιακή αλληλεπίδραση με τους χρήστες του συστήματος. Τα περισσότερα θέματα στο χώρο αυτό αφορούν τη φυσική γλώσσα, καθώς και τις διαδικασίες απόφασης, όπως είναι ο σχεδιασμός ενεργειών (planning) (Badler et al., 1996). Τέτοιοι πράκτορες συνήθως λαμβάνουν μέσω των αισθήσεών τους συμβολική πληροφορία απευθείας από το περιβάλλον, με αποτέλεσμα να αμφισβητείται το κατά πόσο η διαδικασία αίσθησης είναι πραγματικά αυτόνομη.

Ο λόγος για τον οποίο γίνεται αναφορά σε φάσμα και όχι σε κατηγορίες αμοιβαίως αποκλειόμενες είναι, γιατί οι πράκτορες που κινούνται περισσότερο σε γνωσιακό επίπεδο απαιτούν ένα βαθμό φυσικής αλληλεπίδρασης με το περιβάλλον, ενώ και οι φυσικοί πράκτορες συχνά απαιτούν ένα είδος ελέγχου σε γνωσιακό επίπεδο.

Θα μπορούσε να χαρακτηρίσει κάποιος την έρευνα που γίνεται στη γνωσιακή άκρη του φάσματος ως δουλειά από το μυαλό προς τα έξω, ενώ την αντίστοιχη έρευνα στο φυσική άκρη ως δουλειά από το σώμα προς τα μέσα. Στην ιδανική τους μορφή οι εικονικοί πράκτορες θα πρέπει να έχουν εντελώς ρεαλιστική κίνηση και φυσική αλληλεπίδραση αλλά και ανθρώπινες γνωσιακές ικανότητες. Στην πραγματικότητα όμως και οι δύο περιοχές απαιτούν την επίλυση αρκετών δύσκολων προβλημάτων, με αποτέλεσμα να υπάρχει μια τάση στις ερευνητικές ομάδες να δίνουν μεγαλύτερη έμφαση στη μία ή την άλλη μεριά του φάσματος. Αυτή η διαφορά εστίασης μπορεί να

εντοπιστεί σε διάφορα θέματα σχετικά με εικονικούς πράκτορες, όπως το θέμα των συναισθημάτων που αφορά και τις δύο άκρες του φάσματος.

Στη συνέχεια αναφέρονται ορισμένα από τα ερευνητικά θέματα που εμφανίζονται στις δύο αυτές κατηγορίες εικονικών πρακτόρων, που καταδεικνύουν και τις δυσκολίες που εμφανίζονται, εφόσον επιθυμεί κάποιος την ανάπτυξη ενός μοντέλου εικονικού πράκτορα που έχει ενισχυμένες τόσο τις φυσικές, όσο και τις γνωσιακές ικανότητές του. Φυσικοί Πράκτορες

Όπως είναι αναμενόμενο, το βάρος στην κατηγορία αυτή πέφτει στη φυσική συμπεριφορά. Πρώτα απ' όλα θα πρέπει να αντιμετωπιστεί η κίνηση του σώματος και η μετακίνηση του πράκτορα στο περιβάλλον (Faloutsos et al., 2001), ένα θέμα που επηρεάζει άμεσα τον τρόπο σχεδιασμού και αναπαράστασης του σώματος. Στο σημείο αυτό η έννοια της ενσάρκωσης είναι εξίσου σημαντική για τους εικονικούς πράκτορες, όσο και για τους πραγματικούς ανθρώπους. Εφόσον υπάρχει μια περίπλοκη φυσική αναπαράσταση, τότε υπάρχει η δυνατότητα να χρησιμοποιηθεί για πιο ρεαλιστικές αλληλεπιδράσεις με τα αντικείμενα αλλά και για ενίσχυση της επικοινωνίας με το χρήστη μέσω βλέμματος, έκφρασης προσώπου και χειρονομιών (Gibet et al., 2001).

Ακόμα, καθώς οι πράκτορες έχουν τη δυνατότητα της μετακίνησης, θα πρέπει να είναι σε θέση να αποφύγουν ανεπιθύμητες συγκρούσεις με άλλα αντικείμενα στο περιβάλλον τους, είτε αυτά είναι στατικά, όπως δέντρα, κτίρια και έπιπλα, είτε κινούμενα, όπως άλλοι πράκτορες. Η ύπαρξη παραπάνω πρακτόρων στο περιβάλλον δημιουργεί ζητήματα συντονισμένης κίνησης μεγάλου αριθμού ατόμων (Reynolds, 1987; Muse and Thalmann, 1997), ενώ θα πρέπει επίσης να αντιμετωπιστούν και άλλα είδη αλληλεπίδρασης, που περιλαμβάνουν επαφή με το περιβάλλον πέρα από απλές συγκρούσεις, όπως το πιάσιμο των αντικειμένων ή οι φυσικές αλληλεπιδράσεις με άλλους πράκτορες, όπως η χειραψία.

Τέλος, με δεδομένη την ύπαρξη μιας πολύπλοκης φυσικής αναπαράστασης και ενός ρεπερτορίου φυσικών συμπεριφορών, προκύπτουν διάφορα θέματα ελέγχου. Για παράδειγμα, θα πρέπει να καθοριστεί το επίπεδο στο οποίο θα πρέπει να ασκηθεί ο έλεγχος, που περιλαμβάνει την πιθανότητα ελέγχου στο επίπεδο των ξεχωριστών μυών ή στο επίπεδο των ολοκληρωμένων ενεργειών συμπεριφοράς (Badler et al., 2000),

όπως το βάδισμα και το πιάσιμο των αντικειμένων ή ακόμα και στα δύο επίπεδα. Το θέμα αυτό είναι ήδη γνωστό από το χώρο της ρομποτικής, και συγκεκριμένα από την έρευνα στον τηλεχειρισμό. Ένα δεύτερο αντίστοιχο θέμα είναι το πώς θα συνδυαστούν οι φυσικές συμπεριφορές έτσι, ώστε ο πράκτορας να είναι σε θέση να κινείται και να πιάνει αντικείμενα ταυτόχρονα.

Παρόλο που στην έρευνα των φυσικών πρακτόρων υπάρχουν πολλά κοινά ζητήματα με το χώρο της ρομποτικής, υπάρχουν και ορισμένες ενδιαφέρουσες διαφορές. Για παράδειγμα, τα προβλήματα που αφορούν στις δυνάμεις της φυσικής και στα διάφορα υλικά ή ακόμα περισσότερο στο χειρισμό μοχλών και μηχανισμών είναι πολύ πιο δύσκολα και πιο ακριβά στην επίλυσή τους σε σχέση με τους εικονικούς πράκτορες. Έτσι, οι εικονικοί πράκτορες μπορούν να κατασκευαστούν με πολύ πιο ενδιαφέρουσες δομές σε σχέση με τα περισσότερα ρομπότ. Επιπλέον οι εικονικοί αισθητήρες δε χρειάζεται να πάσχουν από τα προβλήματα που υπάρχουν στους αληθινούς. Οι εικονικοί πράκτορες μπορούν να ξέρουν πάντα με ακρίβεια τη θέση τους στον κόσμο, αντίθετα με τα ρομπότ, για τα οποία η εύρεση της θέσης τους είναι ένα μεγάλο πρόβλημα. Από την άλλη μεριά, στον πραγματικό κόσμο η βαρύτητα, η τριβή, η αδράνεια και όλες οι άλλες φυσικές ιδιότητες είναι διαθέσιμες ελεύθερα, ενώ σε έναν εικονικό κόσμο η φυσική θα πρέπει να προστεθεί από το σχεδιαστή. Γνωσιακοί πράκτορες

Τα θέματα που σχετίζονται με το σχεδιασμό και την ανάπτυξη πρακτόρων που βρίσκονται στη γνωσιακή άκρη του φάσματος, μπορούν να ομαδοποιηθούν σε τρεις περιοχές κλειδιά. Πρώτον, θα πρέπει να υπάρχει κάποιο τμήμα του πράκτορα που είναι υπεύθυνο για τις σημαντικές γνωσιακές ικανότητες της συλλογιστικής, ανάληψης αποφάσεων, σχεδιασμού ενεργειών, εκμάθησης, κλπ, ανεξάρτητα από το περιβάλλον στο οποίο βρίσκεται (Funge et al., 1999). Αυτό είναι ακριβώς το ίδιο πρόβλημα που αντιμετωπίζεται από τους ευφυείς πράκτορες και σε άλλες περιπτώσεις, και μπορεί να θεωρηθεί ως η βάση πάνω στην οποία στηρίζονται τα υπόλοιπα τμήματα της αρχιτεκτονικής ενός γνωσιακού πράκτορα.

Το δεύτερο βασικό θέμα αφορά στο ρεαλισμό του πράκτορα στο περιβάλλον του, σε ό,τι έχει όμως να κάνει με τη συμπεριφορά του και όχι με την εμφάνιση και την κίνησή του. Για να είναι πρακτικά τα ευφυή εικονικά περιβάλλοντα, θα πρέπει να 39 είναι αληθοφανή, τόσο σε ό,τι αφορά τις ενέργειες των πρακτόρων, όσο και στις αλληλεπιδράσεις τους με άλλους πράκτορες ή χρήστες. Αυτό πρακτικά σημαίνει ότι το γνωσιακό τμήμα δε θα πρέπει να είναι αποκομμένο από τα στοιχεία που επηρεάζουν τις αποφάσεις, δηλαδή κίνητρα, συναισθήματα, προσωπικότητα, κλπ., τα οποία θα πρέπει να είναι σε θέση να εκφραστούν στο εικονικό περιβάλλον. Έχει γίνει αρκετή έρευνα στο χώρο των ευφύων πρακτόρων με συναισθήματα και σε μοντέλα συναισθημάτων και κινήτρων. Παραδείγματα είναι η εργασία των Moffat και Fridja στην ανάπτυξη ενός υπολογιστικού μοντέλου συναισθημάτων (Moffat et al, 1993; Moffat and Frijda, 1995) και η πιο περίπλοκη «υπολογιστική θεωρία του μυαλού», που

αναπτύχθηκε για πολλά χρόνια από τον Sloman (Sloman and Croucher, 1981; Sloman, 1987; Sloman, 1997).

Το τρίτο θέμα είναι η αναπαράσταση του πράκτορα και των ενεργειών του στον κόσμο, που ολοκληρώνει τον κύκλο, καθώς επιστρέφει στο θέμα της μοντελοποίησης γνωσιακών και συναισθηματικών μοντέλων σε φυσικά (Miranda et al., 2001). Η οπτικοποίηση των πρακτόρων είναι σημαντική και μπορεί να είναι αποδοτική μόνο, εφόσον τα μοντέλα πρακτόρων παρέχουν αρκετά λεπτομερείς πληροφορίες. Για παράδειγμα οι Badler et al (1997) περιγράφουν τους τρόπους με τους οποίους η προσωπικότητα, που εκδηλώνεται με διάφορους τρόπους, όπως για παράδειγμα η περιέργεια ή η κούραση, μπορεί να συσχετιστεί με διαφόρους παραμέτρους της κίνησης ενός εικονικού πράκτορα, όπως ταχύτητα και πρόβλεψη. Περιγράφουν ένα απλό μοντέλο στο οποίο η προσωπικότητα μπορεί να επηρεάσει τις ενέργειες του πράκτορα. Ενώ το ίδιο το μοντέλο είναι σχετικά απλό και είναι μόνο μια αφορμή για μελλοντική εξερεύνηση του πεδίου αυτού, είναι παράλληλα και ένα παράδειγμα σύνδεσης των υψηλού επιπέδου νοητικών τμημάτων του πράκτορα με τη χαμηλού επιπέδου φυσική έκφρασή του. 3.2 Μοντελοποίηση Εικονικών Πρακτόρων

Έχοντας ήδη αναφερθεί σε γενικές γραμμές στα θέματα που προκύπτουν κατά την ανάπτυξη εικονικών πρακτόρων τόσο σε γνωσιακό όσο και σε φυσικό επίπεδο, θα περιγραφεί ενιαία και με περισσότερη λεπτομέρεια η έρευνα σε όλα τα στάδια μοντελοποίησης ενός εικονικού πράκτορα, από τη χαμηλού επιπέδου κίνησή του μέχρι την υψηλού επιπέδου συμπεριφορά του. Όπως και στη γενικότερη περίπτωση των εικονικών περιβαλλόντων, έτσι και στο χώρο των εικονικών πρακτόρων δεν 40 υπάρχουν κάποιες γενικού τύπου μεθοδολογίες για την ανάπτυξή τους. Έχουν προταθεί πολλά διαφορετικά συστήματα και προσεγγίσεις, καθένα από τα οποία διαφέρει σε εμφάνιση, λειτουργία και αυτονομία ανάλογα με το πεδίο εφαρμογής και την απαιτούμενη λεπτομέρεια και ακρίβεια. Σε ό,τι αφορά τους εικονικούς ανθρώπους, ο Norman Badler, ένας από τους σημαντικούς ερευνητές του χώρου, προτείνει τρία διαφορετικά στάδια για την κατασκευή τους (Badler et al., 1999):

- μοντελοποίηση σώματος: η οπτικοποίηση του συνόλου του ανθρώπινου σώματος και ο ορισμός του σκελετού με τις κατάλληλες αρθρώσεις
- αλληλεπίδραση με το χώρο: η απευθείας μετακίνηση των μελών του σώματος ή η δημιουργία κίνησης προσανατολισμένης στο στόχο με τη χρήση περιορισμών και αλγορίθμων αντίστροφης κινηματικής
- έλεγχος συμπεριφοράς: η δυνατότητα δημιουργίας σύνθετης μετακίνησης ακολουθώντας ένα απλό σύνολο κανόνων που καθορίζει τη συμπεριφορά του μοντέλου

Αν και τα στάδια αυτά αναφέρονται συγκεκριμένα σε ανθρώπινα σώματα, βρίσκουν εφαρμογή και στη γενικότερη περίπτωση των εικονικών πρακτόρων.

Ακολουθώντας τα παραπάνω στάδια θα γίνει αναφορά στα ζητήματα της απεικόνισης και κίνησης των εικονικών πρακτόρων, στην αλληλεπίδρασή τους με το περιβάλλον και στους διάφορους τρόπους μοντελοποίησης της συμπεριφοράς τους.

5.5.2 ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΟΥ ΣΩΜΑΤΟΣ

Το πιο βασικό στάδιο μοντελοποίησης του σώματος ενός εικονικού πράκτορα είναι η γεωμετρική αναπαράστασή του. Αυτό το στάδιο μπορεί να είναι το μοναδικό που χρειάζεται, για να παραχθούν στατικά αντικείμενα, αλλά είναι αδύνατο να δημιουργηθεί σύνθετη μετακίνηση του σώματος με βάση μόνο το τρισδιάστατο μοντέλο του. Για το λόγο αυτό ένα ακόμη σημαντικό στάδιο είναι η μοντελοποίηση του σκελετού, ο οποίος θα ορίσει τα κινούμενα μέρη του σώματος και τον τύπο της μετακίνησης που μπορούν να εκτελέσουν. Τα δύο αυτά στάδια είναι αρκετά, για να είναι σε θέση ο χαρακτήρας να εκτελέσει συνθετικές κινήσεις. Παρόλα αυτά, η παραγωγή φυσικής και αληθοφανούς μετακίνησης απαιτεί επιπλέον εργασία. Τόσο το δέρμα όσο και τα ρούχα του εικονικού ανθρώπου θα πρέπει να μετακινούνται και να παραμορφώνονται με φυσικό τρόπο. Ο υπολογισμός της κίνησης του δέρματος και

των υφασμάτων είναι η πιο απαιτητική υπολογιστικά διαδικασία και για το λόγο αυτό δεν ενδείκνυται για συνθετική κίνηση σε πραγματικό χρόνο, ιδιαίτερα σε περιβάλλοντα που περιέχουν περισσότερους από έναν πράκτορες.



Εικόνα 4.3: Στιγμιότυπο από το πρόγραμμα Poser

Η αναπαράσταση του σώματος γίνεται με τη διαδικασία της τρισδιάστατης μοντελοποίησης, όπως και σε οποιοδήποτε άλλο τρισδιάστατο αντικείμενο. Υπάρχουν πολλοί τρόποι για να αναπαραστήσει κάποιος γεωμετρικά μια φιγούρα (καμπύλες επιφάνειες, voxels κ.α.), αλλά ο πιο αποδοτικός τρόπος, ειδικά για εφαρμογές σε πραγματικό χρόνο, είναι να μοντελοποιηθεί το σώμα ως πλέγμα (mesh), δηλαδή ως σύνολο από πολύγωνα που συνδέονται με κοινές ακμές. Το επίπεδο λεπτομέρειας και πολυπλοκότητας του σώματος εξαρτάται από το πεδίο εφαρμογής. Τα παιχνίδια και οι προσομοιώσεις σε πραγματικό χρόνο τείνουν να χρησιμοποιούν μοντέλα με σχετικά μικρό αριθμό πολυγώνων και χρησιμοποιούν εναλλακτικές μεθόδους (π.χ. υφές), για να αναπαραστήσουν λεπτομέρειες. Από την άλλη μεριά, για την παραγωγή συνθετικής κίνησης υψηλής ποιότητας το σώμα μοντελοποιείται με τη μέγιστη δυνατή λεπτομέρεια. Μια καλή τεχνική για το συμβιβασμό μεταξύ λεπτομέρειας και απόδοσης είναι η χρήση επιπέδων λεπτομέρειας (Aubel et al, 2000), δηλαδή η χρήση διαφορετικών μοντέλων λεπτομέρειας που εναλλάσσονται ανάλογα με την απόσταση του θεατή από αυτά.

Υπάρχουν ορισμένα εμπορικά πακέτα που επιτρέπουν στο χρήστη να ορίσει και να κατασκευάσει ένα μοντέλο ανθρώπινου σώματος και να το εξαγάγει σε κάποιο αρχείο για τρισδιάστατα αντικείμενα. Πιθανώς το πιο διαδεδομένο είναι το Poser (εικ. 4.3) της εταιρίας Curiouslabs, με το οποίο ο χρήστης μπορεί να επιλέξει κάποιο βασικό μοντέλο από μια βιβλιοθήκη, να ορίσει έναν αριθμό χαρακτηριστικών (όπως ύψος, όγκος, χαρακτηριστικά προσώπου) και να χρησιμοποιήσει το μοντέλο αυτό στα δικά του προγράμματα ή για παραγωγή βίντεο.

Μοντελοποίηση σκελετού

Η σωστή μετακίνηση του σώματος απαιτεί να υπάρχει η μοντελοποίηση ενός σκελετού που θα ορίζει τα κινούμενα μέρη της φιγούρας. Στην περίπτωση των εικονικών ανθρώπων, παρόλο που είναι δυνατό να μοντελοποιηθεί κάθε κόκαλο του σκελετού και να κωδικοποιηθεί ο τρόπος με τον οποίον κινείται σε σχέση με τα υπόλοιπα, συνήθως είναι αρκετό να μοντελοποιηθούν τα τμήματα (segments) του σώματος με βάση τη θέση και τις διαστάσεις τους και οι αρθρώσεις (joints) με βάση τις περιστροφές που μπορούν να εκτελέσουν. Οι αρθρώσεις είναι συνήθως περιστροφικές, αλλά σε κάποιες περιπτώσεις μπορούν και να μετακινούνται. Κάθε περιστροφική άρθρωση μπορεί να επιτρέπει περιστροφή σε έναν, δύο ή τρεις ορθογώνιους άξονες, που είναι οι βαθμοί ελευθερίας της άρθρωσης. Μια λεπτομερής προσέγγιση του ανθρώπινου σκελετού μπορεί να έχει γύρω στους διακόσιους βαθμούς ελευθερίας, παρόλο που συνήθως ένας μικρότερος αριθμός είναι αρκετός. Το επιτρεπτό εύρος κινήσεων κάθε άρθρωσης μπορεί να περιοριστεί με τον ορισμό ελάχιστης και μέγιστης γωνίας για κάθε περιστροφή.

Τα επιμέρους αντικείμενα που απαρτίζουν το σκελετό ορίζονται το καθένα στο δικό τους τοπικό σύστημα συντεταγμένων και δημιουργούν μια αναγνωρίσιμη φιγούρα σε ένα γενικό σύστημα συντεταγμένων του κόσμου με μια σειρά από ένθετους (nested)

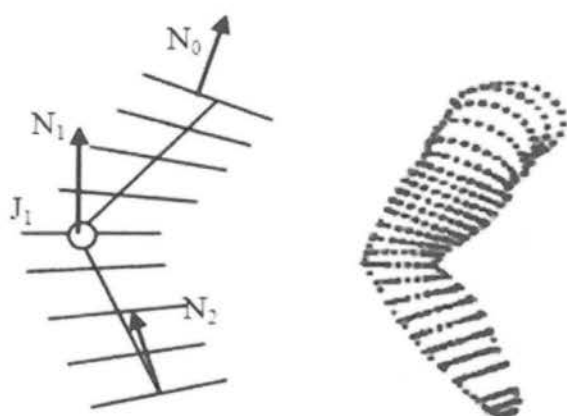
μετασχηματισμούς. Το γεγονός ότι η αλλαγή στη γωνία περιστροφής μιας άρθρωσης επηρεάζει τη θέση αρκετών άλλων καθιστά αναγκαία τη σύνδεσή τους σε τοπολογία δέντρου (Badler et al., 1993b). Η τοπολογία αυτή έχει ως συνέπεια η μεταβολή της περιστροφής ενός κόμβου να προκαλεί γεωμετρικούς μετασχηματισμούς σε όλους τους απογονικούς κόμβους.

Πρόσφατα έχει γίνει μια προσπάθεια να οριστεί μια πρότυπος ιεραρχία αρθρώσεων για ανθρώπινα μοντέλα στη γλώσσα VRML97. Η ομάδα εργασίας Humanoid Animation Working Group πρότεινε τις προδιαγραφές για ένα VRML ανθρωποειδές (h-anim), που περιλαμβάνει ένα πολύ λεπτομερές δέντρο αρθρώσεων. Ο πρωτεύων στόχος της ομάδας εργασίας είναι να ενθαρρύνει τους δημιουργούς να κατασκευάσουν μοντέλα σκελετού με ένα δεδομένο τρόπο, ώστε να μπορούν να χρησιμοποιηθούν σε διαφορετικές εφαρμογές. Μοντελοποίηση του δέρματος και των ρούχων

Η μοντελοποίηση του ανθρώπινου δέρματος θα πρέπει να γίνει μετά από προσεκτική μελέτη, γιατί, αντίθετα με τα μηχανικά αντικείμενα, το δέρμα παραμορφώνεται, όταν το σώμα κινείται. Για παράδειγμα, όταν οι ρόδες ενός αυτοκινήτου περιστρέφονται, η περιστροφή αυτή δεν επηρεάζει οποιοδήποτε άλλο τμήμα του, αλλά, όταν το ανθρώπινο χέρι κινείται, προκαλεί παραμόρφωση του δέρματος γύρω από τον ώμο και το στήθος. Εάν τα μέλη του σώματος έχουν μοντελοποιηθεί ως στερεά αντικείμενα, ακόμα και μικρές περιστροφές μπορούν να προκαλέσουν ανεπιθύμητες «διασπάσεις» στο δέρμα και να κάνουν το μοντέλο να μη δείχνει ρεαλιστικό.

Ο απλούστερος και ευκολότερος τρόπος, για να αποφευχθεί το πρόβλημα αυτό, είναι να χρησιμοποιηθούν σταθερά αντικείμενα στη θέση των αρθρώσεων. Τα αντικείμενα αυτά (συνήθως σφαίρες ή ελλειψοειδή) εξυπηρετούν ως ένα μέσο για να γεμίσουν οι διασπάσεις, όταν τα τμήματα μετακινούνται, δίνοντας έτσι την ψευδαίσθηση συνεχούς δέρματος. Τα οπτικά αποτελέσματα της τεχνικής αυτής δεν είναι ιδιαίτερα κομψά, αλλά λόγω της απλότητάς της έχει υιοθετηθεί από πολλά συστήματα πραγματικού χρόνου, κυρίως σε περιπτώσεις που δεν είναι διαθέσιμη μεγάλη υπολογιστική ισχύς.

Μια ακόμη πιο αποδοτική λύση είναι η παραμόρφωση με το χειρισμό περιγραμμάτων δέρματος (Kalra et al., 1998). Η τεχνική αυτή δεν αντιμετωπίζει τα τμήματα του σώματος ως στατικά πλέγματα, αλλά παραμορφώνει τη γεωμετρία του σώματος, όταν ένα τμήμα περιστρέφεται. Η ιδέα είναι να μεταβληθούν τα παράλληλα περιγράμματα (σύνολο σημείων που βρίσκονται στο ίδιο επίπεδο) του μοντέλου του σώματος. Κάθε άρθρωση θα πρέπει να αντιστοιχεί σε ένα τέτοιο περίγραμμα και να ανήκει στο επίπεδο του περιγράμματος αυτού. Όταν ένα τμήμα περιστρέφεται γύρω από μια άρθρωση, η παραμόρφωση του δέρματος γύρω από την άρθρωση αυτή υπολογίζεται από τις γειτονικές αρθρώσεις (τη γονική και τις απογονικές στην ιεραρχία αρθρώσεων). Εάν το περίγραμμα της άρθρωσης J_1 , που προκάλεσε την περιστροφή,



Εικόνα 4.4 : Παραμόρφωση δέρματος με χειρισμό περιγραμμάτων

έχει κάθετο διάνυσμα N_1 (το κάθετο διάνυσμα του επιπέδου στο οποίο βρίσκεται το περίγραμμα) και οι δύο γειτονικές αρθρώσεις J_0 και J_2 έχουν αντίστοιχα διανύσματα N_0 και N_2 , τα ενδιάμεσα περιγράμματα στα τμήματα $J_0 J_1$ και $J_1 J_2$ υπολογίζονται με τη χρήση γραμμικής παρεμβολής μεταξύ των κάθετων διανυσμάτων (εικόνα 3.2). Μόλις υπολογιστεί το κάθετο διάνυσμα ενός περιγράμματος, όλες οι κορυφές του μετασχηματίζονται έτσι, ώστε να ανήκουν στο νέο επίπεδο. Η τεχνική αυτή είναι γρήγορη και αποδοτική, αλλά δεν μπορεί να εφαρμοστεί σε όλα τα τρισδιάστατα μοντέλα σώματος. Το σώμα θα πρέπει να έχει κατασκευαστεί με τη χρήση περιγραμμάτων που ανήκουν σε επίπεδα που είναι κάθετα στα τμήματα του σκελετού.

Στις περιπτώσεις εκείνες που απαιτείται υψηλό επίπεδο ρεαλισμού, η κίνηση του δέρματος θα πρέπει να ακολουθεί τους φυσικούς νόμους, θα πρέπει δηλαδή να συμπεριφέρεται ως ελαστική επιφάνεια. Ένας τρόπος για να επιτευχθεί αυτό είναι να μοντελοποιηθεί το δέρμα ως ένα σύνολο από ελατήρια και μάζες (Vince, 1995). Κάθε κορυφή του πλέγματος μπορεί να ληφθεί ως μια μικρή μάζα και κάθε ακμή ως ένα ελατήριο με σχετικά μικρή ελαστικότητα. Η δομή αυτή εγγυάται την ελαστική παραμόρφωση του δέρματος, καθώς το σώμα κινείται. Παρόλα αυτά απαιτείται το πλέγμα της επιφάνειας να είναι αρκετά πυκνό και σχετικά κανονικό.

Υπάρχουν επίσης και πιο πολύπλοκες προσεγγίσεις που χρησιμοποιούν βιολογικά και μηχανικά μοντέλα, για να υπολογίσουν τη συμπεριφορά του δέρματος και να δημιουργήσουν ρυτίδες, όταν το δέρμα παραμορφώνεται (Wu et al., 1997 Magnetat Thalmann et al., 1996). Τα μοντέλα αυτά λαμβάνουν υπόψη τους και τη δομή των μυών του σώματος και υπολογίζουν και τη δική τους παραμόρφωση.

Η προσομοίωση του υφάσματος ακολουθεί σχεδόν την ίδια διαδικασία χρησιμοποιώντας και πάλι ένα μοντέλο με ελατήρια και μάζες. Η επιφάνεια των

ρούχων πρέπει να είναι ένα πυκνό σύνολο από μικρές μάζες, που θα πρέπει να έχουν βαρύτητα και τη δυνατότητα να συγκρούονται τόσο μεταξύ τους όσο και με το δέρμα. Όσο πιο λεπτομερές είναι το μοντέλο (μεγάλος αριθμός ελατηρίων και μαζών) τόσο πιο ρεαλιστική είναι η παραμόρφωση και η δημιουργία πτυχώσεων στα ρούχα. Η σημερινή υπολογιστική ισχύς επιτρέπει συνθετική κίνηση ρούχων σε πραγματικό χρόνο με την χρήση ενός απλουστευμένου μοντέλου ελατηρίων - μαζών (Lander, 1999). 3.2.2 Στάσεις του σώματος: Ευθεία και Αντίστροφη Κινηματική

Μια στάση του σώματος μπορεί να δημιουργηθεί με τη μεταβολή των τοπικών περιστροφών στις αρθρώσεις καθώς και με τη γενική μετατόπιση της κορυφής του δέντρου αρθρώσεων. Υπάρχουν δύο θεμελιώδεις προσεγγίσεις σε ό,τι αφορά τον ορισμό μιας στάσης του σώματος: η χαμηλού επιπέδου ευθεία κινηματική (forward kinematics), και η πιο κομψή, αντίστροφη κινηματική (inverse kinematics).

Στην ευθεία κινηματική θα πρέπει να ορίσει κάποιος επακριβώς τη θέση και τον προσανατολισμό των αντικειμένων σε συγκεκριμένους χρόνους. Για δομές που περιλαμβάνουν σκελετό (όπως ο άνθρωπος), αυτό σημαίνει να ορίσει κάποιος απευθείας τις περιστροφές σε επιλεγμένες αρθρώσεις και πιθανώς τη γενική μετακίνηση του αντικειμένου, ώστε να δημιουργήσει μια στάση. Χρησιμοποιώντας ευθεία κινηματική, η θέση κάθε τμήματος σε ένα σκελετό μπορεί να ελεγχθεί μόνο έμμεσα μεταβάλλοντας τις περιστροφές μεταξύ των αρθρώσεων από την κορυφή του δέντρου μέχρι το συγκεκριμένο τμήμα.

Η αντίστροφη κινηματική, που πολλές φορές καλείται κίνηση προσανατολισμένη στο στόχο, παρέχει μια ενδιαφέρουσα εναλλακτική λύση σε σχέση με την περιστροφή των επιμέρους αρθρώσεων σε ένα σκελετό. Ο δημιουργός μπορεί να ορίσει απευθείας τη θέση ενός τελικού στοιχείου δράσης (end-effector), ενώ το σύστημα υπολογίζει αυτόματα τις γωνίες των αρθρώσεων που απαιτούνται, για να τοποθετηθεί το τμήμα αυτό στην κατάλληλη θέση (Watt and Watt, 1992). Το πρόβλημα της αντίστροφης κινηματικής έχει μελετηθεί εκτενώς στο χώρο της ρομποτικής, ενώ μόλις πρόσφατα εφαρμόστηκαν οι τεχνικές αυτές στη συνθετική κίνηση με υπολογιστή.

Στην περίπτωση της ευθείας κινηματικής ο δημιουργός έχει πάρα πολλούς μετασχηματισμούς να ελέγξει, οι οποίοι, ενώ του δίνουν παραπάνω ελευθερία για πιο εκφραστικές κινήσεις, μπορούν να αποδειχτούν ιδιαίτερα πολύπλοκοι στην πράξη. Από την άλλη μεριά, οι αλγόριθμοι ανάστροφης κινηματικής έχουν μεγάλο υπολογιστικό κόστος καθιστώντας απίθανο τον υπολογισμό των γωνιών των αρθρώσεων για όλους τους βαθμούς ελευθερίας ενός εικονικού ανθρώπου σε πραγματικό χρόνο. Στις περισσότερες περιπτώσεις απαιτείται μια ισορροπία μεταξύ των δύο αυτών προσεγγίσεων. 3.2.3 Κίνηση

Οι άνθρωποι είναι ιδιαίτερα ικανοί στο να αντιλαμβάνονται τις μικρές λεπτομέρειες της ανθρώπινης κίνησης. Για παράδειγμα, κάποιος μπορεί συχνά να αναγνωρίζει τους γνωστούς του από μεγάλη απόσταση μόνο και μόνο από το βάδισμά

τους. Λόγω της ικανότητας αυτής, είναι ιδιαίτερα υψηλές οι απαιτήσεις από συνθετικές κινήσεις που περιλαμβάνουν ανθρώπους ή γενικότερα συνθετικούς χαρακτήρες. Για να είναι ρεαλιστική και ενδιαφέρουσα η κίνηση που δημιουργείται από υπολογιστή, θα πρέπει οι εικονικοί χαρακτήρες να κινούνται με φυσικό τρόπο.

Ο ορισμός της κίνησης σε υπολογιστή είναι εξαιρετικά δύσκολος. Ακόμα και μια απλή μπάλα που αναπηδά είναι δύσκολο να αναπαρασταθεί πειστικά, κυρίως επειδή οι άνθρωποι εύκολα ξεχωρίζουν ενέργειες που δείχνουν λανθασμένες και αφύσικες, χωρίς κατ' ανάγκη να γνωρίζουν τι ακριβώς φταίει. Η συνθετική κίνηση ενός χαρακτήρα είναι εξαιρετικά επίπονη διαδικασία, διότι πρέπει να οριστούν πολλές λεπτομέρειες της κίνησης, για να αποδώσουν την προσωπικότητα και τη διάθεση.

Οι τεχνικές συνθετικής κίνησης σε υπολογιστή κατατάσσονται σε τρεις βασικές κατηγορίες: διαδικαστική συνθετική κίνηση (procedural animation), συνθετική κίνηση απόδοσης (performance animation) και συνθετική κίνηση συμπεριφοράς (behavioral animation). Και οι τρεις περιλαμβάνουν ένα συμβιβασμό μεταξύ του επιπέδου ελέγχου που έχει ο δημιουργός στις μικρές λεπτομέρειες της κίνησης και της ποσότητας εργασίας που εκτελεί ο υπολογιστής μόνος του. Η τεχνική της κίνησης με θέσεις-κλειδιά που χρησιμοποιείται στη διαδικαστική συνθετική κίνηση επιτρέπει 47

λεπτομερή έλεγχο, αλλά ο δημιουργός πρέπει να σιγουρέψει μόνος του τη φυσικότητα του αποτελέσματος. Αντίθετα, η συνθετική κίνηση απόδοσης και η συνθετική κίνηση συμπεριφοράς παράγουν κίνηση με πιο αυτοματοποιημένο τρόπο, αλλά παρέχουν λίγες δυνατότητες για διόρθωση εκ των υστέρων. Διαδικαστική συνθετική κίνηση

Μεγάλο μέρος της ώθησης στην τεχνολογία των φυσικών πρακτόρων οφείλεται στην παραγωγή φιλμ. Η συνθετική κίνηση με το χέρι είναι μια ιδιαίτερα χρονοβόρα και ακριβή διαδικασία, με αποτέλεσμα να φαίνεται προφανής ιδέα η χρήση υπολογιστικής ισχύος. Ο πιο απλός αυτοματισμός είναι η παραγωγή ενδιάμεσων σταδίων μεταξύ των εικόνων που είχαν ζωγραφιστεί με το χέρι, έτσι ώστε να χρειάζεται σημαντικά μικρότερος αριθμός σχεδίων. Η τεχνική αυτή συχνά ονομάζεται διαδικαστική συνθετική κίνηση.

Η διαδικασία παραγωγής διαδικαστικής συνθετικής κίνησης ονομάζεται κίνηση με θέσεις-κλειδιά (keyframing) και απαιτεί από το δημιουργό να ορίσει κρίσιμες θέσεις για τα μέρη του σώματος του χαρακτήρα ανά τακτά χρονικά διαστήματα. Ο υπολογιστής στη συνέχεια «γεμίζει» τις ενδιάμεσες σκηνές με ομαλή παρεμβολή μεταξύ των θέσεων αυτών.

Ο ορισμός των θέσεων κλειδιών μπορεί να είναι σχετικά αυτοματοποιημένος με τεχνικές που βοηθούν στο μετασχηματισμό ορισμένων αρθρώσεων. Εάν για παράδειγμα το χέρι ενός χαρακτήρα πρέπει να βρίσκεται σε μια συγκεκριμένη θέση, ο υπολογιστής θα μπορούσε να υπολογίσει τις κατάλληλες γωνίες του ώμου και του αγκώνα χρησιμοποιώντας αντίστροφη κινηματική. Παρόλο που τέτοιες τεχνικές

διευκολύνουν τη διαδικασία, ο ορισμός στιγμιότυπων απαιτεί από το δημιουργό να έχει λεπτομερή κατανόηση του τρόπου με τον οποίο τα κινούμενα μέρη θα πρέπει να συμπεριφέρονται στο χρόνο, καθώς και το ταλέντο να εκφράσει την πληροφορία αυτή μέσα από θέσεις κλειδιά. Το γεγονός, ότι η τεχνική του ορισμού στιγμιότυπων εξακολουθεί και είναι δημοφιλής, οφείλεται καθαρά στο βαθμό ελέγχου που δίνει πάνω στις λεπτομέρειες της κίνησης.

Μία επέκταση της ιδέας αυτής βρίσκεται στο προγραμματιστικό περιβάλλον Microsoft AGENT, που έχει προκύψει από το πρόγραμμα Persona (Ball et al, 1997), και αφορά χαρακτήρες σε δύο διαστάσεις. Στην περίπτωση αυτή παρέχονται προσχεδιασμένα τμήματα συνθετικής κίνησης, (π.χ. ο χαρακτήρας κοιτάζει αριστερά, ο χαρακτήρας εξαφανίζεται και ο χαρακτήρας πετάει ψηλά) και ο προγραμματιστής μπορεί να δηλώσει ακολουθίες τέτοιων κινήσεων, τις οποίες το σύστημα παράγει με ομαλή μετάβαση μεταξύ των επιμέρους συνθετικών κινήσεων. Μια παρόμοια προσέγγιση σε τρεις διαστάσεις υπάρχει στο σύστημα εικονικών ηθοποιών IMPROV (Goldberg, 1997), που ορίζει ως ενέργεια μια απλή ή επαναλαμβανόμενη διαδικασία που δεν απαιτεί υψηλού επιπέδου επίγνωση ή συνειδητές αποφάσεις. Στο IMPROV οι ενέργειες χωρίζονται σε κατηγορίες, όπου ενέργειες της ίδιας κατηγορίας είναι αμοιβαίως αποκλειόμενες. Αυτό επιτρέπει διαφορετικές ενέργειες να εκτελούνται ταυτόχρονα για παράδειγμα, ένας εικονικός ηθοποιός να περπατάει και να χαιρετάει.

Ο προσδιορισμός των ακολουθιών μπορεί να είναι ιδιαίτερα πολύπλοκος. Το IMPROV παρέχει ένα μηχανισμό γλώσσας σεναρίων με τη δυνατότητα να εκτελούνται σενάρια μέσα από άλλα σενάρια ή να καλούνται από ενέργειες του χρήστη. Στο πρόγραμμα Persona, το πρόβλημα αυτό αντιμετωπίστηκε με τη χρήση ενός γραμμικού σχεδιαστή ενεργειών (planner) σε κατάσταση off-line, ο οποίος παρήγαγε αρχικά τις ακολουθίες ως πλάνα που εκτελούνταν στη διάρκεια του προγράμματος. Συνθετική κίνηση απόδοσης

Μια δεύτερη προσέγγιση που χρησιμοποιείται στη συνθετική κίνηση αλλά και στο χειρισμό των ενσαρκώσεων (Emerging et al, 1997), είναι να προσδιορίζονται οι κινήσεις του πράκτορα από έναν πραγματικό άνθρωπο μια προσέγγιση που είναι γνωστή ως συνθετική κίνηση απόδοσης.

Η συνθετική κίνηση απόδοσης επιτυγχάνεται με μια τεχνική που ονομάζεται σύλληψη κίνησης (motion capture) και περιλαμβάνει τη μέτρηση της θέσης και του προσανατολισμού αντικειμένων του φυσικού κόσμου και την καταγραφή της πληροφορίας αυτής σε μορφή χρησιμοποιήσιμη από υπολογιστή. Εφόσον τα δεδομένα καταγραφούν, ο υπολογιστής τα αξιοποιεί σε αντίστοιχα αντικείμενα στο εικονικό περιβάλλον. Στην περίπτωση των εικονικών ανθρώπων η συνθετική κίνηση που παράγεται χρησιμοποιεί τις καταγεγραμμένες θέσεις και προσανατολισμούς των μελών του σώματος, ώστε να ορίσει τα στιγμιότυπα που θα χρησιμοποιηθούν στη συνέχεια στο τρισδιάστατο σκηνικό.

Η σύλληψη κίνησης έχει αποδειχτεί μια πολύ σημαντική τεχνική για τη συνθετική κίνηση ανθρώπων και ανθρωποειδών χαρακτήρων. Τα δεδομένα σύλληψης κίνησης κρατούν πολλές από τις μικρές λεπτομέρειες που αντικατοπτρίζουν το στυλ του ηθοποιού κάνοντας δυνατές τις ψηφιακές παραστάσεις, όπου η ιδιαίτερη προσωπικότητα του ανθρώπου είναι αναγνωρίσιμη στο τελικό προϊόν. Επειδή η βασική κίνηση ορίζεται σε πραγματικό χρόνο από το υποκείμενο του οποίου η θέση συλλαμβάνεται, η σύλληψη κίνησης παρέχει μια ισχυρή λύση για εφαρμογές, όπου συνθετική κίνηση με τη χαρακτηριστική ποιότητα της ανθρώπινης κίνησης πρέπει να δημιουργηθεί σε σύντομο χρονικό διάστημα. Υπάρχουν τρεις διαφορετικές τεχνικές που χρησιμοποιούνται σήμερα για να καταγράψουν την κίνηση του σώματος: τα μαγνητικά συστήματα, τα οπτικά συστήματα, και τα ψηφιακά ανδρείκελα (Dyer et al., 1995). Συνθετική κίνηση συμπεριφοράς

Εάν ο πράκτορας έχει ρεαλιστική δομή του σώματος, ένα προφανές βήμα είναι να παρέχεται αυτόνομος έλεγχος σε αυτό, έτσι ώστε η κίνηση του πράκτορα να ελέγχεται από εσωτερική ενεργοποίηση της δομής του σώματος αντί για εξωτερική συνθετική κίνηση της επιφάνειάς του. Η τεχνική αυτή περιγράφηκε από το Reynolds (1987) ως συνθετική κίνηση συμπεριφοράς. Παρόλα αυτά, ανάλογα με τη φυσική ακρίβεια αυτού του ελέγχου και τη δομική πολυπλοκότητα του σώματος η τεχνική αυτή μπορεί να απαιτήσει πολλή υπολογιστική ισχύ.

Αντίθετα με τις δύο παραπάνω τεχνικές, η συνθετική κίνηση συμπεριφοράς προσομοιώνει σε ένα βαθμό τους νόμους της φυσικής, για να δημιουργήσει την κίνηση του σώματος. Οι εικονικοί χαρακτήρες αναπαρίστανται συνήθως ως ένα σύνολο από στερεά μέλη του σώματος. Παρόλο που τα μοντέλα αυτά είναι αληθοφανή, αποτελούν μια απλοϊκή προσέγγιση του ανθρώπινου σώματος, καθώς αγνοούν την κίνηση των μυών σε σχέση με τα κόκαλα. Πρόσφατα έχουν αρχίσει να κατασκευάζονται πιο περίπλοκα μοντέλα και οι προκύπτουσες προσομοιώσεις γίνονται όλο και πιο ζωντανές, όσο περισσότερη λεπτομέρεια προστίθεται (Hodgins and Wooten, 1998; Kalra et al., 1998).

Όταν τα μοντέλα είναι άψυχα αντικείμενα, όπως ενδύματα ή νερό, ο υπολογιστής μπορεί να καθορίσει τις κινήσεις τους κάνοντάς τα να υπακούουν στους νόμους της κίνησης, όπως ορίζονται στους φυσικούς νόμους. Στην περίπτωση μιας μπάλας που κυλάει από ένα λόφο, η προσομοίωση θα μπορούσε να υπολογίσει την κίνηση λαμβάνοντας υπόψη τη βαρύτητα και δυνάμεις όπως η τριβή που δημιουργείται από την επαφή μεταξύ της μπάλας και του εδάφους. Αλλά οι άνθρωποι έχουν εσωτερικές πηγές ενέργειας και δεν είναι παθητικά, άψυχα αντικείμενα. Για το λόγο αυτό οι εικονικοί άνθρωποι απαιτούν μια πηγή που θα παράγει δυνάμεις στους μύες, ένα «σύστημα ελέγχου». Το τμήμα αυτό λογισμικού υπολογίζει και εφαρμόζει ροπές σε κάθε κλείδωση του προσομοιωμένου σώματος, ώστε να μπορέσει ο χαρακτήρας να εκτελέσει την επιθυμητή ενέργεια. Για παράδειγμα, ένα σύστημα ελέγχου για τρέξιμο θα πρέπει να υπολογίζει τις ροπές, για να κινηθεί προς τα εμπρός το πόδι, πριν ακουμπήσει στο έδαφος, ώστε να αποτρέψει το δρομέα από το να

σκοντάψει (Hodgins and Wooten, 1998). Τα περισσότερα συστήματα ελέγχου χρησιμοποιούν μηχανές κατάστασης (state machines): αλγορίθμους που υπολογίζουν τι θα πρέπει να κάνει κάθε άρθρωση την κάθε στιγμή και στη συνέχεια επιβεβαιώνουν ότι οι κλειδώσεις εκτελούν τις πράξεις αυτές σε κατάλληλους χρόνους (Badler and Webber, 1995; Hodgins, 1996). Το τρέξιμο, για παράδειγμα, είναι μια κυκλική διαδικασία που μεταβάλλεται μεταξύ μιας φάσης, όπου το ένα πόδι παρέχει στήριξη, και μιας άλλης, όπου και τα δύο είναι στον αέρα. Κατά τη διάρκεια της πρώτης φάσης ο αστράγαλος, το γόνατο και ο γοφός του ποδιού που είναι σε επαφή με το έδαφος, πρέπει να παρέχουν στήριξη και ισορροπία. Όταν όμως αυτό το πόδι είναι στον αέρα, ο γοφός έχει μια διαφορετική λειτουργία, αυτή της προώθησης του ποδιού προς τα εμπρός, για να προετοιμαστεί για την επόμενη επαφή με το έδαφος. Μια μηχανή κατάστασης επιλέγει μεταξύ των διαφόρων ρόλων του γοφού και επιλέγει την κατάλληλη ενέργεια για την τρέχουσα φάση της κίνησης. Σε κάθε φάση υπάρχουν οι αντίστοιχοι «νόμοι ελέγχου» που υπολογίζουν τις επιθυμητές γωνίες για κάθε άρθρωση του προσομοιωμένου ανθρώπινου σώματος. Οι νόμοι ελέγχου είναι εξισώσεις που αναπαριστούν πώς θα έπρεπε να κινείται κάθε μέλος του σώματος, για να εκπληρώσει τη λειτουργία του σε κάθε φάση της κίνησης. Για να μετακινηθούν οι αρθρώσεις στις επιθυμητές θέσεις, το σύστημα ελέγχου υπολογίζει τις απαιτούμενες ροπές με εξισώσεις, που επιδρούν ως ελατήρια, τραβώντας τις αρθρώσεις προς τις επιθυμητές γωνίες. Στην ουσία οι εξισώσεις αυτές 51 είναι εικονικοί μύες που κινούν τα διάφορα μέρη του σώματος στις κατάλληλες θέσεις.

5.5.3 ΑΛΛΗΛΕΠΙΔΡΑΣΗ ΜΕ ΤΟ ΠΕΡΙΒΑΛΛΟΝ

Ως τώρα δεν έχουν αναφερθεί θέματα που προκύπτουν από τη δυνατότητα των εικονικών πρακτόρων να αντιλαμβάνονται τον κόσμο και οι πράξεις τους να έχουν μια αντίστοιχη επίδραση στον κόσμο. Αυτές είναι οι περιπτώσεις αλληλεπίδρασης μεταξύ ενός πράκτορα και του κόσμου, που εξαρτώνται όμως σε μεγάλο βαθμό από το ίδιο το μοντέλο του περιβάλλοντος και τη λειτουργικότητά του. Στη συνέχεια θα εξεταστεί το ζήτημα της αυτονομίας εικονικών πρακτόρων, η μοντελοποίηση του περιβάλλοντος και τα θέματα της αντίληψης και ενέργειας των εικονικών πρακτόρων στο περιβάλλον.

Αυτονομία εικονικών πρακτόρων

Η έννοια της αυτονομίας είναι πολύ σημαντική και αφορά στους πράκτορες που πρέπει να λειτουργούν αποτελεσματικά και ανεξάρτητα σε ένα δυναμικό περιβάλλον (Caicedo and Thalmann, 2000; Goncalves and Silva, 2001). Έχει γίνει αρκετή δουλειά σε θέματα αυτονομίας, όπως για παράδειγμα σχετικά με τη φύση της και όσα αυτή συνεπάγεται (Luck and d'Inverno, 1995), με το πώς μπορεί να προσδιοριστεί από αρχιτεκτονικές πρακτόρων (Castelfranchi, 1995) και με τις λεπτομέρειες της χρήσης τέτοιων αρχιτεκτονικών (Franklin and Graesser, 1996). Μια ερώτηση που προκύπτει αυτόματα είναι το κατά πόσο η αυτονομία είναι χρήσιμη και κατάλληλη για πράκτορες σε εικονικά περιβάλλοντα, όσο είναι για πράκτορες στον πραγματικό κόσμο (π.χ. ρομπότ). Στον πραγματικό κόσμο το περιβάλλον λειτουργεί ανεξάρτητα από τους πράκτορες που βρίσκονται μέσα σε αυτό. Ένας μεμονωμένος πράκτορας μπορεί να αντιληφθεί μόνο μέρος του περιβάλλοντος, το οποίο επηρεάζεται τόσο από τις δικές

του ενέργειες, όσο και από αυτές άλλων πρακτόρων. Υπό αυτές τις συνθήκες οι προβλέψεις για την τρέχουσα κατάσταση του κόσμου μπορούν κάλλιστα να είναι λανθασμένες. Η αυτονομία στην περίπτωση αυτή είναι απολύτως απαραίτητη, καθώς ο πράκτορας θα πρέπει να παίρνει αποφάσεις με βάση την τρέχουσα κατάσταση του κόσμου, η οποία δεν μπορεί να είναι εκ των προτέρων προβλεπόμενη.

Σε ένα εικονικό περιβάλλον η κατάσταση είναι αρκετά διαφορετική. Ο σχεδιαστής έχει μια ολοκληρωμένη άποψη τόσο του κόσμου όσο και του πράκτορα, και συνεπώς δεν χρειάζεται να κάνει διάκριση μεταξύ των δύο. Επιπλέον, ολόκληρο το περιβάλλον μπορεί να είναι διαθέσιμο στον πράκτορα - εφόσον δεν υπάρχει λόγος να γίνεται διάκριση μεταξύ του μοντέλου του κόσμου στο μυαλό του πράκτορα και του ίδιου του κόσμου. Από την πρακτική σκοπιά η αυτονομία αντιμετωπίζεται ως μια επιπλέον επιβάρυνση του συστήματος, που είναι χρήσιμη μόνο ως βάση για επιστημονικές έρευνες πάνω σε θεωρίες πρακτόρων.

Παρόλα αυτά, όπως ισχυρίζονται οι Petta και Trapl (1997), η χρήση ενός πράκτορα παντογνώστη μπορεί τελικά να αποδειχτεί ακόμα λιγότερο αποδοτική. Το πρόβλημα είναι ότι, εάν οι εικονικοί πράκτορες πρέπει να συμπεριφέρονται με έναν τρόπο που είναι πειστικός στο χρήστη και διατηρεί το αίσθημα της παρουσίας στον εικονικό κόσμο, θα πρέπει να συμπεριφέρονται σαν να είχαν τους ίδιους περιορισμούς με αυτούς των αντίστοιχων οντοτήτων στον πραγματικό κόσμο. Θα πρέπει να φαίνονται σαν να συλλέγουν πληροφορίες, μόνο όταν αυτές είναι προσβάσιμες, και να αλληλεπιδρούν με τον κόσμο και τα αντικείμενα με λογικοφανή τρόπο. Η χρήση πρακτόρων - παντογνωστών καταλήγει σε προβλήματα, όταν θα πρέπει το σύστημα να καταγράψει το τι υποτίθεται ότι έχει αντιληφθεί και το τι γνωρίζει ο κάθε πράκτορας. Σε τέτοιες περιπτώσεις είναι τελικά πολύ πιο απλό να εφοδιαστεί κάθε πράκτορας με εικονικούς αισθητήρες (Thalmann et al, 1997) και να τους αξιοποιήσει, για να πραγματοποιεί αυτόνομα τις ενέργειές του χρησιμοποιώντας εικονικούς επιδραστής. Έτσι, ένα μεγάλο μέρος της ερευνητικής δουλειάς που έχει γίνει σχετικά με εικονικούς πράκτορες υιοθετεί τελικά την προσέγγιση της αυτονομίας.

Υπάρχουν, όμως, και άλλα πλεονεκτήματα της χρήσης αυτόνομων πρακτόρων, που εντοπίζονται κυρίως στο επίπεδο υλοποίησης, με πιο σημαντικά τις δυνατότητες επαναχρησιμοποίησής τους σε διαφορετικά περιβάλλοντα και κατανομής της λειτουργίας μεμονωμένων πρακτόρων σε διαφορετικούς επεξεργαστές. Τα παραπάνω, βέβαια, είναι σε μεγάλο βαθμό θεωρητικά, καθώς στην πράξη υπάρχουν λίγες ενδείξεις επαναχρησιμοποίησης ή κατανομής, γεγονός που αντανακλά την τρέχουσα ανωριμότητα του πεδίου και την ποικιλία των προβλημάτων που συναντιούνται. Είναι επίσης αλήθεια ότι, ενώ η αυτονομία δείχνει να είναι αναγκαία συνθήκη για επαναχρησιμοποιήσιμους πράκτορες, απέχει πολύ από το να είναι και ικανή, καθώς υπάρχουν πολλά θέματα που πρέπει να λυθούν σε ό,τι αφορά την εννοιολογική αναπαράσταση των πρακτόρων, του περιβάλλοντος και των διαθέσιμων αλληλεπιδράσεων. Μοντελοποίηση του περιβάλλοντος

Τα συνηθισμένα εργαλεία για εικονικά περιβάλλοντα χρησιμοποιούν το γράφημα σκηνικού, για να αναπαραστήσουν τον κόσμο. Αυτός είναι ένας βολικός τρόπος αναπαράστασης της γραφικής άποψης των αντικειμένων, καθώς τα φύλλα στο γράφημα σκηνικού αναπαριστούν συνήθως πρωτογενή αντικείμενα γραφικών ως συλλογή από πολύγωνα. Τα πρωτογενή αυτά αντικείμενα ομαδοποιούνται στη συνέχεια σε πιο πολύπλοκα γραφικά αντικείμενα χρησιμοποιώντας κόμβους ομαδοποίησης στους οποίους ενσωματώνονται.

Παρόλα αυτά, αν θελήσει κάποιος να προσθέσει γνώση στα αντικείμενα, και συγκεκριμένα να μπορούν να τα χειριστούν οι πράκτορες σε επίπεδο γνώσης (Levison and Badler, 1994), η αναπαράσταση γραφήματος σκηνικού δεν είναι ιδιαίτερα βολική, καθώς δεν είναι πάντα ξεκάθαρος ο τρόπος με τον οποίο ένα εννοιολογικό αντικείμενο αντιστοιχίζεται σε συλλογή από πρωτογενή γραφικά αντικείμενα. Κατά συνέπεια, υπάρχει σήμερα η άποψη ότι είναι απαραίτητο οι σχεδιαστές εργαλείων για εικονικά περιβάλλοντα να αρχίσουν να ασχολούνται με την ενσωμάτωση συγκεκριμένων δυνατοτήτων αναπαράστασης γνώσης (West and Hubbard, 1998), μια περιοχή στην οποία η τεχνητή νοημοσύνη έχει πολλά να προσφέρει. Αντίληψη

Ο βαθμός στον οποίον η αντίληψη σε ένα εικονικό περιβάλλον μοντελοποιείται με βάση την αντίληψη στον πραγματικό κόσμο διαφέρει πολύ ανάλογα με τους πράκτορες. Στον πραγματικό κόσμο η αντίληψη είναι ένα πολύ δύσκολο πρόβλημα, όπως είναι ήδη γνωστό από το χώρο των ρομπότ. Αντίθετα, στον εικονικό κόσμο η αντίληψη δε χρειάζεται να έχει προβλήματα, όπως η αμφιβολία, ο θόρυβος ή η χρονοβόρος επεξεργασία. Οι πιο απλοϊκοί εικονικοί αισθητήρες θα μπορούσαν να προβάλλουν μια ακτίνα από τα μάτια του πράκτορα και να επιστρέφουν πληροφορία σχετικά με οποιοδήποτε αντικείμενο του εικονικού κόσμου τέμνει, λαμβάνοντας το όνομα και τις ιδιότητές του απευθείας από τα δεδομένα του κόσμου (Noser et al., 1995).

Παρόλα αυτά, έχει γίνει αρκετή δουλειά και στην έρευνα για ένα βιολογικά ακριβές σύστημα αντίληψης για εικονικούς πράκτορες (Terzopoulos et al, 1996), στο οποίο το πεδίο όρασης του πράκτορα προβάλλεται σε έναν προσομοιωμένο αμφιβληστροειδή και χρησιμοποιούνται αλγόριθμοι από την τεχνητή όραση για να επεξεργαστούν τα pixels και να μετατραπούν σε μια αναγνωρίσιμη μορφή. Μεταξύ των δύο αυτών ακραίων περιπτώσεων υπάρχουν συστήματα εικονικών ρομπότ, στα οποία μοντελοποιούνται υπέρυθροι και υπερηχητικοί αισθητήρες με ένα βαθμό ρεαλισμού (Michel, 1998) και προστίθεται θόρυβος στο προσομοιωμένο σήμα.

Άσχετα από την ακρίβεια του βιολογικού μοντέλου οποιουδήποτε συστήματος εικονικών αισθητήρων είναι σημαντικό η αντίληψη του πράκτορα να είναι περιορισμένη όσο και η αντίστοιχη αντίληψη του χρήστη σε ένα εικονικό περιβάλλον. Για παράδειγμα, κάποιος θα περίμενε ότι ένας πράκτορας σε ένα σκοτεινό δωμάτιο έχει τη δυνατότητα να δει λιγότερα πράγματα σε σχέση με ένα φωτεινότερο. Όσο πιο απλό είναι το σύστημα αντίληψης ενός πράκτορα, τόσο περισσότερη επιπρόσθετη

γνώση πρέπει να μεταφερθεί στους πράκτορες από τα αντικείμενα που γίνονται αντιληπτά. Αντίθετα, ένα καλύτερα μοντελοποιημένο σύστημα αντίληψης απαιτεί πολύ λιγότερη γνώση από το περιβάλλον, αλλά χρησιμοποιεί περισσότερη υπολογιστική ισχύ από τη μεριά του πράκτορα, ώστε να παραχθεί η απαραίτητη γνώση.

Ενώ τα εργαλεία εικονικών περιβαλλόντων παρέχουν αισθητήρες, αυτοί δεν αντιστοιχούν σε εικονικές αναπαραστάσεις των αντίστοιχων αισθητήρων που θα είχε ένα ρομπότ αλλά σε διαδικασίες, που ελέγχουν θέματα αλληλεπίδρασης του χρήστη, όπως σήμα ότι ο χρήστης προσέκρουσε σε έναν τοίχο ή ότι επέλεξε με το ποντίκι κάποια περιοχή. Η κατασκευή συνεπώς εικονικών αισθητήρων απαιτεί επιπλέον προγραμματισμό σε οποιαδήποτε γλώσσα προσφέρει το εργαλείο. Ενέργεια

Εάν η αντίληψη δημιουργεί το ερώτημα στο κατά πόσο θα πρέπει ο πράκτορας να είναι ενσωματωμένος σε ένα συγκεκριμένο κόσμο, το ερώτημα ισχύει ακόμα περισσότερο στην περίπτωση της ενέργειας. Έτσι, ενώ η αντίληψη είναι παθητική, η ενέργεια προκαλεί αλλαγές στον κόσμο που εξαρτώνται από τη λειτουργικότητα του πράκτορα αλλά και τη λειτουργικότητα και την τρέχουσα κατάσταση του κόσμου (Prokopenko and Jauregui, 1997). Για παράδειγμα, αν ένας πράκτορας πιάσει ένα αντικείμενο, η ικανότητά του να το σηκώσει εξαρτάται τόσο από τη σωματική κατασκευή του, όσο και από το μέγεθος και βάρος του αντικειμένου. Επιπλέον, το βάρος του αντικειμένου μπορεί να εξαρτάται από το κατά πόσο ο κόσμος είναι στο επίπεδο της γης, μέσα σε έναν διαστημικό σταθμό ή στο βυθό του ωκεανού.

Σε πιο λεπτομερές επίπεδο, ενέργειες, όπως το πιάσιμο, θα πρέπει να είναι οπτικά πειστικές, για παράδειγμα ένα χέρι δε θα πρέπει να διαπερνά το αντικείμενο που πιάνει, ενώ ο τρόπος που ακουμπάει το χέρι το αντικείμενο θα πρέπει να είναι λογικοφανής. Η πολυπλοκότητα της αλληλεπίδρασης σε αυτό το επίπεδο εξαρτάται από τη λεπτομέρεια και την ακρίβεια με την οποία έχει μοντελοποιηθεί το χέρι, αλλά και από το κατά πόσο οι φυσικές δυνάμεις και οι περιορισμοί, όπως η τραχύτητα των επιφανειών έχουν αναπαρασταθεί στον κόσμο ή το πιάσιμο είναι απλώς μια συνθετική κίνηση, που αφορά ένα σύνολο πρωτογενών γραφικών αντικειμένων.

Ένα άλλο σημαντικό ζήτημα αφορά στην αλληλεπίδραση με σύνθετα αντικείμενα. Το ερώτημα είναι σε τι ποσοστό θα ελέγχεται η αλληλεπίδραση από το περιβάλλον και σε τι ποσοστό θα είναι ενσωματωμένη ιδιότητα του ίδιου του αντικειμένου. Ως τώρα δεν υπάρχει σαφής λύση στο πρόβλημα αυτό. Μια προσέγγιση είναι να ενσωματωθούν στο αντικείμενο οι απαραίτητες ιδιότητες και η απαραίτητη γνώση σχετικά με τον τρόπο αλληλεπίδρασης με αυτό. Για παράδειγμα, το σύστημα IMPROV (Goldberg, 1997) υιοθετεί αυτό που ο Goldberg ονομάζει αντίστροφη αιτιότητα (inverse causality) και αποθηκεύει τις συνθετικές κινήσεις της αλληλεπίδρασης μεταξύ ενός αντικειμένου και ενός εικονικού ηθοποιού μέσα στο ίδιο το αντικείμενο, έτσι ώστε να μην υπάρχουν απαιτήσεις εκμάθησης από τους εικονικούς ηθοποιούς. Κατά συνέπεια, ένας εικονικός ηθοποιός που επιλέγει ένα εικονικό

μπουκάλι μπύρας έχει τη δυνατότητα να πει από αυτό χωρίς να χρειάζεται να μάθει τις απαραίτητες ενέργειες, για να κάνει κάτι τέτοιο.

Αυτή η προσέγγιση δε φαίνεται ιδιαίτερα χρήσιμη σε περιπτώσεις που ενσωματώνονται και φυσικές ιδιότητες, όπως η βαρύτητα (Aylett et al., 1999), στο εικονικό περιβάλλον. Στην περίπτωση αυτή φαίνεται προτιμότερο όλα τα αντικείμενα που υπάρχουν στο περιβάλλον να υπακούουν σε έναν αριθμό από νόμους, όπως η ελεύθερη πτώση, όταν δεν έχουν στήριξη και υπάρχει βαρύτητα, ή η επίπλευση, όταν δεν υπάρχει βαρύτητα. Ένα τέτοιο παράδειγμα είναι η περίπτωση της συμπεριφοράς ενός ψαριού σε ένα εικονικό περιβάλλον που δεν είναι γεμάτο νερό (West and Hubbard, 1998). Μια προσέγγιση είναι η προσθήκη ιδιοτήτων, όπως η δύναμη, που θα επιτρέψουν σε έναν πράκτορα να αλληλεπιδράσει με έναν λογικοφανή τρόπο σε οποιοδήποτε εικονικό περιβάλλον κι αν ενσωματωθεί.

4.5.3 ΕΛΕΓΧΟΣ ΚΑΙ ΣΥΜΠΕΡΙΦΟΡΑ

Υπάρχει ένα κενό μεταξύ των μεθόδων που υιοθετούνται από τους ερευνητές από το πεδίο των γραφικών για τον έλεγχο των εικονικών πρακτόρων και αυτών που προτιμώνται από ερευνητές με ειδίκευση σε τεχνητή νοημοσύνη και τεχνητή ζωή. Το ζήτημα που διχάζει είναι ο σκηνοθετικός ή καλλιτεχνικός έλεγχος του πράκτορα σε σχέση με την αυτονομία που θα έχει στο περιβάλλον, όπου πολλοί ερευνητές από το χώρο της συνθετικής κίνησης συχνά προτιμούν διαφόρων ειδών γλώσσες σεναρίων για την περιγραφή της συμπεριφοράς, ενώ οι ερευνητές της τεχνητής νοημοσύνης συχνά σκέφτονται με τη λογική του ελέγχου συμπεριφοράς που βασίζεται σε αισθητήρες και επιδραστές ή σε ενέργειες που εξαρτώνται από τους στόχους και υποστηρίζονται από συμβολική συμπερασματολογία. Στη δεύτερη περίπτωση ακολουθείται η προσέγγιση της γενικότερης θεωρίας των ευφυών πρακτόρων. Η συμπεριφορά ενός εικονικού πράκτορα μπορεί να ενισχυθεί παραπέρα με την εισαγωγή αληθοφανών στοιχείων, όπως τα συναισθήματα και η προσωπικότητα (Badler et al., 2002).

Σενάρια

Οι γλώσσες σεναρίων είναι μια πολύ δημοφιλής μέθοδος, που επιτρέπει στους σχεδιαστές να ελέγχουν ευφυείς εικονικούς πράκτορες από υψηλότερο επίπεδο σε σχέση με τον παραδοσιακό τρόπο συνθετικής κίνησης. Ένα από τα πιο δημοφιλή συστήματα με γλώσσα σεναρίου είναι το Improm, το οποίο υποστηρίζει το σκηνοθετικό έλεγχο των εικονικών πρακτόρων. Παρόλο που δεν είναι τρισδιάστατο σύστημα, η γλώσσα σεναρίου του Microsoft AGENT είναι επίσης ευρέως χρησιμοποιούμενη, ενώ και το δημοφιλές σύστημα κίνησης και προσομοίωσης εικονικών ανθρώπων JACK περιλαμβάνει γλώσσα σεναρίου.

Οι γλώσσες σεναρίων βασίζονται συνήθως σε προ-κατασκευασμένες βιβλιοθήκες συμπεριφοράς, αλλά είναι σημαντικό να επεξηγηθεί ότι η συμπεριφορά εδώ δεν αντιστοιχεί στον αντίστοιχο όρο της τεχνητής νοημοσύνης. Εδώ συνήθως αναφέρεται σε προαποφασισμένες συνθετικές κινήσεις ή και σε μετακινήσεις που

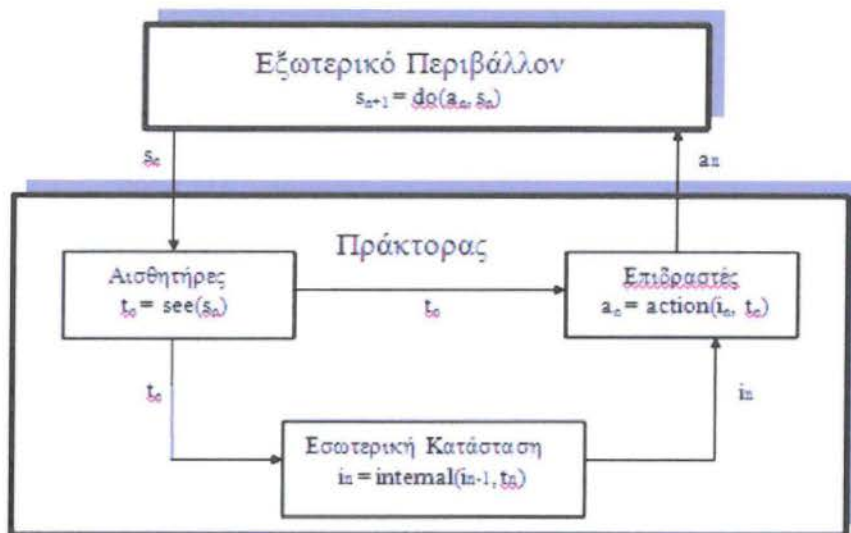
έχουν προέρθει από σύλληψη κίνησης, όπως βάδισμα, τρέξιμο, κοίταγμα, κλπ. Ευφυείς Πράκτορες

Ο όρος συμπεριφορά στο πεδίο της τεχνητής νοημοσύνης αναφέρεται σε ένα σύστημα ελέγχου που βασίζεται στις αισθήσεις (Brooks, 1986; Brooks, 1991), στο οποίο τα εισερχόμενα ερεθίσματα αντιστοιχίζονται σε εξερχόμενες αντιδράσεις. Στο χώρο αυτό, οι αρχιτεκτονικές συμπεριφορές στηρίζονται σε αρχιτεκτονικές ευφυών πρακτόρων (Wooldridge and Jennings, 1995), προσαρμοσμένες στο περιβάλλον με τους κατάλληλους αισθητήρες και επιδραστής.

Είναι αρκετά δύσκολο να ορίσουμε τι ακριβώς είναι αυτό που κάνει έναν πράκτορα «ευφυή». Ωστόσο, μπορούμε γενικά να πούμε ότι ένας πράκτορας είναι ευφυής όταν έχει την ικανότητα να επιτελεί τους στόχους και τα καθήκοντα που έχει επιφορτιστεί. Έτσι, σε ένα ελάχιστο επίπεδο νοημοσύνης μπορεί να δίνονται στον πράκτορα εντολές με τη μορφή κανόνων και αυτός να ενεργεί με τη βοήθεια κάποιου μηχανισμού εξαγωγής συμπεράσματος. Σε ένα ανώτερο επίπεδο ο πράκτορας θα είναι ικανός να μαθαίνει και να προσαρμόζεται αυτόματα στο περιβάλλον έτσι, ώστε να πετυχαίνει τους σκοπούς του.

Ένας ευφυής πράκτορας θα πρέπει να έχει ικανότητα αντίδρασης στα ερεθίσματα που δέχεται, επομένως πρέπει να έχει τη δυνατότητα να αντιλαμβάνεται το περιβάλλον, αλλά και να επιδρά πάνω σ' αυτό. Αυτό συμβαίνει με τη χρήση των αισθητήρων και επιδραστήων, που έχουν ήδη αναφερθεί.

Εκτός από τη δυνατότητα να αντιλαμβάνεται το περιβάλλον, στις περισσότερες περιπτώσεις ένας πράκτορας χρειάζεται να έχει μία εσωτερική κατάσταση, η οποία μπορεί να επηρεάζει τις ενέργειες και τη συμπεριφορά του. Η εσωτερική αυτή κατάσταση είναι ένα σύστημα λήψης αποφάσεων. Στην περίπτωση που ο πράκτορας δεν έχει εσωτερική κατάσταση, έχουμε μία απλουστευμένη αρχιτεκτονική, όπου οι ενέργειές του εξαρτώνται αποκλειστικά από τα ερεθίσματα που δέχεται από το περιβάλλον. Ένας τέτοιος πράκτορας ονομάζεται τροπιστικός (tropistic) Αντίθετα με έναν τροπιστικό πράκτορα, ένας πράκτορας με εσωτερική κατάσταση έχει τη δυνατότητα να διατηρεί και να χρησιμοποιεί εσωτερικές πληροφορίες. Η εσωτερική κατάσταση μπορεί να μεταβληθεί από εξωτερικά ερεθίσματα και επηρεάζει τις αποφάσεις που λαμβάνει ο πράκτορας.



Εικόνα 4.5 : Γενική αρχιτεκτονική ευφυών πρακτόρων

Οι παραπάνω περιγραφές της αρχιτεκτονικής ενός πράκτορα (εικ. 4.5) θα μπορούσαν να περιγραφούν με μία πιο αυστηρή μαθηματική μορφή θεωρώντας τα παρακάτω σύνολα.

S: Είναι το σύνολο των δυνατών καταστάσεων του περιβάλλοντος στο οποίο βρίσκεται ο πράκτορας.

T: Είναι το σύνολο των ερεθισμάτων που μπορεί να δεχτεί ο πράκτορας από το εξωτερικό περιβάλλον. Το σύνολο T καθορίζει ποιες από τις καταστάσεις του περιβάλλοντος φαίνονται ίδιες για τον πράκτορα και ποιες μπορεί να ξεχωρίσει μεταξύ τους. Έτσι το σύνολο T έχει ως στοιχεία υποσύνολα του S και, πιο συγκεκριμένα, διαμερίζει το σύνολο S.

A: Το σύνολο των δυνατών ενεργειών που μπορεί να κάνει ο πράκτορας.

I: Το σύνολο των δυνατών εσωτερικών καταστάσεων του πράκτορα.

Αφού ορίσαμε τα παραπάνω σύνολα, μπορούμε επίσης να θεωρήσουμε τις παρακάτω συναρτήσεις οι οποίες συσχετίζουν τα παραπάνω σύνολα μεταξύ τους: 59

$see : S \rightarrow T$

Η συνάρτηση see αντιστοιχεί τις εξωτερικές καταστάσεις του περιβάλλοντος σε ερεθίσματα που δέχεται ο πράκτορας μέσω των αισθητήρων. Δηλαδή η see καθορίζει το πώς αντιλαμβάνεται ο πράκτορας το εξωτερικό περιβάλλον. $do: A \times S \rightarrow S$

Η συνάρτηση `do` παίρνει ως είσοδο την παρούσα κατάσταση του περιβάλλοντος και μία ενέργεια και δίνει την επόμενη κατάσταση του περιβάλλοντος. Επομένως η `do` καθορίζει τα αποτελέσματα που έχουν οι ενέργειες στο εξωτερικό περιβάλλον. $action : I \times T \rightarrow A$

Η συνάρτηση `action` καθορίζει, δοθέντος ενός εξωτερικού ερεθίσματος και μιας εσωτερικής κατάστασης την ενέργεια που θα κάνει ο πράκτορας. $internal : I \times T \rightarrow I$

Η συνάρτηση `internal` καθορίζει ποια θα είναι η επόμενη εσωτερική κατάσταση δοθέντος του εξωτερικού ερεθίσματος που δέχτηκε ο πράκτορας, και της παρούσας εσωτερικής κατάστασης.

Με βάση τα παραπάνω μπορούμε να πούμε ότι ο πράκτορας καθορίζεται από την παρακάτω οκτάδα:

Πράκτορας = (I, S, T, A, see, do, action, internal)

Ο κύκλος λειτουργίας ενός πράκτορα της παραπάνω μορφής γραμμένος σε μορφή ψευδοκώδικα είναι ο παρακάτω:

```
repeat
  begin
    t = see(s)
    i = internal(i, t)
    a = action(i, t)
    s = do(a, s)
  end
```

Σε κάθε κύκλο ο πράκτορας δέχεται μέσω των αισθητήρων του το εξωτερικό ερέθισμα από το περιβάλλον, ενημερώνει την εσωτερική του κατάσταση, αποφασίζει για την ενέργεια που θα εκτελέσει και τέλος εκτελεί την ενέργεια με τους επιδραστές του αλλάζοντας έτσι το περιβάλλον.

Έχοντας ορίσει με τον παραπάνω τρόπο την αρχιτεκτονική ενός ευφυούς πράκτορα, το επόμενο βήμα είναι να καθοριστεί ο τρόπος με τον οποίο τα ερεθίσματα του περιβάλλοντος σε συνδυασμό με την παρούσα εσωτερική κατάσταση θα καθορίσουν την επόμενη εσωτερική κατάσταση καθώς και την ενέργεια που θα κάνει ο πράκτορας με τη βοήθεια των επιδραστών του. Πρέπει δηλαδή να οριστεί ο τρόπος με τον οποίο λειτουργούν οι συναρτήσεις `internal` και `action`. Πάνω σ' αυτό το θέμα υπάρχουν αρκετές προσεγγίσεις, που μπορούν να διακριθούν σε δύο βασικές κατηγορίες: τις κλασσικές και τις εναλλακτικές προσεγγίσεις.

Στην κλασσική προσέγγιση θεωρείται ο πράκτορας ως έναν ειδικός τύπος συστήματος βασισμένου σε γνώση (knowledge-based system) χρησιμοποιώντας έτσι μία αρχιτεκτονική που είναι ευρύτερα γνωστή ως θεληματική (deliberative). Με τον όρο «θεληματική αρχιτεκτονική πράκτορα» εννοείται η αρχιτεκτονική, η οποία αναπαριστά τον κόσμο με ένα συμβολικό μοντέλο και οι αποφάσεις του πράκτορα παίρνονται μέσω λογικών συμπερασμάτων βασισμένων στην επεξεργασία συμβόλων. Η κατασκευή πρακτόρων με βάση τη θεληματική αρχιτεκτονική θέτει δύο σημαντικά προβλήματα:

Το πρόβλημα της μετάφρασης (transduction problem), δηλαδή πώς θα μεταφραστεί ο πραγματικός κόσμος σε μία επαρκή και ακριβή συμβολική περιγραφή.

Το πρόβλημα της αναπαράστασης και της συμπερασματολογίας (representation and reasoning problem), πώς δηλαδή θα αναπαρασταθεί συμβολικά η πληροφορία για σύνθετες οντότητες και διαδικασίες του πραγματικού κόσμου και πώς θα εξαγάγει συμπεράσματα ο πράκτορας βασισμένος στις πληροφορίες αυτές.

Το πρώτο πρόβλημα απαιτεί τεχνικές από πεδία όπως αναγνώριση ομιλίας, αναγνώριση εικόνας κλπ, ενώ για το δεύτερο πρόβλημα χρειάζονται γνώσεις πάνω σε αναπαράσταση γνώσης, αυτόματη συμπερασματολογία κλπ. Ωστόσο φαίνεται ότι κανένα από τα δύο αυτά προβλήματα δεν έχει βρει ακόμα τη λύση του, αφού ακόμα και απλά προβλήματα, όπως για παράδειγμα η συμπερασματολογία απλής λογικής (common sense reasoning), έχουν αποδειχθεί εξαιρετικά πολύπλοκα. Το βασικό πρόβλημα φαίνεται ότι είναι η δυσκολία εφαρμογής τεχνικών απόδειξης θεωρημάτων (theorem proving), ακόμα και σε απλές λογικές, και η πολυπλοκότητα των αλγορίθμων επεξεργασίας συμβόλων. Έτσι, η ιδέα της κατασκευής ενός έξυπνου πράκτορα χρησιμοποιώντας συμβολική λογική και αλγορίθμους απόδειξης θεωρημάτων, αν και αρκετά ελκυστική ως ιδέα, αποδεικνύεται δύσκολη στην πρακτική εφαρμογή.

Τα πολλά άλυτα προβλήματα που έχουν οι κλασσικές προσεγγίσεις, που είδαμε παραπάνω, οδήγησαν πολλούς ερευνητές σε αρχιτεκτονικές που είναι ευρύτερα γνωστές ως αρχιτεκτονικές αντίδρασης (reactive architectures). Τέτοιου είδους αρχιτεκτονικές έχουν το χαρακτηριστικό ότι δε χρησιμοποιούν συμβολικό μοντέλο αναπαράστασης του κόσμου ούτε και πολύπλοκες τεχνικές συμπερασματολογίας. Βασική ιδέα των αρχιτεκτονικών αντίδρασης είναι ότι η νοήμων συμπεριφορά μπορεί να δημιουργηθεί χωρίς σαφείς αναπαραστάσεις και χωρίς τεχνικές σαφούς συμπερασματολογίας, αλλά απλώς μπορεί να προκύψει μέσω της αλληλεπίδρασης του πράκτορα με το περιβάλλον του.

Τέλος, έχουν προταθεί και πολλές υβριδικές αρχιτεκτονικές, που προσπαθούν να συνδυάσουν τη θεληματική αρχιτεκτονική με την αρχιτεκτονική αντίδρασης, εκμεταλλευόμενες τα πλεονεκτήματα της κάθε μίας. Έτσι ένας υβριδικός πράκτορας αποτελείται από δύο υποσυστήματα: το γνωσιακό υποσύστημα, όπου υπάρχει μία συμβολική αναπαράσταση του κόσμου, και το υποσύστημα αντίδρασης, με το οποίο ο

πράκτορας μπορεί να αντιδρά στα ερεθίσματα του περιβάλλοντος χωρίς να καταφεύγει σε τεχνικές απόδειξης θεωρημάτων. Συνήθως δίνεται προτεραιότητα στο υποσύστημα αντίδρασης έτσι, ώστε ο πράκτορας να μπορεί να αποκρίνεται άμεσα σε ερεθίσματα του περιβάλλοντος. Οι υβριδικές αρχιτεκτονικές έχουν αρκετά πλεονεκτήματα συγκρινόμενες με τις γνωστικές αρχιτεκτονικές και τις αρχιτεκτονικές αντίδρασης. Ωστόσο, το σημαντικότερο πρόβλημα αυτών των αρχιτεκτονικών είναι ότι δεν είναι πάντα εύκολο να συνδυαστούν μεταξύ τους τα δύο υποσυστήματα. Συναισθήματα

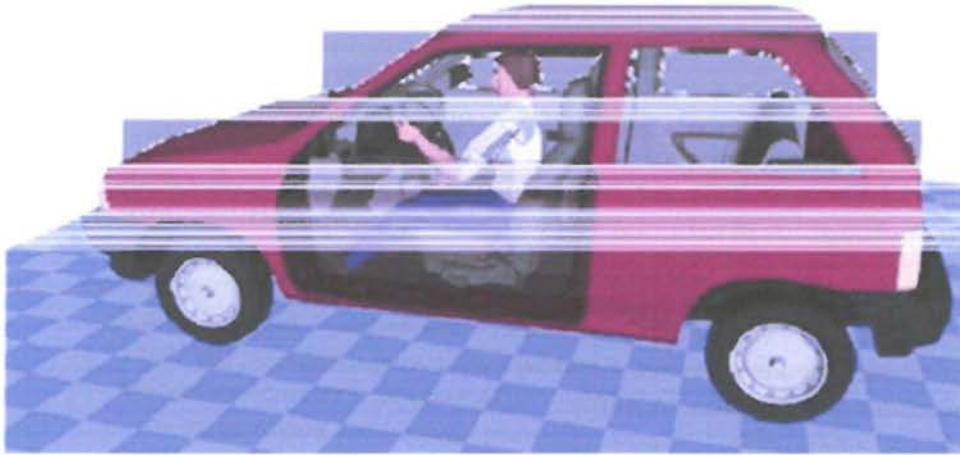
Η έρευνα στους εικονικούς πράκτορες έχει δώσει μια νέα ώθηση στο χώρο των κινήτρων και συναισθημάτων των πρακτόρων (Allbeck and Badler, 2002), πράγμα το οποίο οφείλεται σε δύο βασικούς λόγους. Πρώτον, ένας ενσωματωμένος εικονικός πράκτορας σε τρισδιάστατο περιβάλλον παρέχει πολλά κανάλια εξωτερίκευσης της συναισθηματικής του κατάστασης (βλέμμα, έκφραση προσώπου, χειρονομίες και γενικότερα γλώσσα του σώματος) σε σχέση με τους «παραδοσιακούς» ευφυείς πράκτορες, όπου ο μόνος τρόπος έκφρασης ήταν το γλωσσικό κείμενο. Δεύτερον, υπάρχουν πολλά πεδία εφαρμογής εικονικών πρακτόρων, στα οποία η έκφραση της συναισθηματικής κατάστασης είναι αναγκαία.

Η έμφαση που δίνουν όσοι εργάζονται σε γνωσιακό επίπεδο είναι στο συναίσθημα ως κατάσταση του μυαλού, ενώ αυτοί που ασχολούνται με το φυσικό επίπεδο αντιμετωπίζουν το συναίσθημα ως κατάσταση του σώματος. Αυτό αφορά βασικά την εσωτερική μοντελοποίηση του συναισθήματος και όχι τον τρόπο έκφρασής του. Οι δύο αυτές προσεγγίσεις αντανακλούν μια μακρά διαμάχη στο χώρο της ψυχολογίας (Picard, 1997).

5.5.4 ΟΛΟΚΛΗΡΩΜΕΝΑ ΣΥΣΤΗΜΑΤΑ

Jack

Μια σημαντική εφαρμογή που υλοποιεί πολλές πτυχές της ανθρώπινης κίνησης και προσομοίωσης είναι ένα εμπορικό σύστημα με το όνομα Jack (Badler et al., 1993b), που αναπτύχθηκε στο πανεπιστήμιο της Πενσυλβανίας και εστιάζει σε εφαρμογές εργονομικού και μηχανολογικού σχεδιασμού. Για παράδειγμα, κάποιος μπορεί να χρησιμοποιήσει το μοντέλο του Jack σε ένα ταμπλό χειρισμού ενός οχήματος, ώστε να αποφασίσει για τη λειτουργικότερη διάταξη των πλήκτρων και μοχλών ελέγχου. Ο έλεγχος του Jack μπορεί να επιτευχθεί από το χρήστη είτε με άμεσο χειρισμό των μελών του σώματός του, είτε με την εκτέλεση υψηλού επιπέδου εντολών συμπεριφοράς, όπως βάδισμα ή πιάσιμο. Επιπλέον, υπάρχει δυνατότητα διασύνδεσης με άλλες γλώσσες προγραμματισμού, όπως η Lisp και η C++, που επιτρέπουν στον Jack να λειτουργήσει ως αυτόνομος πράκτορας, αν αυτό είναι επιθυμητό.



Εικόνα 4.6 : Το μοντέλο του Jack μέσα σε ένα εικονικό αυτοκίνητο

Οι εφαρμογές στις οποίες στοχεύει το σύστημα χρειάζονται ένα βαθμό βιολογικού ρεαλισμού. Τα μοντέλα του Jack έχουν 68 αρθρώσεις με περίπου 120 βαθμούς ελευθερίας. Παρόλο που αυτό βιολογικά δεν είναι εντελώς ακριβές, οι σημαντικές αρθρώσεις, όπως ο ώμος, έχουν μοντελοποιηθεί με ακρίβεια, ενώ μπορούν να περιστραφούν και οι βολβοί των ματιών, ώστε να ελεγχθεί το κοίταγμα του μοντέλου. Μπορούν να μετρηθούν οι ροπές σε συγκεκριμένα σημεία του μοντέλου, ενώ δίνεται βάση και στο βάρος των αντικειμένων που κρατάει ο πράκτορας. Οι φιγούρες του Jack γεννιούνται από μια βάση δεδομένων, που περιέχει ένα μεγάλο εύρος ανθρώπινων χαρακτηριστικών, αντρικών, γυναικείων και παιδικών, και έτσι είναι δυνατόν να κατασκευαστούν άνθρωποι με ιδιαίτερα χαρακτηριστικά και ικανότητες ανάλογα με την ηλικιακή τους ομάδα, το φύλο, τη σωματική κατασκευή κλπ.

Οι εντολές υψηλού επιπέδου συμπεριφοράς μεταφράζονται σε ενέργειες χαμηλού επιπέδου, που εκτελούνται στους μύς του μοντέλου με τη χρήση μοντέλων κινηματικής και αντίστροφης κινηματικής, όπως συμβαίνει και στο χώρο της ρομποτικής. Έχει δοθεί ιδιαίτερη έμφαση στην ακρίβεια της χαμηλού επιπέδου συμπεριφοράς, έτσι που, για παράδειγμα, το βάδισμα να περιλαμβάνει χτύπημα της φτέρνας στο έδαφος. Επιπροσθέτως, έχει ενσωματωθεί και ένας μικρός αριθμός αυτόνομων ενεργειών χαμηλού επιπέδου, έτσι ώστε, για παράδειγμα, εάν μια φιγούρα χάσει την ισορροπία της και γέρνει προς τα εμπρός, θα κάνει αυτόματα ένα βήμα μπροστά για να ισορροπήσει. Παρόμοια, αν μια φιγούρα έχει εντολή να παρατηρεί ένα αντικείμενο, θα το κάνει με το βλέμμα του, όσο το αντικείμενο βρίσκεται στο

Εικονικοί Πράκτορες πεδίο όρασής του, αλλά θα στρίψει και το κεφάλι του, εάν το αντικείμενο μετακινηθεί έξω από αυτό.

Η ακρίβεια που υπάρχει στον έλεγχο του Jack έχει μεγάλο υπολογιστικό κόστος, αλλά καθώς το σύστημα αντίληψης έχει μοντελοποιηθεί με πολύ λιγότερη ακρίβεια σε

σχέση με το σύστημα μυών και ο υψηλού επιπέδου έλεγχος αφήνεται στο χρήστη, είναι δυνατό να εκτελεστεί το σύστημα σε πραγματικό χρόνο, εφόσον τρέχει σε ισχυρό μηχάνημα και το περιβάλλον δεν είναι ιδιαίτερα πολύπλοκο. Στα ισχυρά μηχανήματα επεξεργασίας γραφικών μπορούν να τρέξουν σε πραγματικό χρόνο μέχρι και τέσσερα ανθρώπινα μοντέλα ταυτόχρονα.

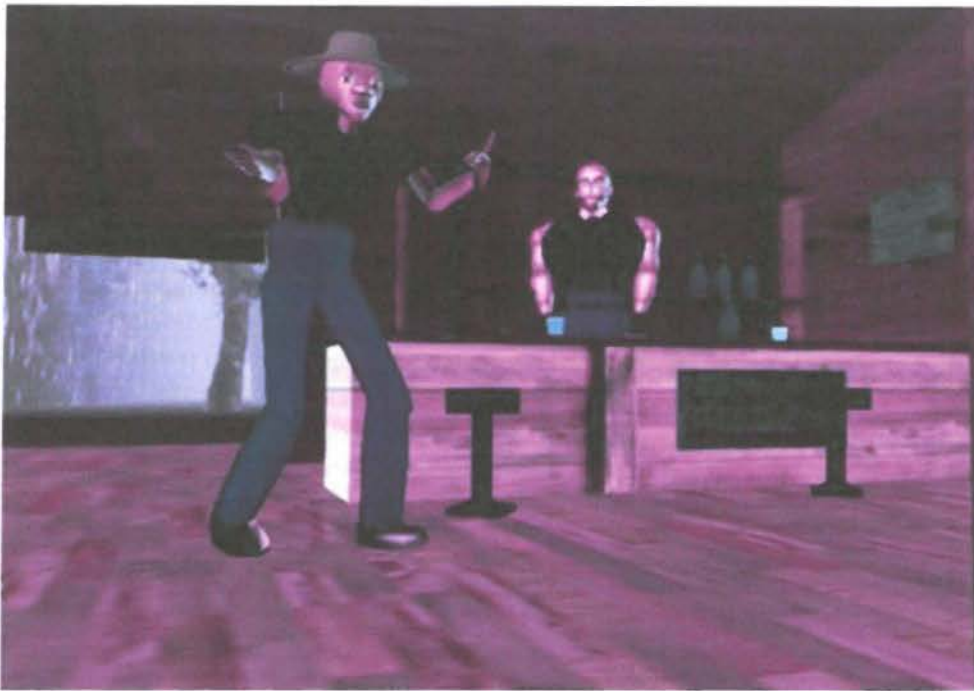
Humanoid

Ένα ευρωπαϊκό αντίστοιχο του συστήματος Jack αναπτύχθηκε στο πρόγραμμα HUMANOID (Boulic et al., 1995). Περιλαμβάνει 75 βαθμούς ελευθερίας για το σώμα και 30 επιπλέον βαθμούς ελευθερίας για κάθε χέρι. Για να γίνει η εμφάνιση του μοντέλου πιο ρεαλιστική, ο σκελετός καλύπτεται από ένα στρώμα που έχει κατασκευαστεί με την τεχνική των metaballs για την αναπαράσταση του δέρματος και των μυών. Το μοντέλο του HUMANOID έχει χρησιμοποιηθεί σε έναν αριθμό από ευρωπαϊκά προγράμματα.

3.3.3 Impron

Ένα άλλο ενδιαφέρον σύστημα, που δίνει έμφαση στην απεικόνιση και στην κίνηση, είναι το Impron (Perling and Goldberg, 1996; Goldberg, 1997). Το Impron είναι ένα σύστημα για την κατασκευή εικονικών ηθοποιών που κινούνται σε πραγματικό χρόνο με βάση κάποιο μοντέλο συμπεριφοράς. Παρέχει εργαλεία για τη δημιουργία ηθοποιών που ανταποκρίνονται στις ενέργειες του χρήστη ή και στις ενέργειες άλλων ηθοποιών σε πραγματικό χρόνο, βασισμένοι σε προσωπικότητες και διαθέσεις και με συνέπεια προς τους στόχους και τις προθέσεις του σχεδιαστή.

Οι χαρακτήρες του Impron αποτελούνται από μια γεωμετρία, που μεταβάλλεται σε πραγματικό χρόνο, μια μηχανή συνθετικής κίνησης, που αξιοποιεί τις περιγραφές των ατομικών ενεργειών (όπως βάδισμα ή χαιρετισμός), για να προκαλέσει αλλαγές στη γεωμετρία, και μια μηχανή συμπεριφοράς, που είναι υπεύθυνη για τις υψηλότερου επιπέδου ικανότητες (όπως μετάβαση σε συγκεκριμένη τοποθεσία ή διάλογος με άλλον ηθοποιό) και τις αποφάσεις για τις συνθετικές κινήσεις που θα ενεργοποιηθούν.



Εικόνα

4.7 : Στιγμιότυπο από ιστορία κατασκευασμένη με το Improv

Η μηχανή συνθετικής κίνησης χρησιμοποιεί διαδικαστικές τεχνικές, ώστε να επιτρέψει στους σχεδιαστές να παράγουν συνεχείς μη-επαναλαμβανόμενες κινήσεις σε στρώματα (layers), καθώς και ομαλές μεταβάσεις μεταξύ αυτών. Αντίθετα, η μηχανή συμπεριφοράς, που αποτελεί το μυαλό του ηθοποιού, χειρίζεται το εσωτερικό μοντέλο του ηθοποιού, που αναπαριστά τη διάθεση, τους στόχους και την προσωπικότητά του, και επιτρέπει στους σχεδιαστές να δημιουργήσουν περίπλοκους κανόνες, που αφορούν τον τρόπο με τον οποίο οι ηθοποιοί επικοινωνούν, μεταβάλλουν την εσωτερική τους κατάσταση και παίρνουν αποφάσεις. Το συνδυασμένο σύστημα παρέχει ένα ενσωματωμένο σύνολο εργαλείων για το σχεδιασμό των «μυαλών» και «σωμάτων» των εικονικών ηθοποιών. Παρέχεται μια γλώσσα σεναρίων που θυμίζει την αγγλική γλώσσα, έτσι ώστε οι δημιουργικοί σχεδιαστές, που δεν είναι κατ' ανάγκη προγραμματιστές, να μπορούν να κατασκευάζουν ισχυρές αλληλεπιδραστικές εφαρμογές.

Το Improv βρίσκει κυρίως εφαρμογή στο χώρο της διασκέδασης, δηλαδή στα παιχνίδια και στις αλληλεπιδραστικές αφηγήσεις ιστοριών (interactive storytelling), καθώς εξυπηρετεί στο σχεδιασμό ιστοριών με εικονικούς ηθοποιούς, που δε χαρακτηρίζονται από ιδιαίτερη αυτονομία ούτε εμπεριέχονται στοιχεία προσομοίωσης.

3.3.4 ACE (Agents Common Environment)

Το ACE (Kallmann et al., 2000) είναι μια πλατφόρμα για προσομοιώσεις εικονικών ανθρώπων, που αξιοποιεί τα πλεονεκτήματα τόσο των σεναρίων, όσο και

των μοντέλων συμπεριφοράς. Ο πυρήνας του συστήματος καταλαβαίνει ένα σύνολο εντολών που μπορούν να ελέγξουν μια προσομοίωση. Μεταξύ άλλων οι εντολές αυτές μπορούν:

- να δημιουργήσουν και να τοποθετήσουν διαφορετικούς εικονικούς ανθρώπους, αντικείμενα και έξυπνα αντικείμενα (Kallmann and Thalmann, 1999; Kallmann and Thalmann, 2002), δηλαδή αντικείμενα με ενσωματωμένη πληροφορία αλληλεπίδρασης.
- να εφαρμόσουν μια μηχανική κίνηση σε έναν εικονικό άνθρωπο. Παραδείγματα τέτοιων κινήσεων είναι: συνθετική κίνηση με θέσεις-κλειδιά, αντίστροφη κινηματική, βάδισμα, εκφράσεις προσώπου κλπ. Οι κινήσεις αυτές μπορούν να εκτελεστούν και παράλληλα, όπου συνδυάζονται ανάλογα με δοσμένες ιδιότητες από ένα συγκεκριμένο τμήμα της εφαρμογής.
- να ενεργοποιήσουν την αλληλεπίδραση ενός εικονικού ανθρώπου με ένα έξυπνο αντικείμενο. Κάθε έξυπνο αντικείμενο περιλαμβάνει μια λίστα από τις διαθέσιμες αλληλεπιδράσεις με αυτό, που εξαρτώνται από την εσωτερική κατάσταση του αντικειμένου. Κάθε αλληλεπίδραση περιγράφεται με απλά πλάνα, που είναι προσχεδιασμένα με τη χρήση του κατάλληλου γραφικού συστήματος διεπαφής. Τα πλάνα αυτά περιγράφουν την κατάλληλη ακολουθία μηχανικών κινήσεων που επιτυγχάνουν μια αλληλεπίδραση. Με το γραφικό σύστημα διεπαφής ορίζονται αλληλεπιδραστικά οι τρισδιάστατες παράμετροι που απαιτούνται για την εκκίνηση των μηχανικών κινήσεων, όπως θέσεις που θα τοποθετηθούν τα χέρια, κινήσεις που θα εφαρμοστούν στα τμήματα του αντικειμένου κλπ. 67
- να υποβάλλουν ερώτηση στο σύστημα αντίληψης ενός δοσμένου εικονικού ανθρώπου. Η αντίληψη μπορεί να διαμορφωθεί σε σύστημα διασωληνώσεων (pipeline) επιμέρους φίλτρων αντίληψης και να προσομοιώσει, για παράδειγμα, συνθετική όραση. Η ερώτηση στην περίπτωση αυτή θα επιστρέψει μια λίστα με όλα τα αντικείμενα που έγιναν αντιληπτά στο δοσμένο οπτικό πεδίο και εύρος.

Τα μοντέλα συμπεριφοράς δεν αποτελούν μέρος του πυρήνα του ACE, αλλά μπορούν να συνδεθούν εξωτερικά, είτε τοπικά είτε από απόσταση μέσω του πρωτοκόλλου TCP/IP, χρησιμοποιώντας σενάρια στην γλώσσα Python που καταλήγουν στις παραπάνω εντολές. Συμπερασματικά, το ACE είναι ένα ολοκληρωμένο σύστημα που αξιοποιεί πολλές πτυχές των ευφυών εικονικών πρακτόρων, η αδυναμία όμως ενσωμάτωσης όλων των στοιχείων του σε ενιαία αρχιτεκτονική δυσκολεύει τη χρήση του για την κατασκευή αυτόνομων περιβαλλόντων, που μπορεί να αποτελούν τμήματα μεγαλύτερων προγραμμάτων.



Εικόνα 4.8: Προσομοίωση χώρου εργασίας με το σύστημα ACE

Καθώς οι εικονικοί πράκτορες «κατοικούν» σε εικονικά περιβάλλοντα, οι εφαρμογές τους έχουν φυσικά άμεση σχέση με τις εφαρμογές των εικονικών περιβαλλόντων που περιγράφηκαν αναλυτικά στο προηγούμενο κεφάλαιο. Τα πιο βασικά πεδία εφαρμογής που χρησιμοποιούν ή πρόκειται να χρησιμοποιήσουν την τρέχουσα έρευνα στους εικονικούς πράκτορες είναι (Badler, 1997):

- Μηχανολογία, σχεδιασμός και συντήρηση: Στο χώρο αυτό αξιοποιούνται οι εικονικοί άνθρωποι που έχουν μοντελοποιηθεί με μεγάλη ακρίβεια για την ανάλυση και προσομοίωση εικονικών πρωτοτύπων και το σχεδιασμό βασισμένο στην προσομοίωση.
- Εικονικά Περιβάλλοντα / Εικονικές Διασκέψεις: Χρησιμοποιούνται μοντέλα εικονικών πρακτόρων για αποδοτικές τηλεδιασκέψεις με τη χρήση αναπαραστάσεων των χρηστών για τη μείωση του εύρους επικοινωνίας. Ακόμα, οι εικονικοί πράκτορες αξιοποιούνται και σε περιβάλλοντα συνεργασίας για ανάλυση, εκπαίδευση και ψυχαγωγία.
- Εκπαίδευση και παρουσιάσεις: Έχει γίνει αρκετή έρευνα στη χρήση εικονικών πρακτόρων για την αλληλεπιδραστική καθοδήγηση του χρήστη στην πραγματοποίηση διαφόρων ενεργειών με σκοπό την εκπαίδευσή του. Μια τέτοια περίπτωση είναι το πρόγραμμα STEVE (Rickel and Johnson, 1999), που αξιοποιήθηκε για την εκπαίδευση του προσωπικού ενός πλοίου στο χειρισμό των μηχανημάτων του. Επιπλέον, εικονικοί πράκτορες χρησιμοποιούνται σε συστήματα παρουσιάσεων, αλλά και στην απόδοση κειμένου στη νοηματική γλώσσα των κωφαλάλων.
- Παιχνίδια και Διασκέδαση: Τόσο σε αλληλεπιδραστικά συστήματα αφήγησης ιστοριών όσο και σε παιχνίδια (Narayek,

2000) υπάρχει η τάση να χρησιμοποιούνται θεωρίες εικονικών πρακτόρων στους συνθετικούς χαρακτήρες, ώστε να αποκτήσουν μεγαλύτερη αληθοφάνεια τόσο στη φυσικότητα της κίνησης, όσο και στη συμπεριφορά τους με τη χρήση συναισθημάτων και προσωπικότητας (Silva et al., 1999)

6 ΒΙΟΝΙΚΑ ΤΕΧΝΗΤΑ ΜΕΛΗ

Τα τεχνητά μέλη διακρίνονται σε προθέσεις κάτω και άνω άκρων, εξυπηρετώντας οποιαδήποτε μορφή ακρωτηριασμού.

Τα ειδικά πέλματα αποτελούν έναν από τους βασικότερους τρόπους για να θεραπευτούν ή να αποφευχθούν τα προβλήματα των ποδιών. Κατασκευάζονται χρησιμοποιώντας αφρώδη αποτυπώματα, πελματογράφημα ή γύψινο πρόπλασμα. Τα υλικά που χρησιμοποιούνται απορροφούν τους κραδασμούς και αποτρέπουν τις πιέσεις. Υπάρχουν 3 κύρια είδη πελμάτων:

1. Τα Διορθωτικά πέλματα, η κατασκευή των οποίων επιδιώκει να διορθώσει και να επαναφέρει το πέλμα στην αρχική – υγιή μορφή του,
2. Τα Υποστηρικτικά πέλματα που συγκρατούν το πέλμα στην σωστή θέση και,
3. Τα Μαλακά πέλματα όπου η κατασκευή τους αποσκοπεί στην αποσυμφόρηση όλων αυτών των περιοχών, στις οποίες πρέπει να αποφευχθεί η πίεση.

Πλέον πραγματοποιείται μια επέμβαση, κατά την οποία επανατοποθετούνται σε άλλο σημείο του σώματος νεύρα όπως, για παράδειγμα, κάτω από το στήθος, τα αισθητήρια νεύρα του χεριού και χάρη σ' αυτό ο ασθενής της μπορεί να νιώσει κάποιον που του σφίγγει το τεχνητό χέρι. Η νέα τεχνική δίνει στον ασθενή τη δυνατότητα να αισθάνεται το άγγιγμα, την άσκηση διαφορετικής πίεσης στα δάχτυλα, την αίσθηση του κρύου και του ζεστού, ακόμη και την ικανότητα να αντιλαμβάνεται την υφή των αντικειμένων. Τα τεχνητά μέλη μοιάζουν ολοένα και περισσότερο με αληθινά και σύντομα θα είναι μονίμως τοποθετημένα, θα συνεργάζονται άριστα με τα οστά, τους ιστούς και το νευρικό σύστημα, και θα ελέγχονται από τον εγκέφαλο.

Όσον αφορά την ιστορική τους προέλευση, σε αιγυπτιακές μούμιες έχουν βρεθεί προσθετικά άκρα φτιαγμένα από ίνες ξύλου. Επίσης σ' έναν ελληνικό μύθο, η Δήμητρα, θεά της γεωργίας, κατασκεύασε ένα προσθετικό ώμο από ελεφαντόδοντο. Το πιο σημαντικό όμως είναι η ρωμαϊκή πρόθεση από το 300 Π.Χ. που ανακαλύφθηκε στην Ιταλία, το 1858 ήταν κατασκευασμένη από ξύλο, χαλκό, και δερμάτινα λουριά.

Διαχρονικά, έχουμε δει σε πολλά μυθιστορήματα, δοξασίες αλλά και κινηματογραφικές ταινίες να δείχνουν ανθρώπους με ιδιαίτερα χαρακτηριστικά φτιαγμένα με τεχνολογικά μέσα. Υπάρχει άραγε δυνατότητα να βελτιωθεί το είδος του ανθρώπου μέσω τεχνολογίας; Μπορεί για παράδειγμα ένας άνθρωπος να βλέπει πολύ μακριά, να ακούει ήχους σε συχνότητες που οι άνθρωποι δεν μπορούν να ακούσουν, να μπορεί να τρέξει πιο γρήγορα από αυτοκίνητο; Βέβαια η αρχή έχει γίνει

ήδη, αφού βλέπουμε ιατρικές εφαρμογές αποκατάστασης οργάνων, ή βελτίωσης όρασης και ακοής να εφαρμόζονται πλέον ευρύτατα.

Βιονικοί τένοντες, τεχνητό δέρμα που επαναφέρει την αίσθηση της αφής, τεχνητός εγκέφαλος και μικροτσιπ που αποκαθιστούν την όραση των τυφλών είναι μερικά από τα τελευταία επιστημονικά επιτεύγματα που συνθέτουν τον βιονικό άνθρωπο. Μέλη και ιστοί που έχουν καταστραφεί ανεπανόρθωτα μπορούν να επιδιορθωθούν χάρη στη προηγμένη τεχνολογία της εποχής μας, σώζοντας ή ακόμα και βελτιώνοντας την ποιότητα της ζωής των ασθενών.

Σε διάφορα πανεπιστήμια και εργαστήρια στον κόσμο πραγματοποιούνται έρευνες για την δημιουργία τέτοιων μελών.

Για παράδειγμα Αυστραλοί επιστήμονες στο πανεπιστήμιο «Λα Τρόμπι» έχουν σχεδιάσει ένα εμφύτευμα το οποίο λειτουργεί ακριβώς όπως το ανθρώπινο αυτί και είναι ικανό να αναγνωρίζει ήχους σε θορυβώδες περιβάλλον. Το εμφύτευμα αυτό βρίσκεται ακόμα σε πειραματικό στάδιο. Στο πανεπιστήμιο της Νότιας Καλιφόρνιας έχει δημιουργηθεί από επιστήμονες ένα μικροτσιπ το οποίο εμφυτεύεται στον εγκέφαλο με σκοπό την αποκατάσταση της μνήμης. Το μικροτσιπ αυτό έχει δοκιμαστεί με ιστό από εγκέφαλο ποντικών.

Οι δύο κυριότερες αιτίες τύφλωσης είναι η μελαχρωστική αμφιβληστροπάθεια και η εκφύλιση ωχράς κιλίδας. Γιατροί στην Γερμανία κατάφεραν να αποκαταστήσουν την όραση σε τρεις τυφλούς ασθενείς εμφυτεύοντας τσιπάκια στον αμφιβληστροειδή τους. Οι μικροεπεξεργαστές αυτοί διαθέτουν ηλεκτρονικούς αισθητήρες παρόμοιους με αυτούς των ψηφιακών μηχανών. Όταν τους χτυπά το φως παράγουν ηλεκτρικές ώσεις που μεταφέρονται στο οπτικό νεύρο και από εκεί στον εγκέφαλο. Στο πανεπιστήμιο του Μάντσεστερ ερευνητές ετοιμάζουν τεχνητούς τένοντες από πλαστικές ίνες που μπορούν να περιστρέφονται. Οι τένοντες αυτοί βρίσκονται ακόμα σε πειραματικό στάδιο.

Όλα αυτά έχουν ως αποτέλεσμα στο ότι σε μερικά χρόνια οι σημερινές συσκευές χειρός θα είναι τόσο μικρές που θα μπορούν να φορεθούν όπως ένα βραχιόλι για παράδειγμα. Αλλά για στοιχεία τα οποία μπορούν να συνοψιστούν μέσα σε πολύ λίγα «bytes» όπως είναι τα διαβατήρια, θα είναι πολύ πιο πρακτική η χειρουργική τοποθέτησή τους μέσα στον οργανισμό μας. Με άλλα λόγια ίσως να μην είναι τόσο μακριά η εποχή κατά την οποία αισθητήρες τοποθετημένοι σε διάφορους δημόσιους χώρους, θα είναι σε θέση να ελέγχουν κάθε κίνησή μας.



Τεχνητή όραση, ρομποτικά μέλη, έξυπνα γόνατα, τεχνητό νεφρό και φορητό πάγκρεας είναι μερικές από τις 10 βιονικές τεχνολογίες που δοκιμάζονται στα επιστημονικά εργαστήρια. Μερικές μάλιστα, έχουν ήδη αρχίσει να εφαρμόζονται σε ασθενείς σε πειραματικό στάδιο.

ΕΓΚΕΦΑΛΟΣ

Δεν θα μπορούσε να λείπει από τη λίστα ο εγκέφαλος. Το τμήμα που ονομάζεται υπόκαμπος ελέγχει τη βραχυπρόθεσμη μνήμη και την αίσθηση του χώρου. Όταν αυτό παθαίνει βλάβη από Αλτσχάιμερ, εγκεφαλικό ή άλλες αιτίες, οι επιπτώσεις σε αυτές τις λειτουργίες είναι σημαντικές. Ο Θίοντορ Μπέργκερ, καθηγητής στο Πανεπιστήμιο της Νότιας Καλιφόρνιας, δημιούργησε έναν ηλεκτρονικό επεξεργαστή που μπορεί να εμφυτευθεί στον εγκέφαλο και να αντικαταστήσει τον υπόκαμπο. Η μέθοδος δεν έχει τελειοποιηθεί ακόμα.



People control The Utah Electrode Array as implanted in a human brain. For a patient and more details, go to CyberWeek.com.

Η ΣΤΥΤΙΚΗ ΔΥΣΛΕΙΤΟΥΡΓΙΑ

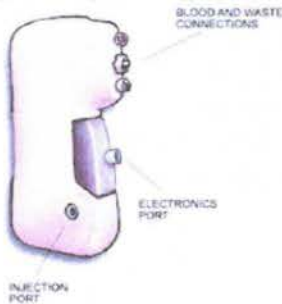
Βρίσκεται επίσης στο στόχαστρο των ερευνητών. Ο Άντονι Ατάλα και η ομάδα του στο Πανεπιστήμιο Wake Forest κατάφεραν να αναπτύξουν στυτικό ιστό στο εργαστήριο, το τμήμα του πέους που γεμίζει αίμα κατά τη σύση. Δοκίμασαν τη μέθοδο σε κουνέλια, τα οποία είχαν χάσει τον δικό τους ιστό και τον αντικατέστησαν με αυτόν που αναπτύχθηκε στο εργαστήριο από δικά τους κύτταρα. Η δοκιμή στέφθηκε με πλήρη επιτυχία.

ΤΕΧΝΗΤΑ ΚΥΤΤΑΡΑ

Μερικές φορές ένα φάρμακο πρέπει να κατευθυνθεί στον στόχο μέσα στο σώμα με απόλυτη ακρίβεια. Τα τεχνητά κύτταρα είναι η απάντηση των ειδικών, εκεί που το χάπι ή η ένεση δεν μπορούν να φτάσουν. Ο Ντάνιελ Χάμερ, καθηγητής Βιομηχανικής στο Πανεπιστήμιο της Πενσυλβάνιας έχει αναπτύξει τεχνητά κύτταρα από πολυμερή, που μιμούνται την ευκολία με την οποία ταξιδεύουν τα λευκά αιμοσφαίρια στο σώμα. Αυτά τα ψεύτικα κύτταρα μπορούν να μεταφέρουν τα φάρμακα εκεί που χρειάζονται, συμβάλλοντας στην καταπολέμηση ασθενειών όπως ο καρκίνος.

ΤΟ ΤΕΧΝΗΤΟ ΝΕΦΡΟ

Είναι χαρμόσυνη είδηση για όσους είναι υποχρεωμένοι να περνούν πολύτιμες ώρες στο νοσοκομείο τρεις φορές την εβδομάδα για τη διαδικασία της αιμοδιάλυσης. Μόνον όταν συνδέονται με το σχετικό μηχάνημα αποβάλλονται οι τοξίνες από το αίμα και διατηρούνται τα σωστά επίπεδα των υγρών στον οργανισμό. Οι Μάρτιν Ρόμπερτς και Ντέιβιντ Λι στο Πανεπιστήμιο του Λος Άντζελες στην Καλιφόρνια (UCLA) κατασκεύασαν ένα φορητό τεχνητό νεφρό, τόσο μικρό και ελαφρύ που να μεταφέρεται εύκολα από τον ασθενή και να μπορεί να χρησιμοποιείται επί 24ώρου βάσεως κάθε ημέρα, όπως ακριβώς συμβαίνει με ένα αληθινό νεφρό.



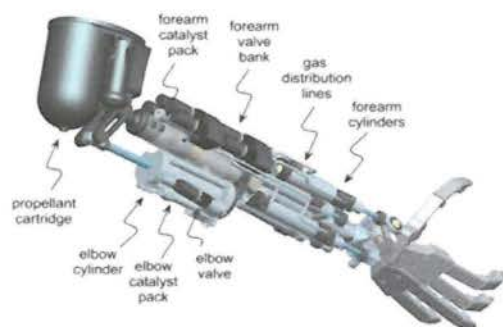
ΠΡΟΣΘΕΤΙΚΟ ΓΟΝΑΤΟ

Οι Χιου Χερ και Άρι Γουίλκενφελντ, ερευνητές τεχνητής ευφυΐας στο MIT, σχεδίασαν ένα προσθετικό γόνατο που... σκέφτεται. Τα προηγούμενα ηλεκτρονικά γόνατα απαιτούσαν προγραμματισμό από ειδικό πριν φορεθούν από τον ασθενή. Το έξυπνο γόνατο του MIT μαθαίνει τον τρόπο που περπατά ο ασθενής και μιμείται την κίνηση όσο πιο ρεαλιστικά γίνεται. Οι ενσωματωμένοι αισθητήρες το βοηθούν να καταλάβει πάνω σε τι υλικό περπατά για να προσαρμόσει την κίνησή του. Το γόνατο που εφαρμόζεται σε ένα προσθετικό μέλος κάνει το περπάτημα πολύ πιο άνετο και ξεκούραστο.



ΒΙΟΝΙΚΑ ΧΕΡΙΑ

Τα προσθετικά βιονικά χέρια που κινούνται με τη σκέψη του ασθενούς χρησιμοποιούνται εδώ και λίγα χρόνια στην Αμερική. Ο Τοντ Κιούκεν στο Ινστιτούτο Αποκατάστασης του Σικάγου ανέπτυξε το μοντέλο που συνδέεται με τον εγκέφαλο μέσω υγιών κινητικών νευρώνων, που ενώνονταν με το χέρι του ασθενούς. Αυτοί οι νευρώνες επανακατευθύνονται σε άλλη περιοχή του σώματος, όπως ο θώρακας, και τα σήματα που μεταφέρουν εντοπίζονται από τα ηλεκτρόδια του βιονικού χεριού. Όταν ο ασθενής αποφασίσει να κουνήσει το χέρι, οι ίδιοι νευρώνες μεταφέρουν την εντολή στο προσθετικό μέλος. Η ομάδα του Κιούκεν προσπαθεί να κάνει βελτιώσεις συνδέοντας το μέλος με τους αισθητήριους νευρώνες για να μεταδώσουν την αίσθηση της θερμοκρασίας, της δόνησης και της πίεσης στον εγκέφαλο.



Η ΓΛΩΣΣΑ

Και περνάμε στη γλώσσα, ένα πανίσχυρο αλλά και υποκειμενικό όργανο γεύσης και αίσθησης. Ο Ντιν Νίκικ, καθηγητής Πληροφορικής και ηλεκτρολόγος μηχανικός στο Πανεπιστήμιο του Τέξας στο Όστιν, σχεδίασε μία ηλεκτρονική γλώσσα, την οποία χρησιμοποιούν οι εταιρείες τροφίμων για να δημιουργήσουν κάθε φορά την ίδια γεύση. Η γλώσσα αυτή αναλύει υγρά και στερεά και καταλαβαίνει την ακριβή χημική τους σύνθεση χρησιμοποιώντας μικροσφαιρίδια. Πρόκειται για μικροσκοπικούς αισθητήρες που αλλάζουν χρώμα όταν εκτίθενται σε συγκεκριμένες ουσίες, όπως για παράδειγμα κάποια σάκχαρα. Ως αποτέλεσμα, εξασφαλίζεται η πανομοιότυπη χημική σύνθεση της τροφής.



ΤΟ ΤΕΧΝΗΤΟ ΠΑΓΚΡΕΑΣ

Θα βγει στην αγορά σε λίγα χρόνια, σύμφωνα με τον δημιουργό του Άαρον Κοβάλσκι, υπεύθυνο Ερευνών στο Ερευνητικό Ίδρυμα Νεανικού Διαβήτη. Το όργανο θα μετρά τα επίπεδα σακχάρου στο αίμα και θα προσαρμόζει ανάλογα τα επίπεδα της ινσουλίνης που χρειάζεται ο οργανισμός. Η συσκευή θα συνδυάζει δύο υπάρχουσες τεχνολογίες: αντλία ινσουλίνης και συνεχή παρακολούθηση της γλυκόζης. Το τεχνητό πάγκρεας θα βοηθήσει τους διαβητικούς να ζήσουν μία «κανονική» ζωή χωρίς τις πολύ επικίνδυνες επιπτώσεις των χαμηλών ή υψηλών επιπέδων σακχάρων στο αίμα.



ΑΝΑΠΛΑΣΗ ΚΥΤΤΑΡΩΝ

Από τη δεκαετία του '60, οι ερευνητές γνωρίζουν ότι κάποιες πρωτεΐνες βοηθούν τα οστά να αναπλάσουν ένα τμήμα που έχει χαθεί ή τραυματιστεί. Η μέθοδος όμως, αποδείχθηκε πολύ δύσκολη στην εφαρμογή της. Το 2005, επιστήμονες στο UCLA έλυσαν το πρόβλημα χρησιμοποιώντας μία ειδικά σχεδιασμένη πρωτεΐνη που μπορεί να ξεκινήσει την ανάπλαση σε συγκεκριμένο τύπο κυττάρων. Με την ονομασία UCB-1 η πρωτεΐνη χρησιμοποιείται για την ανάπτυξη νέου οστού που μπορεί- για παράδειγμα- να ακινητοποιήσει τμήματα των σπονδύλων ανακουφίζοντας τον ασθενή από οξύ πόνο στη μέση.

ΤΕΧΝΗΤΟ ΜΑΤΙ

Χάρη σε έναν τεχνητό αμφιβληστροειδή, οι τυφλοί έχουν ελπίδα να ανακτήσουν τη χαμένη τους όραση. Ο αμφιβληστροειδής αποτελείται από μία μικροκάμερα και μία δέσμη ηλεκτροδίων που εμφυτεύονται στο μάτι. Ο ασθενής μπορεί να διακρίνει αντικείμενα και ανθρώπους που βρίσκονται σε κοντινή τους απόσταση. Η μέθοδος δοκιμάστηκε με επιτυχία στην 57χρονη Μπάρμπαρα Κάμπελ, που είχε χάσει την όρασή της στα 40 της. Οι επιστήμονες εμφύτευσαν ηλεκτρόδια στο μάτι της. Η ασθενής φορά ειδικά γυαλιά που στη ράχη της μύτης έχουν μία μικροκάμερα και στη μέση έχει δεμένο έναν επεξεργαστή εικόνας βίντεο που μετατρέπει τις εικόνες σε φωτεινά και σκοτεινά σχέδια και τα μεταφέρει στον εγκέφαλο μέσω των οπτικών νεύρων.

Η λειτουργία του βασίζεται στη διέγερση των οπτικών νεύρων ενός προβληματικού οφθαλμού, με ηλεκτρικούς παλμούς από εμφυτευμένο microchip στο εμπρόσθιο ή στο οπίσθιο μέρος του αμφιβληστροειδούς.

- 1.Εμφύτευμα στον οφθαλμό με βοηθητικό εξωτερικό εξοπλισμό
- 2.Εμφύτευμα εξ ολοκλήρου μέσα στον οφθαλμό
- 3.Ενσωματωμένη κάμερα πάνω σε γυαλιά

Ο ασθενής σχηματίζει τελικά εικόνα για το προς παρατήρηση αντικείμενο με τη συνένωση εικόνων, από διάφορες οπτικές γωνίες, με ασαφή περιγράμματα και χρώματα.

Γίνεται προσπάθεια εκμετάλλευσης, εκτός του οπτικού φάσματος, και του υπερύθρου, με χρήση διόδων υπερύθρου, αφού είναι γνωστό ότι κάθε σώμα ανάλογα με τη θερμοκρασία του εκπέμπει στην υπέρυθρη περιοχή του ηλεκτρομαγνητικού φάσματος.

Η δίοδος τοποθετείται στην κόρη του ματιού και η αποστολή του λαμβανομένου σήματος σε κατάλληλη ηλεκτρονική διάταξη στοναμφιβληστροειδή γίνεται με πολυμερές νημάτιο μήκους 3cm και πάχους 1μm.

Αυτό θα έδινε τη δυνατότητα αναγνώρισης αντικειμένων ακόμα και στο σκοτάδι. Απαραίτητη προϋπόθεση είναι η δημιουργία κατάλληλου λογισμικού, που θα οργανώνει τις πληροφορίες από το υπέρυθρο και ορατό φάσμα.

Όσο μικραίνει η κλίμακα των διαστάσεων του εμφυτεύματος, τόσο περισσότερα ηλεκτρόδια ανά μονάδα τετραγωνικής επιφάνειας μπορούν να 'επέμβουν' σε 'λεπτότερα' σημεία του νευρικού ιστού και συνεπώς να δώσουν καλύτερη ανάλυση εικόνας.



ΒΙΟΝΙΚΗ ΚΑΡΔΙΑ

Ενα νέο υβριδικό υλικό καταφέρνει να συνδυάσει ανθεκτικούς νανοσωληνες άνθρακα και ζωντανά καρδιακά κύτταρα, που μιμούνται το φυσικό καρδιοαγγειακό ιστό με μεγάλη επιτυχία.

Ο βιονικός ιστός, που δημιουργήθηκε από τον Αλί Καντεμοσέιι, καθηγητή της ιατρικής σχολής του Πανεπιστημίου Χάρβαρντ, μπορεί να χρησιμοποιηθεί ως μυς για

βιολογικές μηχανές. Πολλές από τις ιδιότητες των φυσικών ιστών και των βιολογικών κυττάρων, όπως η αίσθηση του περιβάλλοντος και η ανταπόκριση σε αυτό, είναι ιδιαίτερα δύσκολο να επιτευχθούν με συνθετικά υλικά στη συμβατική ρομποτική. Οι ερευνητές ελπίζουν πως η κατασκευή μηχανών από βιολογικά υλικά, όπως ο καρδιακός ιστός, θα επεκτείνουν τις δυνατότητές τους. Αυτοί οι νέοι βιονικοί ιστοί μπορούν να κινηθούν ελεύθερα στο νερό, να ταλαντευτούν μπρος-πίσω και να εκτελέσουν μια σειρά άλλων κινήσεων, ελέγχοντας το σχήμα και το πάχος τους.

Εάν αυτά τα υλικά αποδειχθούν ασφαλή για το ανθρώπινο σώμα, μπορούν να χρησιμοποιηθούν για την επιδιόρθωση ιστών που έχουν καταστραφεί από καρδιακές προσβολές. Οι ερευνητές που ασχολούνται εργαστηριακά με τον καρδιοαγγειακό ιστό συχνά χρησιμοποιούν πολυμερή και τζελ, για να προσφέρουν στα καρδιακά κύτταρα ένα περιβάλλον μέσα στο οποίο μπορούν να αναπτυχθούν και να συμπεριφερθούν, όπως θα έκαναν μέσα στο ανθρώπινο σώμα. Το αποτέλεσμα έχει συνήθως δύο κρίσιμα λάθη, λέει ο Καντεμοσέι. Δεν έχουν την ηλεκτρική αγωγιμότητα της ανθρώπινης καρδιάς ούτε τη μηχανική δύναμή της.

«Όταν η καρδιά χτυπά, τα κύτταρα ανταποκρίνονται σε αυτή τη μηχανική δύναμη, απελευθερώνοντας χημικές ουσίες που ευνοούν την ανάπτυξη», εξηγεί ο Τόμας Γουέμπστερ, χημικός μηχανικός στο Northeastern University της Βοστώνης. Εάν ο βιονικός ιστός δεν έχει την αγωγιμότητα που έχει η υπόλοιπη καρδιά, τα ηλεκτρικά σήματα αντιμετωπίζουν καθυστερήσεις. Ενα τεχνητό επίθεμα που δεν έχει τις σωστές ιδιότητες, μπορεί να μην αναπτυχθεί κανονικά, με αποτέλεσμα να μην πάλλεται στον ίδιο χρόνο με την υπόλοιπη καρδιά, λέει ο Γουέμπστερ.

Η ομάδα του Καντεμοσέι λύνει το πρόβλημα αυτό, τοποθετώντας ανθεκτικούς, ηλεκτρικά αγώγιμους νανοσωλήνες άνθρακα μέσα σε ειδικά τζελ ανάπλασης ιστού. Ο Καντεμοσέι εμφύτευσε καρδιακά κύτταρα μέσα σε αυτά τα τζελ και μελέτησε τις ιδιότητές τους. Οι βιονικοί ιστοί είχαν αντίστοιχη ελαστικότητα με την καρδιά ποντικών - δηλαδή πολύ μεγαλύτερη συγκριτικά με τα τωρινά τεχνητά υλικά. Είχαν επίσης πολύ καλύτερη αγωγιμότητα και συγχρονισμένο παλμό.

Η ανθεκτικότητά τους ήταν επίσης βελτιωμένη, αφού οι νανοσωλήνες άνθρακα δημιουργούν ηλεκτρικούς συνδέσμους μεταξύ των κυττάρων, τα οποία παραμένουν σε επαφή ακόμα και όταν βρίσκονται υπό πίεση. Ο Γουέμπστερ διευκρινίζει όμως πως, πριν προχωρήσουν σε ιατρικές εφαρμογές, οι ερευνητές πρέπει να αποδείξουν ότι οι νανοσωλήνες άνθρακα δεν είναι τοξικοί. Αυτό είναι ιδιαίτερα σημαντικό, αφού δεν απορροφώνται από τον οργανισμό.

Η πρώτη χρήση των υλικών αυτών μπορεί να γίνει σε βιολογικές μηχανές, που χρησιμοποιούνται για την αποκατάσταση μολυσμένων περιοχών ή την επιδιόρθωση κτηρίων, προτείνει ο Καντεμοσέι. Πέρυσι, ερευνητές επέδειξαν μέδουσες-ρομπότ που κολυμπούσαν ελεύθερα και βιολογικές μηχανές που περπατούσαν μόνες τους, κατασκευασμένες από καρδιακό ιστό και πολυμερή υλικά. Όμως, χωρίς ηλεκτρική

αγωγιμότητα, οι χρήσεις τους είναι περιορισμένες, λέει ο Ρασίντ Μπασίρ, ο ερευνητής του Πανεπιστημίου του Ιλινόις, ο οποίος κατασκεύασε το ρομπότ που περπατά.



Τεχνητό πόδι που μιμείται με ακρίβεια την κίνηση ενός κανονικού ποδιού.

Τα προσθετικά άκρα είναι η αλήθεια ότι έχουν προσφέρει στους ανθρώπους που έχουν χάσει κάποιο μέλος τους, την δυνατότητα να το υποκαταστήσουν με ένα τεχνητό που ως ένα βαθμό επιτρέπει ίσως και κάποιες κινήσεις ή λειτουργίες που ήταν εφικτές πριν την απώλεια του μέλους. Όσον αφορά το πόδι αν και έχουν βγει διάφορα προσθετικά βοηθήματα, είναι η αλήθεια ότι αποτελεί ένα τμήμα του σώματος με κάποιες ιδιαιτερότητες και συγκεκριμένες απαιτήσεις.

Το πόδι χρησιμοποιείται για να στηρίξει το βάρος μας και να μας βοηθά να περπατάμε δίνοντας την απαραίτητη ώθηση για να κινηθούμε προς τα εμπρός. Αποτελεί ένα επίτευγμα εξελικτικής διαδικασίας εκατοντάδων χιλιάδων ετών το οποίο είναι δύσκολο στον άνθρωπο να το αντικαταστήσει με κάτι μηχανικό ώστε να κάνει την ίδια δουλειά το ίδιο καλά.

Ανάμεσα στα προσθετικά πόδια που κυκλοφορούν υπάρχουν ακόμα και μηχανικά τα οποία προσπαθούν να μιμηθούν τον φυσικό τρόπο βάδισης. Κανένα όμως δεν μπορούσε μέχρι στιγμής να είναι τόσο καλό ώστε να μπορεί να ανταποκριθεί σε διάφορα εδάφη και ανωμαλίες του οδοστρώματος, ούτε να επιτρέπει στον κάτοχό του να μπορεί να βαδίζει με μεγάλη φυσικότητα. Ένας εφευρέτης φαίνεται ότι μάλλον το κατάφερε όμως και το προσθετικό πόδι που ίσως να κάνει την διαφορά είναι γεγονός.

Ο εφευρέτης του Jerome Rifkin, ανέπτυξε μέσα σε 8 χρόνια ένα πρωτότυπο τεχνητό μηχανικό πόδι που έχει την ονομασία K3 Promoter και το οποίο λειτουργεί μόνο με το βάρος του σώματος και μιμείται τέλεια την φυσική βάδιση. Το ενδιαφέρον του για τα προσθετικά μέλη προκλήθηκε μετά από ένα ατύχημα το οποίο είχε και χρειάστηκε 3 χρόνια για να καταφέρει να περπατήσει ξανά. Τότε ξεκίνησε να σκέφτεται για μια καλύτερη εναλλακτική στα προσθετικά μέλη που υπήρχαν.

Το πόδι που κατασκεύασε είναι μηχανικό (αποτελείται από μέταλλο) αλλά έχει και αρθρώσεις σε διάφορα σημεία ώστε να μιμείται το κανονικό πόδι. Διαθέτει φτέρνα, πέλμα και μεγάλο δάκτυλο, τα άκρως απαραίτητα δηλαδή για την στήριξη του σώματος και την βάρδια. Όσο και αν φαίνεται απίστευτο το μηχανικό αυτό πόδι διαθέτει και τένοντες (συρματόσχοινα) οι οποίοι λειτουργούν σαν ελατήρια για να επαναφέρουν το πόδι κατά την βάρδια.

Το τεχνητό πόδι είναι κατασκευασμένο από μαγνήσιο και για αυτό το λόγο είναι πολύ ελαφρύ. Παρακάτω μπορείτε να δείτε ένα βίντεο με το τεχνητό πόδι που το δοκιμάζει ένας αρτιμελής. Εννοείται ότι το πόδι του από το γόνατο έως τα δάκτυλα είναι ακινητοποιημένο κατά τέτοιον τρόπο ώστε να προσομοιώσει την απώλεια των κάτω άκρων. Επίσης θα παρατηρήσετε την συμπεριφορά του σε διαφορετικές επιφάνειες ακόμα και ανέβασμα σκαλοπατιών.

Αγγίγμα με τη δύναμη του μυαλού

Πιθήκοι «αγγίζουν» εικονικά αντικείμενα: ελπίδα για ρομποτικά μέλη με αφή.

Δυο μακάκοι μπορούν «αγγίζουν» εικονικά αντικείμενα με τα ηλεκτρικά σήματα του εγκεφάλου τους οδηγώντας στην ανάπτυξη ρομποτικών «στολών» που θα προσφέρουν κίνηση και αφή στους τετραπληγικούς. Credit Katie Zhuang.

Όχι μόνο κίνηση αλλά και αφή για τους τετραπληγικούς με τη βοήθεια της τεχνολογίας υπόσχεται μια νέα μελέτη. Διεθνής ομάδα επιστημόνων κατόρθωσε να εκπαιδεύσει δυο πιθήκους ρέζους μακάκους ώστε να «πιάνουν» με ένα εικονικό χέρι διάφορα αντικείμενα και να «αισθάνονται» την υφή τους μόνο και μόνο με την «ηλεκτρική» δύναμη του μυαλού τους.

Ο επικεφαλής της έρευνας **Μιγκέλ Νικολέλις**, διευθυντής του κέντρου Νευρομηχανικής του αμερικανικού Πανεπιστημίου Ντιουκ είναι γνωστός για τις μελέτες του στην «ανάγνωση» του εγκεφάλου των πιθήκων και στο παρελθόν έχει κατορθώσει να μάθει τα πειραματόζωα να ελέγχουν τις κινήσεις ενός εικονικού χεριού στον ηλεκτρονικό υπολογιστή.

Όπως δήλωσε, η βελτίωση της τεχνικής ως προς την αίσθηση της αφής μάς φέρνει ένα ακόμη βήμα πιο κοντά στον απώτερο στόχο _ την ανάπτυξη ρομποτικών τεχνητών μελών ή και «εξωσκελετών» που θα κινούνται νοητικά και θα προσφέρουν αυτονομία και εμπειρία του κόσμου σε ανθρώπους που πάσχουν από σοβαρή παράλυση.

«Κάποια μέρα στο κοντινό μέλλον» τόνισε *«οι τετραπληγικοί ασθενείς θα επωφεληθούν από αυτήν την τεχνολογία όχι μόνο για να κινούν τα χέρια τους και να περπατούν ξανά αλλά και για να αισθάνονται την υφή των αντικειμένων που πιάνουν ή τις ανωμαλίες του εδάφους στο οποίο πατούν με τη βοήθεια ενός ρομποτικού εξωσκελετού»*.

Ηλεκτρόδια-ανιχνευτές

Όπως περιγράφουν στη μελέτη τους, η οποία δημοσιεύθηκε στην επιθεώρηση «Nature», οι ερευνητές εισήγαγαν στον εγκέφαλο των δυο πιθήκων ηλεκτρόδια τα οποία και «τοποθετήθηκαν» στον κινητικό και στον σωματοαισθητικό φλοιό _ τις περιοχές που σχετίζονται με την εκτέλεση εκούσιων κινήσεων και την επεξεργασία των αισθητικών ερεθισμάτων αντιστοίχως.

Οι πίθηκοι εκπαιδεύθηκαν ώστε να χρησιμοποιούν μόνο τον εγκέφαλό τους για να κινήσουν ένα εικονικό χέρι στην οθόνη του ηλεκτρονικού υπολογιστή και να «αγγίξουν» με αυτό διάφορα αντικείμενα. Τα ηλεκτρόδια που ήταν τοποθετημένα στον κινητικό φλοιό κατέγραφαν, μέσω ηλεκτρικών σημάτων, την πρόθεση κίνησης των πιθήκων και μετέδιδαν την πληροφορία στο πρόγραμμα του υπολογιστή.

Όταν το εικονικό χέρι άγγιζε τα αντικείμενα στην οθόνη ηλεκτρικά σήματα μεταδίδονταν από το πρόγραμμα στον σωματοαισθητικό φλοιό των πειραματόζων δίνοντάς τους πληροφορίες σχετικές με την αφή _ πώς δηλαδή θα αισθάνονταν στο χέρι τους αυτό το αντικείμενο αν το έπιαναν πραγματικά.

Αφή χωρίς άγγιγμα

Σε μια από τις ασκήσεις, κατά την οποία οι πίθηκοι έπρεπε να διαλέξουν ανάμεσα σε δυο αντικείμενα οπτικά ταυτόσημα αλλά διαφορετικά στην αφή, τα ζώα ξεχώριζαν το αντικείμενο που είχε συνδεθεί με μια ανταμοιβή (το οποίο παρήγαγε ένα ηλεκτρικό σήμα «αφής») από εκείνο που δεν συνοδευόταν από ανταμοιβή (ούτε από ηλεκτρικό σήμα).

Αυτό, σύμφωνα με τον κ. Νικολέλις, αποδεικνύει ότι ο εγκέφαλος των πιθήκων μπορούσε να αποκωδικοποιήσει πληροφορίες σχετικές με την αφή χωρίς να υπάρχει κανένα πραγματικό ερέθισμα στο δέρμα των πειραματόζων.

«Δεν ξέρουμε τι αντιλήφθηκαν τα ζώα, αλλά επρόκειτο για μια αίσθηση η οποία δημιουργήθηκε τεχνητά, συνδέοντας τα εικονικά δάχτυλα απ' ευθείας στον εγκέφαλο» δήλωσε.

Στολή που κινείται με το μυαλό σε τρία χρόνια

Απώτερος στόχος του ερευνητή, όπως και του Walk Again Project _ της διεθνούς πρωτοβουλίας στην οποία εμπύπτουν οι σχετικές μελέτες _ είναι η ανάπτυξη μιας «εξωσκελετικής στολής» που θα προσφέρει κινητική αυτονομία στους ανθρώπους με σοβαρή παράλυση. Οι αισθητικές πληροφορίες θα αποτελούν ένα από τα βασικά συστατικά της.

Βιβλιογραφία

Λύκειο Αριστοτελείου Κολλεγίου Θεσσαλονίκης

lykeio@aristotelio.edu.gr

<http://thumbs.dreamstime.com>

<http://www.scienceclarified.com>

<http://movies-theodorecalebkenneth.blogspot.gr/2011/03/examples-of-expert-systems.html>

www.ess.co.at
<http://www.niscair.res.in/sciencecommunication/researchjournals>
<http://library.thinkquest.org>
roboticsblog.org
www.murraystate.edu
www.digitaltrends.com
<http://tvxs.gr/news/sci-tec>
<http://www.pinterest.com>
www.marshall.edu
<http://miloslife.gr>
www.nature.com
www.mayo.edu
eduspace.free.fr
[el.wikipedia.org/wiki
users.sch.gr](http://el.wikipedia.org/wiki/users.sch.gr)
www.eetn.gr
www.sciencedirect.com/science/
www.learnartificialneuralnetworks.com
www.cse.msu.edu
psych.utoronto.ca/users/reingold
psifiakoskosmos.files.wordpress.com
www.perfectlogic.com/Articles/AI/ExpertSystems/ExpertSystems.html
www.researchgate.net/
nemertes.lis.upatras.gr
www.davincisurgery.com
robotic-surgery.med.nyu.edu
online.wsj.com
www.cts.usc.edu/roboticsurgery.htm
www.srobotics.org
spectrum.ieee.org/robotics/military-robots
www.bbc.co.uk/news/technology-2539598
<http://www.popsci.com/category/tags/military-robots>
<http://www.isciencetimes.com/articles/6219/20131022/incredible-bionic-man-smithsonian-walks-talks-breathes.htm>
<http://www.ncbi.nlm.nih.gov/pubmed/1816705>
www.ortholine.gr/index.php?cPath=52
revealedtheninthwave.blogspot.com/2013/11/blog-post_4927.html
www.physics.ntua.gr/pdf_doc_files/
www.radionefeli.com/ViewArticle.aspx?aid=346
<http://www.spyrospapadimitriou.gr/?view=page&id=8>
<http://osarena.net/logismiko/applications/neural-networks-ti-einai.html>
<http://users.auth.gr/~voyatzis/SeniorThesis/mTsouxnika.pdf>
<http://www.spyrospapadimitriou.gr/?view=page&id=8>
ADAM, J.A., 1993. Virtual reality is for real. IEEE Spectrum, vol. 30(10), pp. 30-33.

ALLBECK, J. and BADLER, N., 2002. Toward Representing Agent Behaviors Modified by Personality and Emotion. Proceedings of Workshop on Embodied Conversational at AAMAS 2002, Bologna, Italy. ARTHUR, K.W., BOOTH, K.S. and WARE, C., 1993. Evaluating 3D task performance for fish tank virtual worlds. *ACM Transactions on Information Systems*, vol. 11(3), pp. 239-265.

ASTHEIMER, P., DAI, F., GABEL, M., KRUSE, R., MOLLER, S. and ZACHMANN, G., 1994. Realism in Virtual Reality. In N.M. Thalmann and D. Thalmann (eds.): *Artificial Life and Virtual Reality*.

AUBEL, A., BOULIC, R. and THALMANN, D., 2000. Real-time Display of Virtual Humans: Levels of Detail and Impostors. *IEEE Trans. Circuits and Systems for Video Technology*, Special Issue on 3D Video Technology.

AYLETT, R. and CAVAZZA, M., 2001. Intelligent Virtual Environments - A State-of-the-art Report. Proceedings of Eurographics 2001, pp. 87-109.

AYLETT, R. and LUCK, M., 2000. Applying Artificial Intelligence to Virtual Reality: Intelligent Virtual Environments. *Applied Artificial Intelligence*, vol. 14, pp. 3-32.

AYLETT, R., HORROBIN, A., O'HARRE, J.J., OSMAN, A. and POLSHAW, M., 1999. Virtual Teletubbies: Reapplying a Robot Architecture to Virtual Agents. Proceedings of the 3rd International Conference on Autonomous Agents, ACM press, pp. 338.

BADLER, N. and WEBBER, B., 1995. Planning and Parallel Transition Networks: Animation's new Frontiers. In S. Y. Shin and T. L. Kunii (eds.): *Computer Graphics and Applications: Proceedings of Pacific Graphics '95*, World Scientific Publishing, pp. 101-117. 189

BADLER, N., 1997. Virtual humans for animation, ergonomics, and simulation. Proceedings of IEEE Workshop on Non-Rigid and Articulated Motion, Puerto Rico.

BADLER, N., ALLBECK, J., ZHAO, L. and BYUN, M., 2002. Representing and Parameterizing Agent Behaviors. Proceedings of Computer Animation 2002, IEEE Computer Society, Geneva, Switzerland, pp. 133-143.

BADLER, N., BINDIGANAVALA, R., BOURNE, J., ALLBECK, J., SHI, J. and PALMER, M., 1999. Proc. Real time virtual humans. International Conference on Digital Media Futures, Bradford, UK.

BADLER, N., BINDIGANAVALA, R., BOURNE, J., PALMER, M., SHI, J. and SCHULER, W., 2000. A Parameterized Action Representation for Virtual Human Agents. In J. Cassell, J. Sullivan, S. Prevost, and E. Churchill (eds.): *Embodied Conversational Agents*, MIT Press, pp 256-284.

BADLER, N., HOLLICK, M. and GRANIERI, J., 1993a. Real-Time Control of a Virtual Human Using Minimal Sensors. *Presence*, vol. 2(1), pp. 82-82

