

**ΑΝΩΤΑΤΟ ΤΕΧΝΟΛΟΓΙΚΟ ΕΠΑΓΓΕΛΜΑΤΙΚΟ ΙΔΡΥΜΑ ΚΡΗΤΗΣ
ΠΑΡΑΡΤΗΜΑ ΧΑΝΙΩΝ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΝΙΚΗΣ**

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

**ΤΕΧΝΙΚΕΣ ΕΚΠΑΙΔΕΥΣΗΣ ΚΑΙ ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ
ΔΙΚΤΥΩΝ**



Σπουδαστής: Χριπάσιος Γεώργιος

Επιβλέπων καθηγητής: Δρ.Αντώνιος Κωνσταντάρας

Χανιά 2011

ΠΕΡΙΕΧΟΜΕΝΑ:

Νευρωνικά Δίκτυα: Περίληψη - Abstract

1. Κεφάλαιο 1^ο: Νευρωνικά Δίκτυα Hopfield

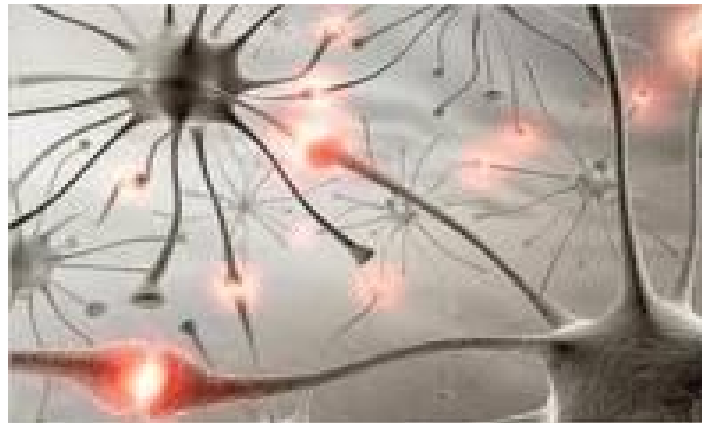
1.1. Εισαγωγή.....	
1.1.1. Παράδειγμα.....	
1.1.2. Ιδιαίτερο παράδειγμα.....	
1.2. Ανάλυση του δικτύου Hopfield:Υπολογιστική ενέργεια.....	
1.2.1. Παράδειγμα.....	
1.2.2. Ιδιαίτερα παραδείγματα.....	
1.3. Hopfield δίκτυα και σταθερές καταστάσεις.....	
1.4. Σχεδιάζοντας σταθερές καταστάσεις σε ένα δίκτυο Hopfield.....	
1.4.1. Ικανοποιώντας τις ανισότητες.....	
1.4.2. Εναλλακτική λύση των ανισοτήτων.....	
1.5. Εκπαιδεύοντας ένα δίκτυο Hopfield για απαιτούμενες σταθερές καταστάσεις.....	
1.6. Περίπτωση μελέτης- Μια πρακτική εφαρμογή των δικτύων Hopfield..	
1.6.1. Μπλοκ Αποκοπής Κωδικοποίησης - (BTC) Block Truncation Coding Σχηματικό Συμπίεσης Εικόνας.....	
1.6.2. Δίκτυο Hopfield (BTC) Σχηματικό συμπίεσης εικόνας.....	
1.6.3. Παράδειγμα.....	
1.6.4. Νευρωνικά δίκτυα Hopfield(παράδειγμα εκμάθησης).....	

2. Κεφάλαιο 2^ο: Η δομή της ανάδρασης με εγγενής συμπεριφορές μάθησης

2.1. Εισαγωγή.....	
2.2. Υπόβαθρο.....	
2.3. Μεταβλητή ανάδρασης.....	
2.3.1. Ένας γενικός κόμβος ανάδρασης.....	
2.3.2. Μεταβλητές ημιτόνου και συνημιτόνου.....	

2.3.3. Παραγωγή και άθροιση μεταβλητών.....
2.3.4. Μεταβλητές καταστάσεις.....
2.4. Παράδειγμα: Μοντελοποίηση ενός απλού ρομποτικού βραχίονα.....
2.5. Συζήτηση και συμπεράσματα.....
Βιβλιογραφία.....

Νευρωνικά Δίκτυα



Περίληψη:

Τα **νευρωνικά δίκτυα** έχουν γνωρίσει μια έκρηξη ενδιαφέροντος τα τελευταία χρόνια και εφαρμόζονται με μεγάλη επιτυχία σε ένα ασυνήθιστα μεγάλο φάσμα κατηγοριών και ποικίλων τομέων, όπως τα χρηματοοικονομικά, η ιατρική, η μηχανική, η γεωλογία και η φυσική. Στην πραγματικότητα, τα νευρωνικά δίκτυα εισάγονται οπουδήποτε τίθεται θέμα πρόβλεψης, ταξινόμησης ή ελέγχου. Η σαρωτική αυτή επιτυχία, μπορεί να αποδοθεί σε δύο βασικά στοιχεία: την ισχύ και την ευχρηστία.

Ισχύς: Τα νευρωνικά δίκτυα είναι πολύ εξελιγμένες τεχνικές μοντελοποίησης ικανές να μοντελοποιήσουν εξαιρετικά πολύπλοκες λειτουργίες. Πιο συγκεκριμένα, τα νευρωνικά δίκτυα είναι μη γραμμικά (non linear) μοντέλα. Για πολλά χρόνια μάλιστα, η γραμμική μοντελοποίηση υπήρξε ευρέως διαδεδομένη και χρησιμοποιείται στην πλειονότητα των κατηγοριών μοντελοποίησης. Αυτό το γεγονός συμβαίνει δεδομένου ότι τα γραμμικά

μοντέλα έχουν πολύ γνωστές στρατηγικές βελτιστοποίησης. Στις συχνές, όμως, περιπτώσεις όπου η γραμμική προσέγγιση δεν ήταν έγκυρη, τα μοντέλα αυτά αποτύγχαναν αναλόγως. Τα νευρωνικά δίκτυα, βέβαια, αντιμετωπίζουν με τη σειρά τους την ίδια δυσκολία στο ζήτημα της διάστασης, το οποίο αποτελεί αγκάθι στις προσπάθειες μοντελοποίησης μη γραμμικών συναρτήσεων με μεγάλο αριθμό μεταβλητών.

Ευχρηστία: Τα νευρωνικά δίκτυα εκπαιδεύονται με παραδείγματα. Ο χρήστης των δικτύων αυτών συγκεντρώνει αντιπροσωπευτικά δεδομένα και στη συνέχεια μέσω των αλγορίθμων, γίνεται αυτομάτως γνωστή η δομή των δεδομένων. Ο χρήστης χρειάζεται να έχει κάποιες ουσιώδεις γνώσεις σχετικά με τον τρόπο επιλογής και προετοιμασίας των δεδομένων, τον τρόπο εκλογής του κατάλληλου νευρωνικού δικτύου και στο πως θα ερμηνευτούν τα αποτελέσματα. Παρά ταύτα, το επίπεδο των γνώσεων του χρήστη που απαιτούνται για μια επιτυχημένη εφαρμογή των νευρωνικών δικτύων, είναι αρκετά χαμηλότερο συγκριτικά με κάποια περίπτωση που θα χρησιμοποιούνταν ορισμένες πιο παραδοσιακές, μη γραμμικές στατιστικές μέθοδοι.

Τα τεχνητά νευρωνικά δίκτυα (Αγγλικά: Artificial Neural Networks), ή απλώς νευρωνικά δίκτυα (Αγγλικά: Neural Networks) είναι ένα μαθηματικό μοντέλο για την επεξεργασία πληροφορίας που προσεγγίζει την υπολογιστική και αναπαραστατική δυνατότητα μέσω συνάψεων. Το μοντέλο είναι εμπνευσμένο από τα βιοηλεκτρικά δίκτυα που δημιουργούνται στον εγκέφαλο ανάμεσα στους νευρώνες (νευρικά κύτταρα) και στις συνάψεις (σημεία επαφής των νευρικών απολήξεων).



Στο μαθηματικό μοντέλο των νευρωνικών δικτύων υπάρχουν κομβικά σημεία (nodes) στα οποία καταλήγουν συνδέσεις από άλλους κόμβους του δικτύου, στις οποίες συνήθως αποδίδεται κάποιο βάρος.

Πρακτικά, ένα νευρωνικό δίκτυο βελτιστοποιεί μία συνάρτηση, σύμφωνα με κάποιους περιορισμούς.

Τα τεχνητά νευρωνικά δίκτυα, ως μαθηματικό μοντέλο, προέκυψαν από τον τομέα της Τεχνητής Νοημοσύνης.

Abstract:

Neural networks have seen an explosion of interest in recent years and applied successfully in an unusually wide range of categories and a variety of areas such as finance, medicine, engineering, geology and physics. In fact, neural networks are introduced wherever it comes to prediction, classification or control. The overwhelming success of this can be attributed to two main factors: the power and usability.

Power: The neural networks are very sophisticated modeling techniques able to model highly complex functions. More specifically, neural networks are nonlinear (non linear) models. For many years, indeed, the linear modeling was widespread and used in most types of modeling. This fact happens because the linear models are well-known optimization strategies. In frequent cases, however, where the linear approximation was not valid, these models fail accordingly. Neural networks, however, face turn the same difficulty on the question of dimension, which bedevils efforts modeling nonlinear functions with many variables.

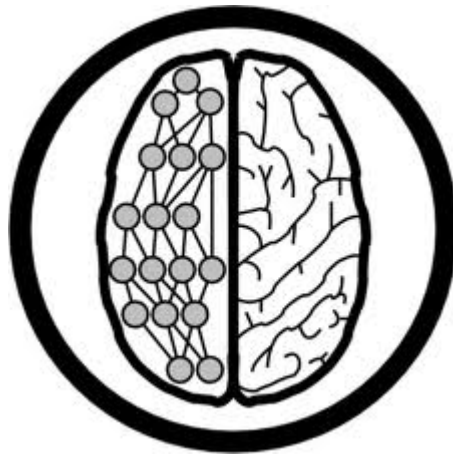
Usability: The neural networks are trained with examples. The user of these networks to collect representative data, and then through algorithms, automatically becomes known data structure. The user needs to have some basic knowledge about how to select and prepare data, how the election of the appropriate neural network and how to interpret the results. However, the level of knowledge of users required for a successful application of neural networks is much lower compared to a case that would be used some more traditional nonlinear statistical methods.

Artificial neural networks (English: Artificial Neural Networks), or simply neural networks (English: Neural Networks) is a mathematical model for information processing close to the representational and computational ability through synapses. The model is inspired by bioelectric networks created in the brain between neurons (nerve cells) and synapses (contact points of nerve endings).

The mathematical model of neural networks are nodal points (nodes) that lead connections from other network nodes, which are usually given some weight. Practically, a neural network optimizes a function under some constraints. Artificial neural networks as a mathematical model derived from the field of Artificial Intelligence.

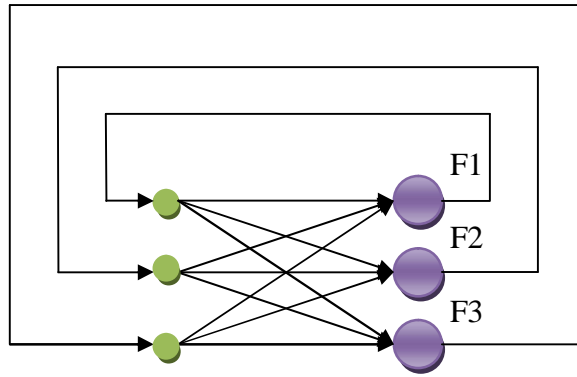
Κεφάλαιο 1^ο: Νευρωνικά Δίκτυα Hopfield

Νευρωνικά Δίκτυα Hopfield



1.1. Εισαγωγή:

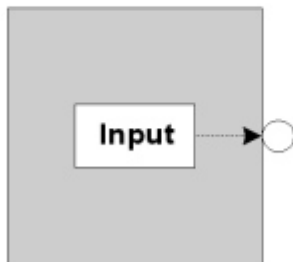
Ένα νευρωνικό δίκτυο Hopfield αποτελείται από ένα ενιαίο στρώμα των νευρώνων MCP με ανάδραση. Κάθε νευρώνας έχει εισόδους που έχουν ληφθεί από τα αποτελέσματα εξόδων των άλλων νευρώνων. Μια απλή αναπαράσταση ενός δικτύου Hopfield με τρεις νευρώνες απεικονίζεται παρακάτω.



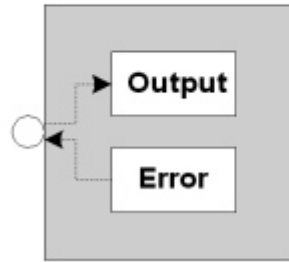
Σημειώστε ότι τα στοιχεία στα αριστερά δεν είναι νευρώνες, αλλά είναι απλώς fan-out (εισόδου-εξόδου) στοιχεία που επιτρέπουν στο διάγραμμα να συναχθεί με σαφήνεια.

Δεδομένου ότι κάθε έξοδο νευρώνα λαμβάνεται ως είσοδο σε κάθε νευρώνα, ο συμβολισμός που χρησιμοποιούμε για τα βάρη του δικτύου πρέπει να εμπεριέχει δύο δείκτες: έναν για να δείχνει από ποιον νευρώνα έχει ληφθεί η έξοδος και το άλλο για να δείχνει σε ποιο νευρώνα καταλήγει η είσοδος. Εμείς το καλούμε βάρος δικτύου το κομμάτι από την έξοδο του νευρώνα i σε μια είσοδο του νευρώνα j W_{ji} .

Neural network input



Neural network output



Κάθε νευρώνας μπορεί επίσης να έχει μια «εξωτερική» είσοδο, η οποία αναφέρεται ως X_i . Τώρα θα εξετάσουμε την περίπτωση όπου $X_i = 0$ για κάθε i και επομένως δεν υπάρχουν εξωτερικές εισόδους.

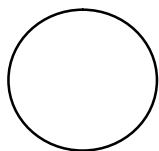
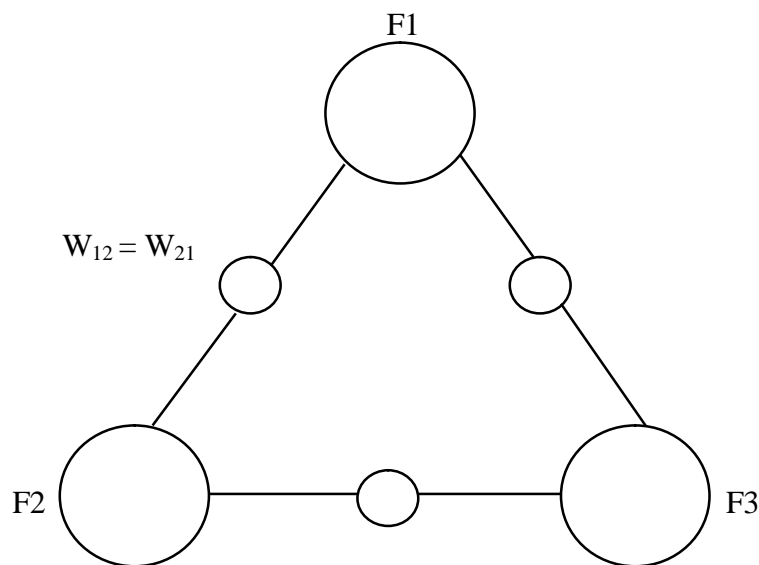
Θα εξετάσουμε τα δίκτυα όπου η έξοδος ενός νευρώνα δεν χρησιμοποιείται ως είσοδος για τον ίδιο νευρώνα. Αυτό μπορεί να επιτευχθεί με την άρση ορισμένων συνδέσεων στο παρακάτω διάγραμμα, ή πιο απλά και πιο εύκολα, θέτοντας:

$W_{ii} =$ για όλα τα i

Επιπλέον θα εξεταστεί η περίπτωση κατά την οποία ο πίνακας συναπτικών βαρών του δικτύου είναι συμμετρικός - που σημαίνει ότι το βάρος από την έξοδο i σε μια είσοδο του νευρώνα j είναι το ίδιο με το βάρος από την έξοδο j σε μια είσοδο του νευρώνα i , δηλαδή:

$W_{ij} =$

Η εισαγωγή αυτών των περιορισμών επιτρέπει σε ένα απλούστερο διάγραμμα να χρησιμοποιηθεί. Αποτελεσματικά υπάρχει μια αμφίδρομη σύνδεση μεταξύ δύο νευρώνων (δηλαδή νευρώνα i και νευρώνα j), οι οποία μπορεί να καθοριστεί με ένα βάρος $W_{ij}=W_{ji}$. Μια απλούστερη αναπαράσταση παρατίθεται στη συνέχεια και πάλι για ένα δίκτυο Hopfield με τρεις νευρώνες.



αντιπροσωπεύει ένα MCP νευρώνα.



αντιπροσωπεύει ένα αμφίδρομο βάρος.

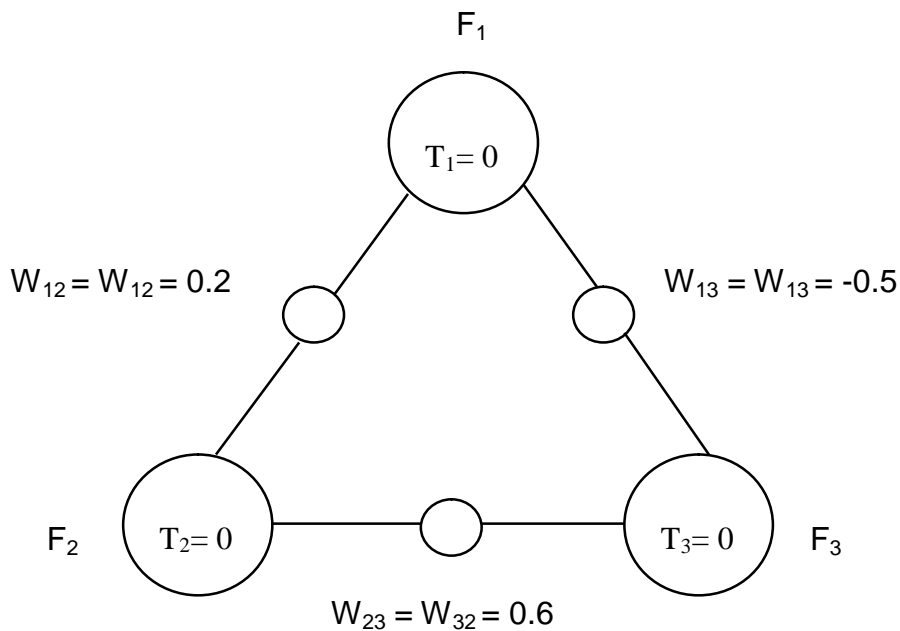
Ένας βασικός παράγοντας για το δίκτυο Hopfield είναι ότι έχει μια ασύγχρονη λειτουργία. Αυτό σημαίνει ότι ένας νευρώνας μπορεί να προσπαθήσει να εκπέμψει στο δικό του χρόνο και δεν είναι συγχρονισμένος με τις απόπειρες εκπομπής των άλλων νευρώνων στο δίκτυο. Μπορούμε να το ερμηνεύσουμε αυτό υπό την έννοια, ότι ανά πάσα στιγμή στο χρόνο κάθε ένας από τους νευρώνες στο δίκτυο έχει ίσες πιθανότητες να προσπαθήσει να εκπέμψει. Για ένα δίκτυο που αποτελείται από τρεις νευρώνες, όπως φαίνεται ανωτέρω, κάθε νευρώνας έχει πιθανότητα $1 / 3$ να είναι ο επόμενος νευρώνας που θα προσπαθήσει να εκπέμψει.

Σε κάθε δεδομένη χρονική στιγμή η έξοδος F_i κάθε νευρώνα είναι 0 ή 1. Έτσι το δίκτυο μπορεί να θεωρηθεί σε μια συγκεκριμένη κατάσταση που ορίζεται από την έξοδό του. Για παράδειγμα, η κατάσταση του δικτύου με τρεις νευρώνες ορίζεται ως $F_1 F_2 F_3$. Δεδομένου ότι κάθε έξοδος είναι δυαδική, υπάρχουν αποτελεσματικά οκτώ διαφορετικά δυνατές καταστάσεις για ένα δίκτυο τριών νευρώνων. Γενικά, για ένα δυαδικό δίκτυο Hopfield με N νευρώνες, υπάρχουν 2^N καταστάσεις που μπορούν να υφίστανται.

1.1.1. Παράδειγμα:



Ένα δίκτυο Hopfield με τρεις νευρώνες έχει τα βάρη και τα κατώτατα όρια (κατώφλια) όπου ορίζονται στο σχεδιάγραμμα που ακολουθεί.



Μπορούμε να αναλύσουμε το δίκτυο αυτό για να καθορίσουμε την επόμενη κατάσταση από οποιαδήποτε και ιδίως την παρούσα κατάσταση. Επειδή δεν γνωρίζουμε εκ των προτέρων ποιος νευρώνας θα είναι ο επόμενος που θα προσπαθήσει να εκπέμψει πρέπει να εξετάσουμε όλες τις περιπτώσεις ξεχωριστά και να επισημανθούν οι μεταβάσεις με τις κατάλληλες πιθανότητες.

Θα εξετάσουμε την παρούσα κατάσταση να είναι 010 αρχικά, δηλαδή:

$$F_1 = 0 \quad F_2 = 1 \quad F_3 = 0$$

Αν ο νευρώνας 1 είναι ο επόμενος για να προσπαθήσει να εκπέμψει, τότε μπορούμε να προβλέψουμε εάν θα εκπέμψει ή όχι εξετάζοντας τον κανόνα ενεργοποίησης για τον νευρώνα 1. Το άθροισμα των βαρών εισόδων συμπεριλαμβανομένου του κατώτατου ορίου - κατωφλίου (γνωστό ως ενεργοποίηση) για τον νευρώνα 1 φαίνεται να είναι :

$$F_2 W_{12} + F_3 W_{13} - T_1 = (1 * 0.2) + (0 * (-0.5)) + 0.1 = 0.3$$

Η τιμή αυτή είναι μεγαλύτερη από το μηδέν, και ως εκ τούτου ο νευρώνας 1 θα εκπέμψει ($F_1 = 1$). Έτσι οι αλλαγές στο δίκτυο δηλώνουν ότι αν η κατάσταση είναι η 110 ο F_1 είναι ο επόμενος νευρώνας για να προσπαθήσει να εκπέμψει (πιθανότητα του $1 / 3$).

Ωστόσο, πρέπει να εξετάσουμε επίσης τις περιπτώσεις κατά τις οποίες ο νευρώνας 2 ή ο νευρώνας 3 είναι ο επόμενος που θα προσπαθήσει να εκπέμψει, από την τωρινή κατάσταση 010. Αν ο νευρώνας 2 είναι ο επόμενος, η ενεργοποίησή του είναι :

$$F_1W_{21} + F_3W_{23} - T_2 = (0 * 0.2) + (0 * 0.6) - 0 = 0$$

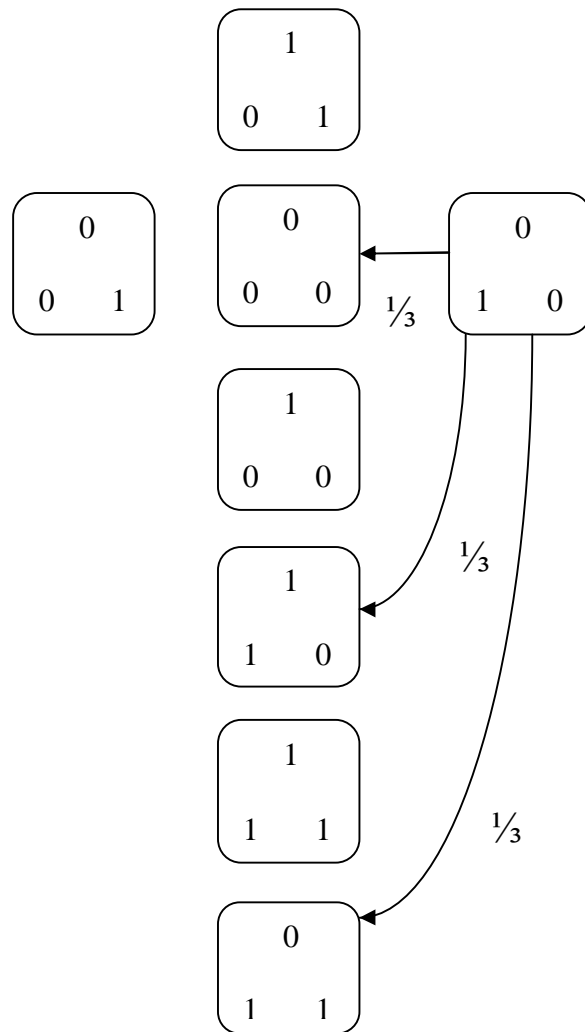
Η ενεργοποίηση είναι ίση με το μηδέν και συνεπώς, ($F_2 = 0$). Έτσι, η επόμενη κατάσταση από την 010 είναι η 000, αν ο F_2 επιχειρήσει να εκπέμψει και πάλι με πιθανότητα $1 / 3$.

Αν ο νευρώνας 3 είναι ο επόμενος για να προσπαθήσει να εκπέμψει, η ενεργοποίησή του είναι:

$$F_1W_{31} + F_2W_{32} - T_3 = (0 * (-0.5)) + (1 * 0.6) - 0 = 0.6$$

Η ενεργοποίηση είναι μεγαλύτερη από το μηδέν και συνεπώς, ($F_3=1$). Έτσι η επόμενη κατάσταση από το 010 είναι η 011 αν ο F_3 επιχειρήσει να εκπέμψει και πάλι με πιθανότητα $1 / 3$.

Ένα διάγραμμα που δείχνει όλες τις δυνατές καταστάσεις και τις μεταβάσεις δίνεται παρακάτω:



1.1.2. Ιδιαίτερο παράδειγμα:

Ολοκληρώστε το δίκτυο Hopfield και τις καταστάσεις μετάβασης του διαγράμματος που αναφέρονται ανωτέρω.

1.2. Ανάλυση του δικτύου Hopfield: Υπολογιστική ενέργεια

Για το παραπάνω παράδειγμα, θεωρήθηκε ότι το δίκτυο σταθεροποιεί τελικά στην τελική κατάσταση του 011. Ένα ιδιαίτερο χαρακτηριστικό των Hopfield δικτύων είναι ότι έχουν τελική κατάσταση(-εις) που είναι σταθερές, δηλαδή η πιθανότητα του δικτύου που παραμένει στην κατάσταση όπου είναι 1 είναι (Οριστική).

Είναι χρήσιμο να συνδέσει μια ποσότητα με κάθε κατάσταση στο δίκτυο, η οποία μειώνεται (ή διατηρείται ίδια) κάθε φορά που ένας νευρώνας αλλάζει κατάσταση. Η ποσότητα αυτή ονομάζεται ποσότητα υπολογιστικής ενέργειας και δίνεται ως σύμβολο E . Έτσι κάθε μετάβαση στην κατάσταση διαγράμματος έχει ως αποτέλεσμα να μειώνεται η E ή να μένει η ίδια. Κάθε μεταβατική κατάσταση περιλαμβάνει μεταβολή της κατάστασης ενός μόνο νευρώνα, εξαιτίας της ασύγχρονης λειτουργίας του δικτύου. Συνεπώς η μεταβολή της ενέργειας ΔE λόγω αλλαγής στην κατάσταση της F_i εξόδου του νευρώνα i , πρέπει να είναι αρνητική ή μηδενική. Ο νευρώνας i αλλάζει κατάσταση σε μια από τις δύο προϋποθέσεις:

1) $F_i = 0$ προτού η μετάβαση και η ενεργοποίηση είναι θετική. Σε αυτή την περίπτωση το F_i αλλάζει σε 1 και η αλλαγή του F_i αποκαλείται ΔF_i και είναι +1.

Ή

2) $F_i = 1$ προτού η μετάβαση και η ενεργοποίηση είναι αρνητική. Σε αυτή την περίπτωση οι αλλαγές είναι F_i σε 0 και του $\Delta F_i = -1$.

Θέλουμε αυτή η αλλαγή κατάστασης να ανταποκρίνεται σε μια αρνητική αξία του ΔE , έτσι ώστε η μετάβαση να έχει ως αποτέλεσμα τη μείωση της συνολικής ενεργειακής E . Αυτό μπορεί να επιτευχθεί με τον καθορισμό ΔE όπως:

$$\begin{aligned}\Delta E &= - \Delta F_i * \text{Ενεργοποίηση} \\ &= - \Delta F_i \left(\sum W_{ij} F_j - T_i \right)\end{aligned}$$

Χρησιμοποιώντας αυτό τον ορισμό το ΔE εγγυάται ότι κάθε αλλαγή κατάστασης αντιστοιχεί σε αρνητική (ή μηδενική) αξία του ΔE , μειώνοντας έτσι την ενέργεια του δικτύου.

Είναι λοιπόν σκόπιμο να οριστεί η ενέργεια του νευρώνα i ως:

$$\begin{aligned}E_i &= - F_i \left(\sum W_{ij} F_j - T_i \right) \\ &= - \sum W_{ij} F_j F_i + F_i T_i\end{aligned}$$

Η συνολική ενέργεια E του δικτύου συνδέεται με το άθροισμα της ενέργειας για όλους τους νευρώνες. Η E Hopfield ορίζεται ως εξής:

$$E = -\frac{1}{2} \sum_i \sum_j W_{ij} F_j F_i + \sum_i F_i T_i$$

Ο μισός όρος συμπεριλαμβάνεται στο λογαριασμό του γεγονότος ότι ορισμένοι από τους όρους λάβανε μέρος δύο φορές κατά τη λήψη συνολικών αθροίσεων (θυμηθείτε ότι $W_{ij} = W_{ji}$). Έτσι μπορεί να θεωρηθεί ότι το δίκτυο Hopfield εγκαθίσταται σε μία σταθερή κατάσταση και κάθε μετάβαση καταστάσεων μειώνει την ενέργεια (ή μετακινείται σε μία κατάσταση με την ίδια ενέργεια).

Το δίκτυο Hopfield είναι χρήσιμο για πρακτικές εφαρμογές, δεδομένου ότι παρέχει ένα μέσο βελτιστοποίησης. Για ένα δεδομένο πρόβλημα, αν μια έκφραση για την E μπορεί να οριστεί έτσι ώστε ένα ελάχιστο του E να αντιστοιχεί σε μια βέλτιστη ή «καλή» λύση στο πρόβλημα, τότε το Hopfield δίκτυο μπορεί να χρησιμοποιηθεί για την ελαχιστοποίηση της E και την «εύρεση», αυτής της καλής λύσης. Ένα παράδειγμα της χρήσης του δικτύου Hopfield είναι για την συμπίεση των ψηφιακών εικόνων που θα συζητηθεί αργότερα.

1.2.1. Παράδειγμα:

Επιστρέφοντας στο παράδειγμα στη σελίδα (9) μπορούμε να υπολογίσουμε τώρα την ενέργεια E του δικτύου για καθεμία από τις οκτώ καταστάσεις και να δούμε ότι κάθε μετάβαση κινείται σε μία κατάσταση με μικρότερη ή ίδια ενέργεια. Δεδομένου ότι το δίκτυο έχει τρεις νευρώνες, μπορούμε να ορίσουμε τα όρια των αθροίσεων διάστημα από 1 έως 3, δηλαδή:

$$E = -\frac{1}{2} \sum_{i=1}^3 \sum_{\substack{j=1 \\ j \neq i}}^3 W_{ij} F_j F_i + \sum_{i=1}^3 F_i T_i$$

Η ενέργεια για την κατάσταση 101 φαίνεται να είναι:

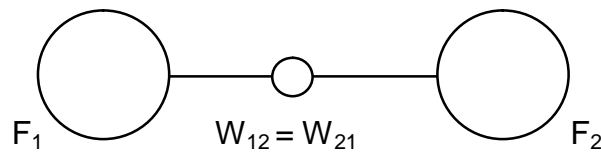
$$E_{101} =$$

1.2.2. Ιδιαίτερα παραδείγματα:

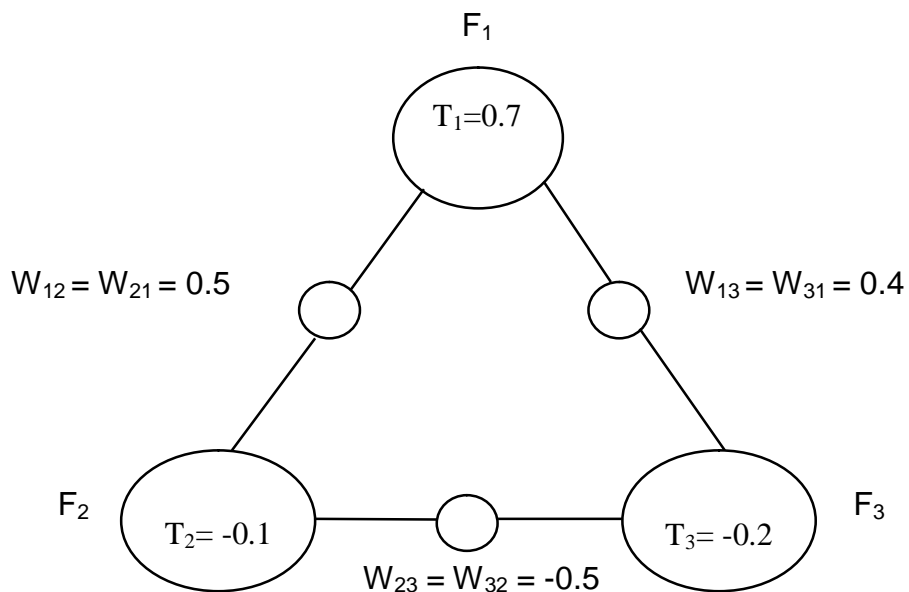
1) Ολοκληρώστε το παραπάνω διάγραμμα υπολογίζοντας την υπολογιστική ενέργεια συνδεδεμένης με καθεμία από τις οκτώ καταστάσεις.

2) Οι δύο νευρώνες Hopfield δικτύου που αναφέρονται κατωτέρω έχουν βάρη $W_{12}=W_{21}=0.4$.

Το κατώφλι για τον νευρώνα 1 είναι $T_1 = 0$ και για τον νευρώνα 2 είναι $T_2 = -0.2$. Σχεδιάστε το διάγραμμα μετάβασης καταστάσεων και υπολογίστε την υπολογιστική ενέργεια για καθεμία από τις τέσσερις καταστάσεις.

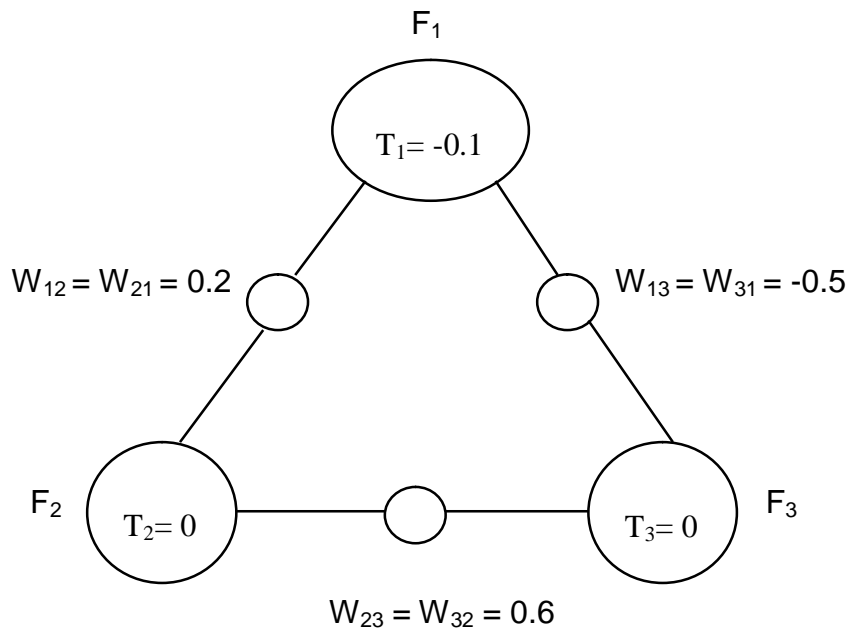


3) Σχεδιάστε το διάγραμμα μετάβασης καταστάσεων και υπολογίστε την ενέργεια για κάθε κατάσταση των τριών νευρώνων του δικτύου Hopfield που φαίνεται παρακάτω.



1.3. Hopfield δίκτυα και σταθερές καταστάσεις:

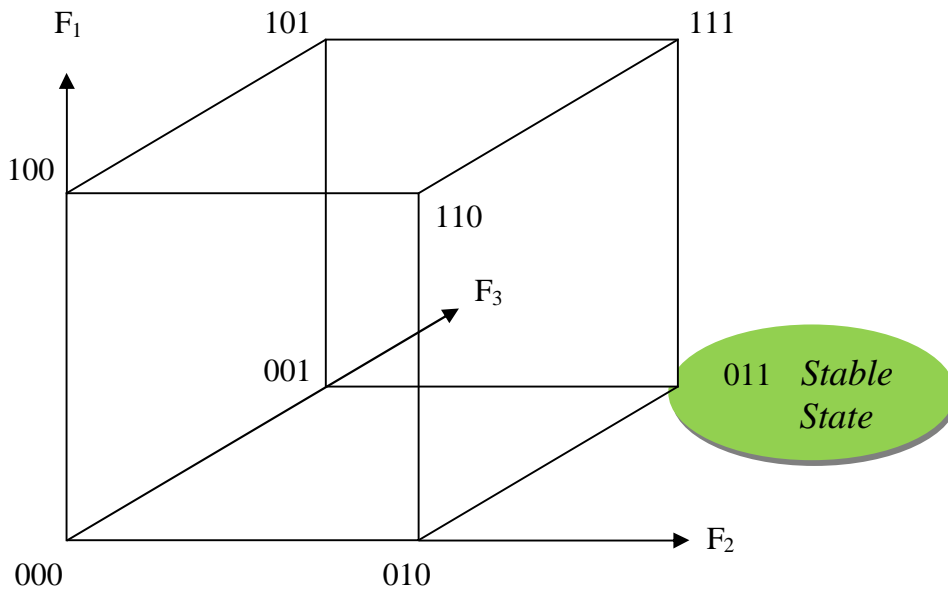
Για το παράδειγμα που μελετήθηκε νωρίτερα (9) θεωρήθηκε ότι το δίκτυο Hopfield τριών νευρώνων που φαίνεται πιο κάτω πάντα σταθεροποιείται στην κατάσταση 011.



Είναι συνεπώς δυνατό να δούμε ότι για το συγκεκριμένο δίκτυο, η τελική κατάσταση (μετά από επαναλήψεις έχουν οδηγήσει σε μια σταθερή κατάσταση) έχει μια πιθανότητα 100% να είναι 011, ανεξάρτητα από την κατάσταση εκκίνησης. Ο παρακάτω πίνακας δείχνει την πιθανότητα από όλες τις τελικές καταστάσεις για κάθε πιθανή κατάσταση εκκίνησης.

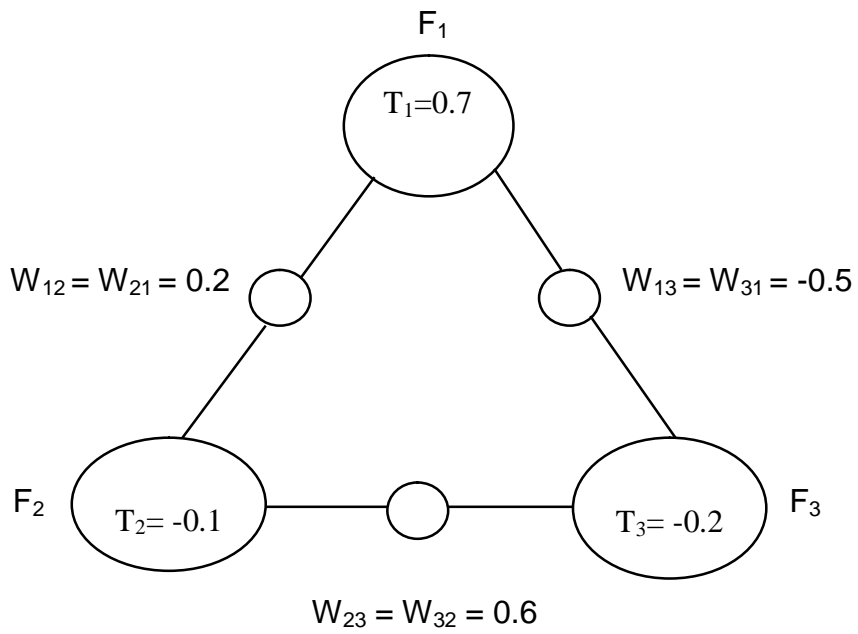
Αρχική Κατάσταση			Τελική Κατάσταση			% Πιθαν.
F_1	F_2	F_3	F_1	F_2	F_3	
0	0	0	0	1	1	100
0	0	1	0	1	1	100
0	1	0	0	1	1	100
0	1	1	0	1	1	100
1	0	0	0	1	1	100
1	0	1	0	1	1	100
1	1	0	0	1	1	100
1	1	1	0	1	1	100

Κάθε μετάβαση από το δίκτυο έχει ως αποτέλεσμα την αλλαγή μιας μόνο εξόδου των νευρώνων. Όπως και με το νευρώνα MCP, ενδέχεται να δούμε όλες τις δυνατές καταστάσεις του δικτύου, όπως γωνίες ενός κύβου στο τρισδιάστατο χώρο εισόδου. Σημειώστε τώρα ότι οι είσοδοι σε ένα νευρώνα είναι οι εξοδοί των άλλων νευρώνων, έτσι ώστε ο χώρος μπορεί να ονομαστεί σωστά «χώρος εξόδου» ή «χώρος κατάστασης». Κάθε διάσταση του κύβου αντιπροσωπεύει μία από τις καταστάσεις ανάλογα με τις τιμές της F_1 , F_2 και F_3 .



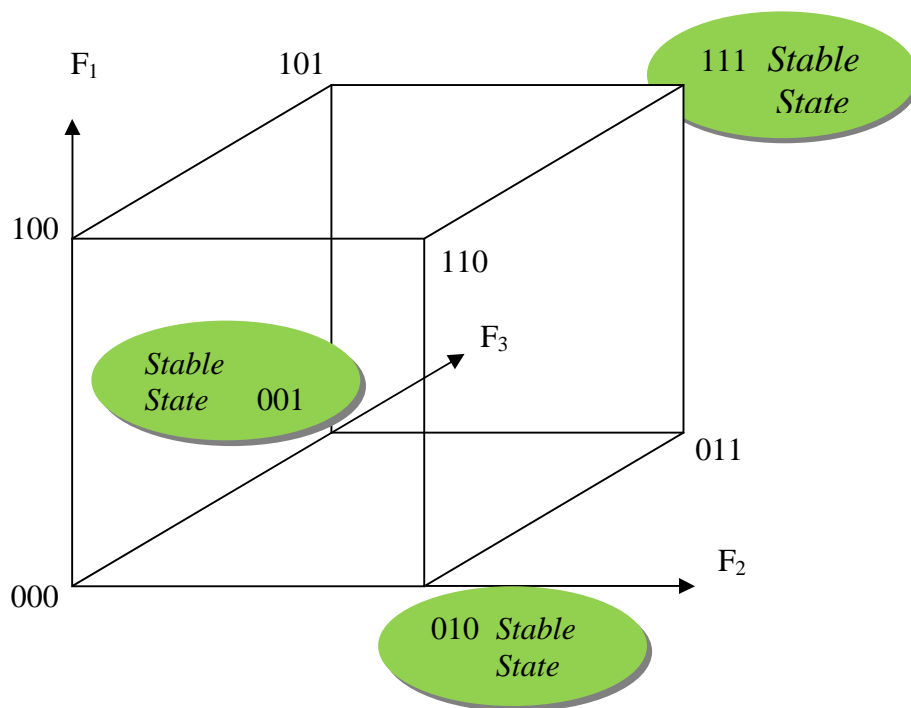
Κάθε μεταβατική κατάσταση αντιστοιχεί σε κίνηση από την παρούσα γωνία σε μία από τις τρεις πιθανές διπλανές κορυφές των γωνιών, δεδομένου ότι μόνο ένας από τους νευρώνες αλλάζει κατάσταση ανά πάσα στιγμή (ασύγχρονη λειτουργία). Δεδομένου ότι για αυτό το παράδειγμα υπάρχει μόνο μία σταθερή κατάσταση, το δίκτυο θα πρέπει τελικά να «επιτευχθεί» σε αυτή την κατάσταση. Τα μονοπάτια μπορούν επίσης να αντληθούν σχετικά με το παραπάνω διάγραμμα σε σχέση με το διάγραμμα μετάβασης καταστάσεων στη σελίδα (4).

Ωστόσο, το δίκτυο τριών νευρώνων από το ιδιαίτερο παράδειγμα 1.2.2 σελίδα (15) (φαίνεται παρακάτω) έχει περισσότερες από μία σταθερή κατάσταση και έτσι η πιθανότητα επίτευξης μίας συγκεκριμένης τελικής κατάστασης είναι μικρότερη από 100% για ορισμένες καταστάσεις εκκίνησης. Αυτό φαίνεται στον παρακάτω πίνακα.



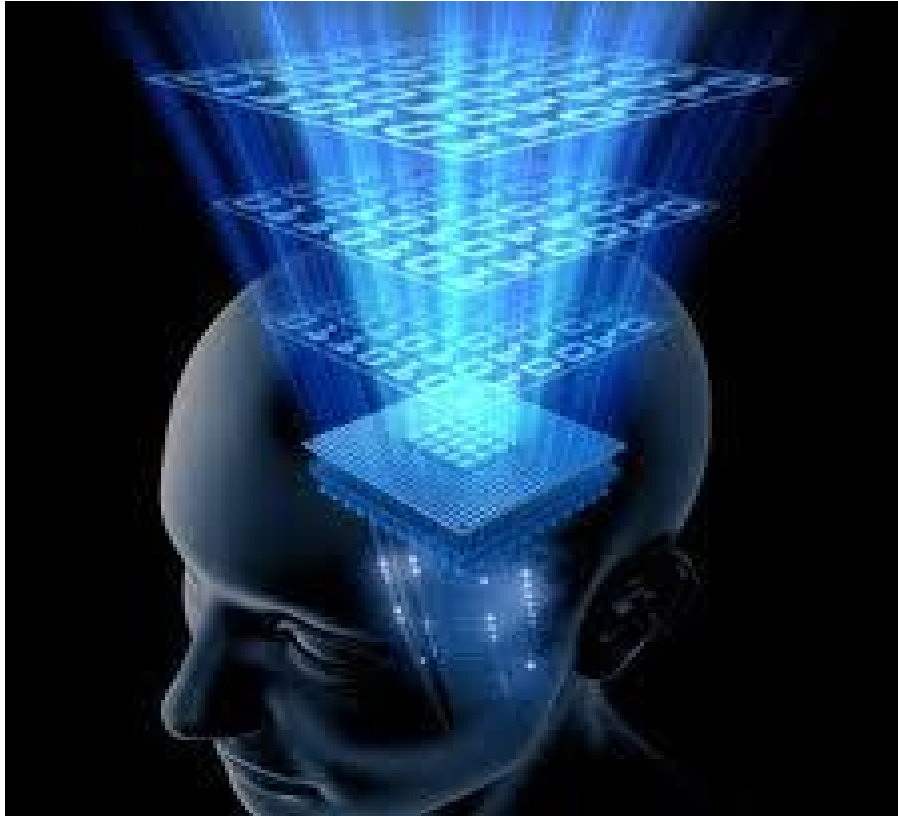
Αρχική κατάσταση			Τελική κατάσταση			% Πιθαν.
F_1	F_2	F_3	F_1	F_2	F_3	
0	0	0				
0	0	1				
0	1	0				
0	1	1				
1	0	0				
1	0	1				
1	1	0				
1	1	1				

Το αντίστοιχο διάγραμμα χώρου κατάστασης μπορεί να σχεδιαστεί:



1.4. Σχεδιάζοντας σταθερές καταστάσεις σε ένα δίκτυο Hopfield:

Ένα δίκτυο Hopfield είναι χρήσιμο να είναι σε θέση να σχεδιαστεί (δηλαδή υπολογισμός κατάλληλων τιμών για τα βάρη και τα κατώφλια), ιδίως όταν υπάρχει προκαθορισμένη σταθερή κατάσταση (εις). Όπως προαναφέρθηκε, αυτές οι σταθερές καταστάσεις αντιστοιχούν σε τοπικά ελάχιστα της υπολογιστικής ενέργειας, και κατά συνέπεια μερικές φορές αναφέρονται ως επίπεδα ενέργειας.



Ως παράδειγμα, θα εξετάσουμε το σχεδιασμό ενός δικτύου Hopfield τριών νευρώνων με δύο πιθανές σταθερές καταστάσεις, που επιλέγονται ως 001 και 111. Το δίκτυο θα πρέπει να είναι σε θέση να καταλήξει σε μία από αυτές τις καταστάσεις από οποιαδήποτε αρχική κατάσταση.

Πρέπει να εξετάσουμε κάθε νευρώνα για κάθε σταθερή κατάσταση, και να εξασφαλίσουμε ότι δεν αλλάζει - αυτό θα εξασφαλίσει ότι οι απαιτούμενες καταστάσεις είναι σταθερές.

Κατάσταση 001

F_1 δεν θα πρέπει να εκπέμπει, δηλαδή

$$W_{12}F_2 + W_{13}F_3 - T_1 \leq 0$$

Αλλά για κατάσταση 001, $F_2 = 0$ και $F_3 = 1$, και ως εκ τούτου μπορούμε να γράψουμε:

$$W_{13} - T_1 \leq 0 \quad (\text{Ανισότητα 1})$$

F_2 δεν θα πρέπει να εκπέμψει, με αποτέλεσμα

$$W_{23} - T_2 \leq 0 \quad (\text{Ανισότητα 2})$$

F_3 πρέπει να εκπέμψει, με αποτέλεσμα

$$-T_3 > 0$$

ή ισοδύναμα

$$T_3 < 0 \quad (\text{Ανισότητα 3})$$

Για την κατάσταση 001 να είναι σταθερή, όπως απαιτείται, οι τρεις ανισότητες παραπάνω πρέπει να την ικανοποιούν.

Κατάσταση 111

F_1 θα πρέπει να εκπέμψει:

$$\quad (\text{Ανισότητα 4})$$

F_2 θα πρέπει να εκπέμψει:

$$\quad (\text{Ανισότητα 5})$$

F_3 θα πρέπει να εκπέμψει:

$$\quad (\text{Ανισότητα 6})$$

Αν βεβαιωθούμε ότι πληρούνται οι έξι ανισότητες, στη συνέχεια, οι καταστάσεις 001 και 111 θα είναι σταθερές, όπως απαιτείται. Σε γενικές γραμμές, για ένα δίκτυο με N νευρώνες και M απαιτούμενες καταστάσεις, υπάρχουν ανισότητες που πρέπει να πληρούνται.

1.4.1. Ικανοποιώντας τις Ανισότητες

Είναι σαφές ότι υπάρχουν πολλές πιθανές τιμές για τα βάρη και τα κατώφλια τα οποία θα πληρούν τις παραπάνω έξι ανισότητες. Θα ερευνήσουμε ένα συγκεκριμένο σύνολο τιμών αρχικά: Θα προσπαθήσουμε να βρούμε ένα σύνολο τιμών, όπως ότι όλοι οι συντελεστές στάθμισης και τα κατώφλια να είναι μεταξύ - 1 και +1.

Ας πούμε ότι $W_{13} = 0.5$ άρα $W_{13} = W_{31} = 0.5$

Η ανισότητα 1 μπορεί πλέον να χρησιμοποιηθεί για να δώσει μια κλίμακα για το T_1 . Μπορεί να διαπιστωθεί ότι το T_1 (πρέπει να κυμαίνεται μεταξύ 0.5 και 1), για αυτό θα θέσουμε το T_1 , για παράδειγμα 0.7.

$$T_1 = 0.7$$

Η ανισότητα 4 δίνει μια σειρά για W_{12} , η οποία πρέπει να είναι μεταξύ

$$W_{12} = W_{21} =$$

Η ανισότητα 3 δίνει το T_3 :

$$T_3 =$$

Ανισότητα 6 δίνει το W_{32} :

$$W_{32} = W_{23} =$$

Οι ανισότητες 2 και 5, δείχνουν ότι η T_2 πρέπει να είναι της κλίμακας

$$T_2 =$$

Ελέγξτε τις έξι ανισότητες για να διασφαλιστεί ότι τηρούνται όλες οι απαραίτητες προϋποθέσεις:

Ανισότητα 1:

Ανισότητα 2:

Ανισότητα 3:

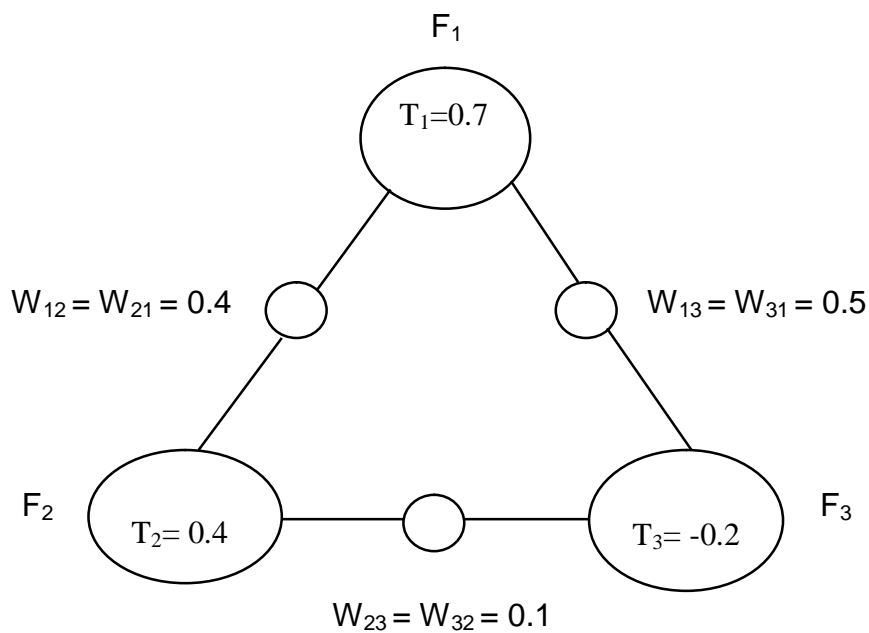
Ανισότητα 4:

Ανισότητα 5:

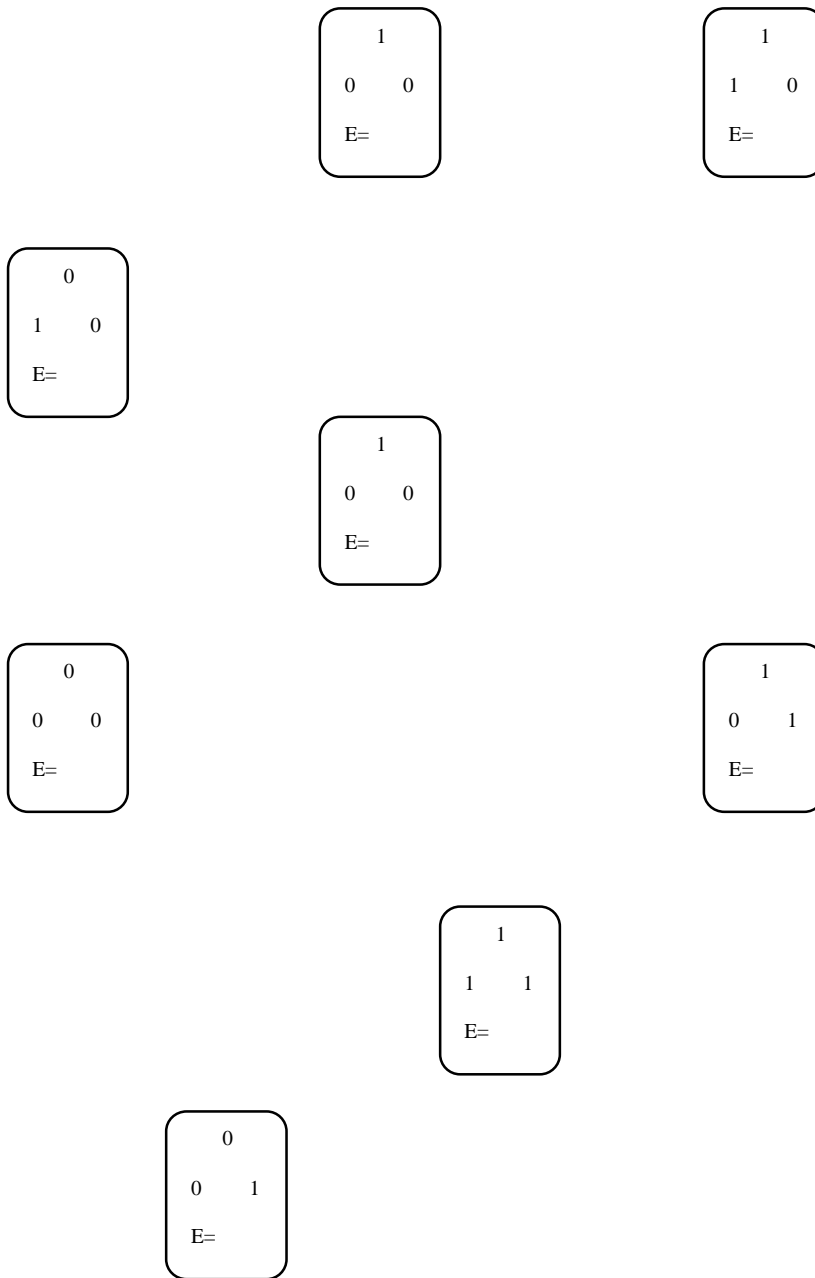
Ανισότητα 6:

Έτσι, μπορεί να διαπιστωθεί ότι τα βάρη και τα κατώφλια έχουν ικανοποιήσει τις έξι ανισότητες.

Ένα διάγραμμα μπορεί να συναχθεί δίνοντας τα βάρη του δικτύου και τα κατώφλια.



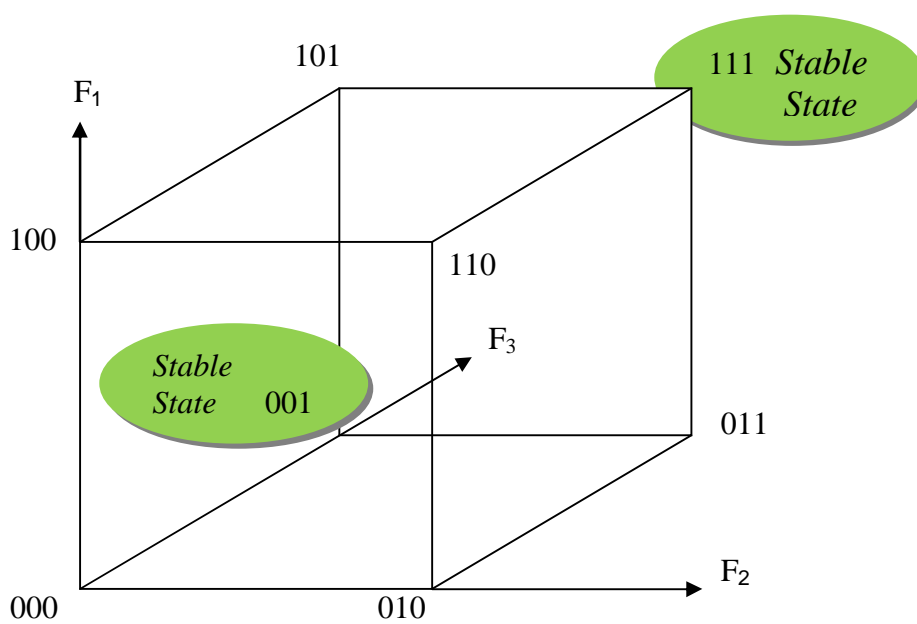
Μπορεί να σχεδιαστεί το διάγραμμα μετάβασης καταστάσεων για αυτό το δίκτυο και να βρεθεί η ενέργεια της κάθε κατάστασης.



Μπορεί να διαπιστωθεί ότι υπάρχουν δύο σταθερές καταστάσεις, όπως απαιτείται.

Το διάγραμμα χώρου καταστάσεων και τέλος πίνακα καταστάσεων πιο κάτω μπορεί να ολοκληρωθεί για να επεξηγήσει τη λειτουργία αυτού του δικτύου.

Αρχική κατάσταση			Τελική κατάσταση			% Πιθαν.
F ₁	F ₂	F ₃	F ₁	F ₂	F ₃	
0	0	0				
0	0	1				
0	1	0				
0	1	1				
1	0	0				
1	0	1				
1	1	0				
1	1	1				



1.4.2.Εναλλακτική λύση των ανισοτήτων

Ένα διαφορετικό σύνολο των βαρών και των κατωφλίων που ικανοποιεί τις ίδιες έξι ανισότητες παρατίθεται παρακάτω.

$$W_{12} = W_{21} = 0.5$$

$$W_{13} = W_{31} = 0.4$$

$$W_{23} = W_{32} = -0.5$$

$$T_1 = 0.7$$

$$T_2 = -0.1$$

$$T_3 = -0.2$$

Μπορεί να διαπιστωθεί ότι οι τιμές αυτές ικανοποιούν επίσης τις ανισότητες, ελέγχοντας και τις έξι:

Ανισότητα 1:

Ανισότητα 2:

Ανισότητα 3:

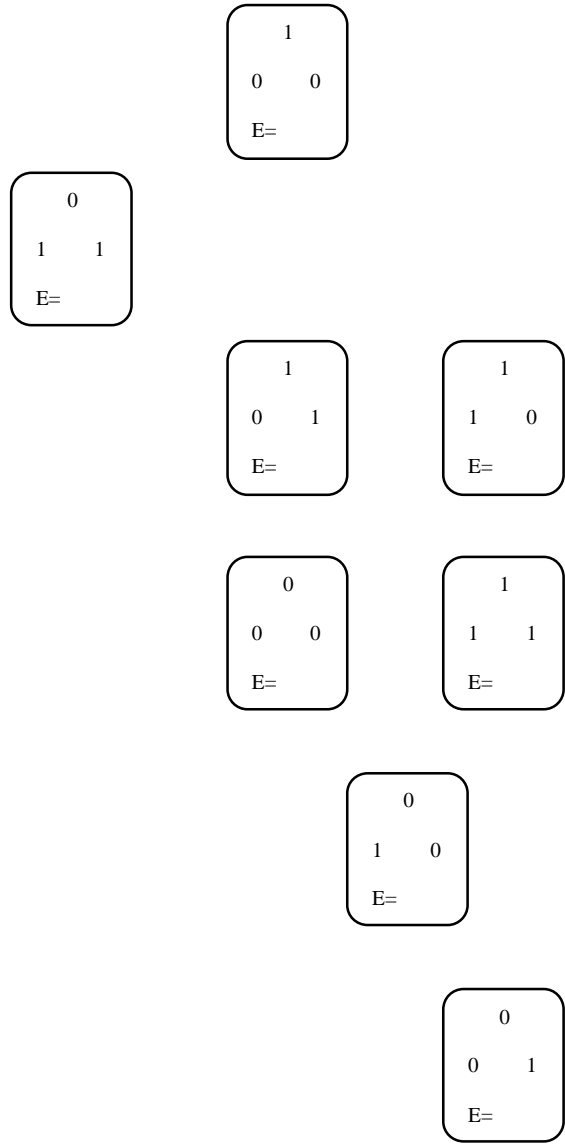
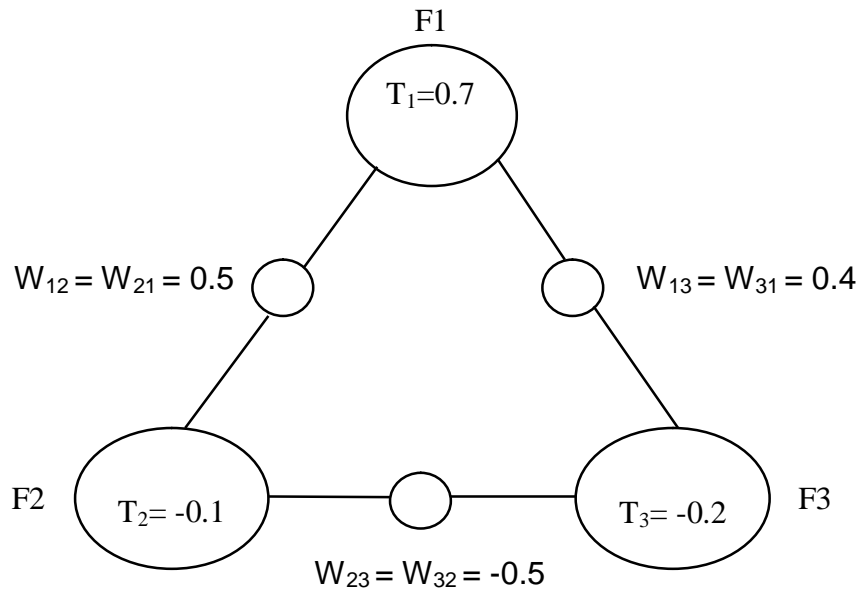
Ανισότητα 4:

Ανισότητα 5:

Ανισότητα 6:

Έτσι, μπορεί να διαπιστωθεί ότι τα βάρη και τα κατώφλια ικανοποιούν εκ νέου τις έξι ανισότητες.

Ωστόσο, αυτά τα βάρη και κατώτατα όρια ελήφθησαν από το δίκτυο της σελίδας (15) που έχει μια τρίτη σταθερή κατάσταση, που δεν κατονομάζεται στο σχεδιασμό. Το δίκτυο και το διάγραμμα καταστάσεως της μετάβασης δίνονται παρακάτω.

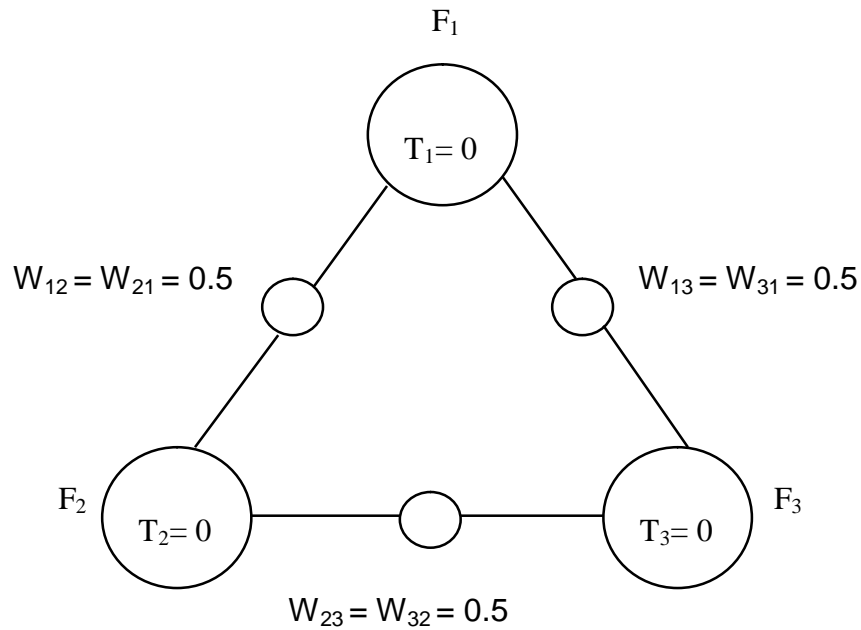


1.5. Η εκπαίδευση ενός δικτύου Hopfield για απαιτούμενες σταθερές καταστάσεις



Ως εναλλακτική λύση προς τη μέθοδο του σχεδιασμού δικτύου που περιγράφεται παραπάνω, μπορούμε να εκπαιδεύσουμε ένα δίκτυο Hopfield για να έχει τις απαιτούμενες σταθερές καταστάσεις. Αυτή η διαδικασία είναι παρόμοια με την κατάρτιση ενός ενιαίου νευρώνα MCP που έχουμε δει νωρίτερα. Όπως και πριν, η εκπαίδευση αποτελεί προσαρμογή των βαρών και κατωφλίων μέχρι το δίκτυο να εκτελεί, όπως απαιτείται. Αυτό μπορεί να επιτευχθεί με την εκπαίδευση αλγορίθμων που συζητήθηκε προηγουμένως. Οι ακόλουθες σημειώσεις καλύπτουν τη χρήση του Δέλτα Κανόνα(Delta Rule), όπως εφαρμόζεται στα Hopfield δίκτυα.

Και πάλι, θα εξετάσουμε το παράδειγμα ενός δικτύου Hopfield τριών νευρώνων με δύο απαιτούμενες σταθερές καταστάσεις: 001 και 111. Η αρχική διαμόρφωση (δηλαδή αρχικές τιμές για τα βάρη και τα κατώφλια) του δικτύου πρέπει να επιλεγεί - θα θέσουμε όλα τα κατώτατα όρια στο 0 και όλα τα βάρη σε 0.5 για αυτό το παράδειγμα. Η αρχική διαμόρφωση του δικτύου φαίνεται παρακάτω.



Εκπαίδευση επιτυγχάνεται με την εφαρμογή του κανόνα της εκπαίδευσης σε κάθε νευρώνα για κάθε μία από τις απαιτούμενες σταθερές καταστάσεις.

Θα καλέσουμε τις σταθερές καταστάσεις A (001) και B (111) και θα εφαρμόσουμε τον κανόνα με τη σειρά που περιέχονται στη λίστα παρακάτω:

- νευρώνας 1, σταθερή κατάσταση A
- νευρώνας 2, σταθερή κατάσταση A
- νευρώνας 3, σταθερή κατάσταση A
- νευρώνας 1, σταθερή κατάσταση B
- νευρώνας 2, σταθερή κατάσταση B
- νευρώνας 3, σταθερή κατάσταση B

Σημειώστε ότι ο κανόνας της εκπαίδευσης μπορεί να εφαρμοστεί στους νευρώνες και τις απαιτούμενες σταθερές καταστάσεις με άλλες σειρές (βλ. ιδιαίτερα παραδείγματα) - αυτό είναι μόνο μια επιλογή της σειράς που χρησιμοποιείται στο παρακάτω παράδειγμα.

Εάν διαπιστωθεί ότι ένας νευρώνας θα εκπέμψει όταν δεν πρέπει, ή δεν θα εκπέμψει όταν θα έπρεπε, στη συνέχεια για να γίνει προσαρμογή(διορθωτικός συντελεστής ΔW) θα πρέπει να εφαρμοστεί στους κατάλληλους συντελεστές

βαρών και κατωφλίων σύμφωνα με το δέλτα κανόνα. Θα πρέπει στη συνέχεια να επιστρέψουμε στην πρώτη θέση της λίστας και να συνεχίσουμε. Αν κανονίσουμε τις επαναλήψεις με αυτό τον τρόπο, όταν φτάσουμε στο τέλος της λίστας τότε γνωρίζουμε ότι όλες οι απαιτήσεις ικανοποιούνται.

Θα επιλέξουμε ένα «ρυθμός εκμάθησης» όπου $\alpha = 0.6$. Ο ΔW_i συντελεστής διόρθωσης, για κάθε βάρος που συνδέεται με νευρώνα i βρίσκεται με τον τύπο:

$$\Delta W_i = \alpha \delta$$

(Σημειώστε ότι ο X_i όρος παραλείπεται, από τη στιγμή που θα εφαρμόσουμε το συντελεστή διόρθωσης μόνο στα βάρη με εκπομπή εισόδου).

Έχοντας βρει το διορθωτικό συντελεστή, θα πρέπει να τον προσθέσουμε στα βάρη με εκπομπή εισόδου, και να τον αφαιρέσουμε το από το κατώφλι του νευρώνα υπό εξέταση.

Η εκπαίδευση (κατάρτιση) του δικτύου φαίνεται στα παρακάτω βήματα.



1) Κατάσταση A: 001

Εδώ βλέπουμε ότι η συνολική ενεργοποίηση για τον νευρώνα 1 είναι 0.5 και ως εκ τούτου ο νευρώνας 1 θα εκπέμπει. Ωστόσο, η απαιτούμενη ενέργεια του νευρώνα 1 δεν είναι για εκπομπή, οπότε κάποια διόρθωση είναι απαραίτητη.

$$\delta =$$

$$\Delta W_1 =$$

Έτσι έχουμε προσαρμόσει όλα τα βάρη στο νευρώνα 1, τα οποία προέρχονται από άλλους νευρώνες που εκπέμπουν - στην προκειμένη περίπτωση, αυτό

σημαίνει ότι έχουμε ρυθμίσει W_{13} . Μπορούμε επίσης να ρυθμίσουμε το κατώφλι T_1 , για νευρώνα 1. Οι νέες τιμές είναι:

$$W_{13} =$$

$$T_1 =$$

Δεδομένου ότι η διόρθωση έγινε σε αυτό το βήμα, πάμε πίσω στην αρχή της λίστας για το επόμενο βήμα (στην προκειμένη περίπτωση αυτό σημαίνει επανάληψη για νευρώνα 1 στην κατάσταση A).



2) Κατάσταση A: 001

Εδώ βλέπουμε ότι η συνολική ενεργοποίηση (συμπεριλαμβανομένου του κατωφλίου) στο νευρώνα 1 είναι -0.7 και ως εκ τούτου ο νευρώνας 1, δεν θα εκπέμψει (όπως απαιτείται). Έτσι $\delta = 0$ και δεν απαιτείται διόρθωση για αυτό το βήμα.



3) Κατάσταση A: 001

Εδώ βλέπουμε ότι η συνολική ενεργοποίηση στο νευρώνα 2 είναι 0.5 και ως εκ τούτου ο νευρώνας 2 θα εκπέμψει. Δεδομένου ότι ο νευρώνας 2 δεν είναι υποχρεωμένος να εκπέμψει, η διόρθωση είναι απαραίτητη.

$$\delta =$$

$$\Delta W_2 =$$

Έτσι θα πρέπει να εφαρμόζεται ο συντελεστής διόρθωσης για την W_{23} και T_2 . Οι νέες τιμές είναι:

$$W_{23} =$$

$$T_2 =$$

Δεδομένου ότι η διόρθωση έχει συμβεί, πάμε πίσω στην πρώτη καταχώρηση στη λίστα και συνεχίζουμε.



4) Κατάσταση A: 001

Και πάλι, η συνολική ενεργοποίηση στο νευρώνα 1 είναι -0.7 και ως εκ τούτου ο νευρώνας 1, δεν θα εκπέμψει (όπως απαιτείται). Έτσι, δεν χρειάζεται καμία διόρθωση για αυτό το βήμα. Παρά το γεγονός ότι αυτό το βήμα φαίνεται να είναι μια επανάληψη του βήματος 2, γενικά οι επιδράσεις της διόρθωσης μπορούν να αλλάξουν τα αποτελέσματα, και έτσι θα πρέπει να επιστρέψουμε στην κορυφή της λίστας κάθε φορά που επέρχεται.



5) Κατάσταση A: 001



6) Κατάσταση A: 001

Μπορεί να διαπιστωθεί ότι ο νευρώνας 3 δεν εκπέμπει όταν πρέπει.

$\delta =$

$\Delta W_3 =$

Κανένα από τα βάρη δεν είναι προσαρμοσμένο (δεδομένου ότι οι νευρώνες 1 και 2 δεν εκπέμπουν), αλλά το κατώφλι για τον νευρώνα 3 είναι.

$T_3 =$

↓

7) Κατάσταση A: 001

↓

8) Κατάσταση A: 001

↓

9) Κατάσταση A: 001

↓

10) Κατάσταση B: 111

$\delta =$

$\Delta W_1 =$

Πρέπει να εφαρμοστεί ο διορθωτικός συντελεστής για την W_{12} , W_{13} και T_1 . Οι νέες τιμές είναι:

$W_{12} =$ $W_{13} =$ $T_1 =$

↓

11) Κατάσταση A: 001

$\delta =$

$\Delta W_1 =$

Πρέπει να εφαρμοστεί ο διορθωτικός συντελεστής για την W_{13} και T_1 . Οι νέες τιμές είναι:

$W_{13} =$

$T_1 =$



12) Κατάσταση A: 001



13) Κατάσταση A: 001



14) Κατάσταση A: 001



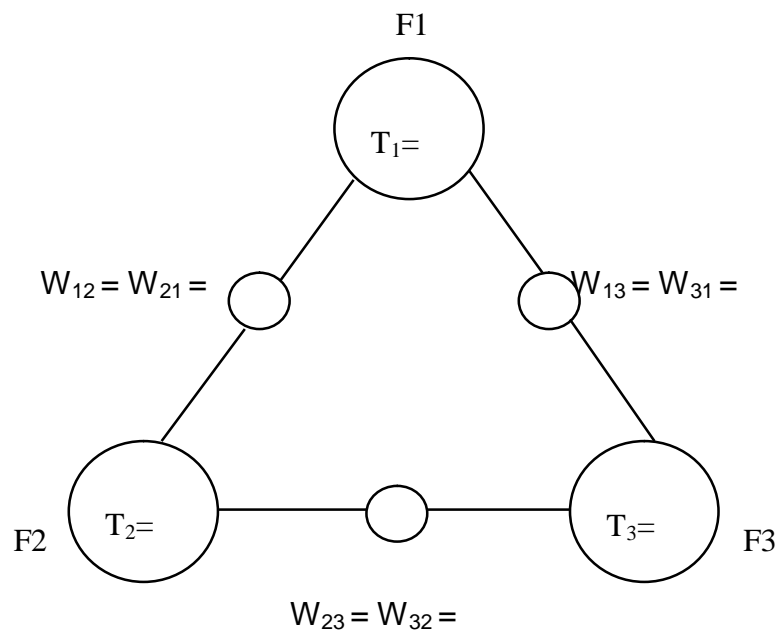
15) Κατάσταση B: 111



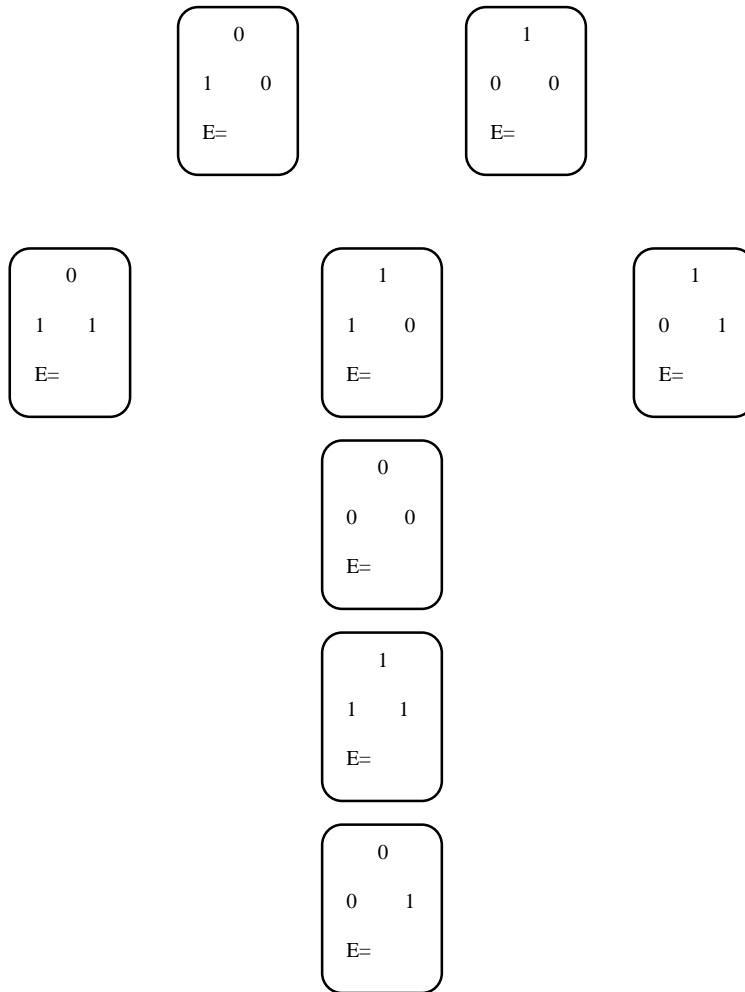
16) Κατάσταση B: 111

↓
17) Κατάσταση B: 111

Έχουμε ολοκληρώσει τώρα τον κατάλογο των συνθηκών, και έτσι το δίκτυο είναι εκπαιδευμένο. Η τελική διαμόρφωση του δικτύου παρατίθεται κατωτέρω.



Η λειτουργία του δικτύου είναι δυνατόν να επιβεβαιωθεί όπως απαιτείται από τη διερεύνηση της μετάβασης καταστάσεων. Ως συνήθως, θα κανονίσουμε το διάγραμμα μετάβασης καταστάσεων να είναι τέτοιο ώστε οι καταστάσεις με χαμηλή ενέργεια να εμφανίζονται προς τα κάτω.



Σημειώστε ότι είτε η μέθοδος του σχεδιασμού ενός δικτύου Hopfield μπορεί να παράγει ανεπιθύμητα επίπεδα ενέργειας εκτός από τα απαιτούμενα. Ορισμένοι συνδυασμοί των απαιτούμενων σταθερών καταστάσεων είναι αδύνατον να επιτευχθούν - για παράδειγμα, βλέπε ερώτημα 2 για τα δίκτυα Hopfield εκχώρησης. Όπως προαναφέρθηκε, επιπλέον νευρώνες μπορούν να συμπεριληφθούν στο δίκτυο για να ξεπεραστεί αυτό το τελευταίο πρόβλημα - βλέπε την εκχώρηση.

1.6. Περίπτωση μελέτης:

Μια πρακτική εφαρμογή των Hopfield δικτύων

Οι σημειώσεις που ακολουθούν δίνουν μια συνοπτική εισαγωγή στη συμπύεση εικόνας, με κάποιες λεπτομέρειες για την εφαρμογή μιας τεχνικής συμπύεσης εικόνας βάση των Hopfield νευρωνικών δικτύων.

Ουσιαστικά, ο στόχος της συμπίεσης εικόνας είναι η μείωση του όγκου των δεδομένων (bits) που απαιτούνται ώστε να αντιπροσωπεύει μια ψηφιακή εικόνα. Μια 256 x 256 pixel (εικονοστοιχείων) εικόνας (για παράδειγμα), με κάθε pixel εκπροσωπούμενο από 8 bit (256 διαφορετικά επίπεδα του γκριζου) απαιτεί πάνω από 500.000 κομμάτια (bits) (64 Kbyte) για να αντιπροσωπεύσουν πλήρως την εικόνα. Είναι σαφές ότι σε ορισμένες εφαρμογές, π.χ. αρχειοθέτησης εικόνας, ιατρική απεικόνιση, δορυφορική επεξεργασία εικόνας κλπ., το ποσό της αποθήκευσης γίνεται σύντομα πρακτικά μεγάλο, δεδομένου ότι πολλές διαφορετικές εικόνες θα πρέπει να αποθηκευτούν. Οι απαιτήσεις αποθήκευσης (π.χ. χώρο στο δίσκο) για μια εικόνα μπορεί να μειωθούν με τη μείωση του αριθμού των κομματιών (bits) που απαιτούνται για να εκπροσωπηθεί η εικόνα - είναι γνωστή ως συμπίεση εικόνας.

Ένα άλλο κίνητρο για συμπίεση εικόνας είναι η αναμετάδοση. Αν μια εικόνα πρέπει να μεταδοθεί σε ένα κανάλι, τότε 64 Kbytes δεδομένων του πρέπει να διαβιβαστούν (για το μέγεθος και την ανάλυση που συζητείται ανωτέρω). Αυτό θα απαιτήσει ένα ορισμένο χρονικό διάστημα, λαμβάνοντας υπόψη ένα εύρος ζώνης καναλιού. Ωστόσο, εάν η ποσότητα των δεδομένων που απαιτούνται μπορεί να μειωθεί, τότε λιγότερος χρόνος απαιτείται για το ίδιο εύρος ζώνης (ή, εναλλακτικά, ο ίδιος χρόνος απαιτείται σε ένα κανάλι με μικρότερο εύρος ζώνης).

Τεχνικές συμπίεσης εικόνας μπορούν να ταξινομηθούν σε δύο τύπους:

1) Τεχνικές με μικρές ή χωρίς απώλειες (διατήρηση πληροφοριών). Όπως υποδηλώνει το όνομα, οι τεχνικές αυτές μπορεί να μειώσουν τον όγκο των δεδομένων που απαιτούνται για να αναπαραστήσει μια εικόνα, ενώ διατηρούν όλες τις πληροφορίες στην εικόνα. Έτσι, η ανακαινισμένη εικόνα είναι ένα ακριβές αντίγραφο του πρωτότυπου εικόνας (ασυμπίεστο) και η εισαγωγή της δεν έχει σφάλματα. Οι τεχνικές αυτές είναι χρήσιμες σε εφαρμογές όπου η απώλεια πληροφοριών είναι ανεπίτρεπτη, όπως π.χ. συμπίεση ιατρικών ή νομικών εικόνων.

2)Τεχνικές απωλειών, οι οποίες επίσης μειώνουν τον όγκο των δεδομένων που απαιτούνται ώστε να αντιπροσωπεύουν μία εικόνα, αλλά παράγουν μία ανακαινισμένη εικόνα η οποία είναι μια προσέγγιση προς το πρωτότυπο. Έτσι, έχουμε λιγότερη από την τέλεια ανασυγκρότηση, η οποία μπορεί να είναι αποδεκτή σε ορισμένες εφαρμογές, π.χ. τηλεόρασης, αποστολές φαξ κλπ. Η διαφορά μεταξύ της αρχικής (ασυμπίεστης) εικόνας και η ανακατασκευασμένη εικόνα μπορεί να θεωρηθεί ως πρόσθετος θόρυβος (ή σφάλμα) και έτσι μπορεί να προσδιοριστεί ποσοτικά για να διαμορφώσει ένα μέτρο της «ποιότητας» της αναπαράστασης.

Η θέσπιση στην κοινή δράση των σφαλμάτων σε μία ανακαινισμένη εικόνα είναι η «μέση τετραγωνική απόκλιση» (MSE), που ορίζεται ως εξής:

$$MSE = \frac{1}{N} \sum_{n=0}^{N-1} (x(n) - x_r(n))^2$$

όπου: $x(n)$ είναι οι τιμές pixel της αρχικής εικόνας,

$X_r(n)$ οι τιμές pixel της ανακατασκευασμένης εικόνας,

N είναι ο συνολικός αριθμός των pixels στην εικόνα.

1.6.1 Μπλοκ Αποκοπής Κωδικοποίησης-Block Truncation Coding (BTC) Σχηματικό συμπίεσης εικόνας

Μπλοκ αποκοπής κωδικοποίησης είναι ένα απλό σύστημα συμπίεσης εικόνας που λειτουργεί με μικρούς κύβους της συνολικής εικόνας. Θα εξετάσουμε ένα τετράγωνο κομμάτι μεγέθους 4 x 4 pixels, ένα παράδειγμα της οποίας παρατίθεται κατωτέρω.

184	180	174	163
183	181	176	163
181	177	176	162
180	176	173	162

Το Μπλοκ Αποκοπής Κωδικοποίησης(BTC)λειτουργεί κατατάσσοντας κάθε pixel στον πίνακα σε μία από τις δύο κατηγορίες (κατηγορία 0 ή κατηγορίας 1) σύμφωνα με την τιμή pixel. Η απλή μέθοδος ταξινόμησης που χρησιμοποιείται στο αρχικό σύστημα του πίνακα για την αλλαγή τιμής είναι η σύγκριση της τιμής κατωφλίου με τον μέσο όρο όλων των κατωφλίων. Κάθε τιμή pixel κάτω από το μέσο όρο κατωφλίου τοποθετείται στην κατηγορία 0, καθώς και κάθε τιμή pixel ανωτέρω αξίας, τοποθετείται στην κατηγορία 1.

Στον πίνακα όπως φαίνεται από τα παραπάνω, η μέση τιμή είναι 174.4 και έτσι τα εικονοστοιχεία (pixel) ταξινομούνται όπως φαίνεται παρακάτω. Αυτό το 1-bit ανά pixel κατάταξης είναι γνωστό ως bitmap.

Στη συγκεκριμένη περίπτωση, τα δέκα εικονοστοιχεία είναι στην κατηγορία 1 και έξι στην κατηγορία 0. Ο μέσος όρος κάθε μιας από τις κατηγορίες υπολογίζεται και στη συνέχεια προσαρμόζεται, ώστε να μπορεί να εκπροσωπείται με 8 bits. Σε αυτό το παράδειγμα, η κατηγορία των 0 pixel έχει μια μέση τιμή των 166 και της κατηγορίας 1 pixel έχει μέση τιμή 179.

Τα συμπιεσμένα δεδομένα για τον πίνακα είναι το bitmap που εμφανίζεται παραπάνω και οι μέσες τιμές των 8-bits για τις δύο κατηγορίες. Έτσι έχουμε μειώσει την ποσότητα των δεδομένων που απαιτούνται για να παρουσιαστεί ο πίνακας, με κόστος κάποια απώλεια πληροφοριών. Το ποσό κατά το οποίο ο πίνακας έχει συμπιεστεί (αναφέρεται ως ο λόγος συμπίεσης) μπορεί να βρεθεί, όπως φαίνεται παρακάτω:

Για έναν πίνακα των 16 pixel με γκρι επίπεδα αντιστοιχούν σε 8 bits:

Αποσυμπιεσμένος πίνακας

απαιτεί bits

Συμπιεσμένος πίνακας

απαιτεί bits

Σχέση συμπίεσης =

Ένας εναλλακτικός τρόπος για να εκφραστεί ο λόγος συμπίεσης είναι το bit rate (ρυθμός μετάδοσης) - εν προκειμένω ο μέσος αριθμός των bits που απαιτούνται για να αντιπροσωπεύσει ένα εικονοστοιχείο είναι 2 bits ανά pixel.

Ανασυγκρότηση του πίνακα επιτυγχάνεται με bits στην εικόνα κατά την τιμή που αντιστοιχεί στο μέσο όρο της αντίστοιχης κατηγορίας. Στο παραπάνω παράδειγμα, αυτό θα προκαλούσε την ανακατασκευή του πίνακα που φαίνεται παρακάτω.

Συγκρίνοντας αυτόν με τον αρχικό πίνακα, μπορεί να θεωρηθεί ότι είναι διαφορετικός και επομένως τα σφάλματα έχουν εισαχθεί στον ανακατασκευασμένο. Τα σφάλματα αυτά μπορεί να είναι ποσοτικά(μετρώντας τα) με τη χρήση του MSE - για παράδειγμα αυτό το MSE (μέση τετραγωνική απόκλιση) φαίνεται να είναι 15.3.

1.6.2. Δίκτυο Hopfield (BTC) Σχηματικό συμπίεσης εικόνας

Μπορούμε να χρησιμοποιήσουμε ένα δίκτυο Hopfield ως ένα εναλλακτικό τρόπο ταξινόμησης των pixels μέσα σε έναν πίνακα. Όπως προαναφέρθηκε,

το δίκτυο Hopfield σταθεροποιείται σε μια ελάχιστη τιμή της υπολογιστικής ενέργειας. Εάν ορίσουμε την υπολογιστική ενέργεια, όπως ότι ένα ελάχιστο αντιστοιχεί σε μια «καλή» κατάταξη, τότε το δίκτυο θα ταξινομήσει τα εικονοστοιχεία καλά. Το Hopfield έδειξε ότι τα βάρη του δικτύου W_{ij} και τα κατώφλια T_i μπορούν να βρεθούν χρησιμοποιώντας τον τύπο:

$$W_{ij} = \frac{\partial^2 E}{\partial F_i \partial F_j}$$

$$T_i = \frac{\partial E}{\partial F_i} \quad F_j=0$$

Για να χρησιμοποιήσουμε το δίκτυο Hopfield για τη συμπίεση εικόνας, χρειαζόμαστε ένα δίκτυο που αποτελείται από 16 νευρώνες να χαρακτηρίσει έναν πίνακα των 4 x 4 pixels. Κάθε νευρώνας συνδέεται με ένα συγκεκριμένο pixel στο μπλοκ και η αντίστοιχος έξοδοι του F_i νευρώνα (0 ή 1), αφού το δίκτυο έχει φτάσει σε μία σταθερή κατάσταση, δίνει την κατάσταση των pixels.

Κατάλληλος ορισμός της υπολογιστικής ενέργειας, η οποία «ενθαρρύνει», παρόμοια pixels να καταταγούν στην ίδια κατηγορία είναι:

$$E = \sum_{i=1} \sum_{j=1} F_i F_j (x(i) - x(j))^2 + \sum_{i=1} \sum_{j=1} (1-F_i)(1-F_j) (x(i) - x(j))^2$$

Ελαχιστοποιώντας την E γίνεται αντιστοίχιση των pixels με παρόμοιες τιμές στην ίδια κατηγορία (κατηγορία 0 ή κατηγορίας 1). Τα Pixels με μεγάλες διαφορές μεταξύ των τιμών τους, τοποθετούνται σε διαφορετικές καταστάσεις. Επομένως, οι συντελεστές βαρών των δικτύων και των κατωφλίων δίνονται από τον τύπο:

$$W_{ij} = -4 [x(i) - x(j)]^2$$

$$T_i = -2 \sum_{j=1}^M [x(i) - x(j)]^2$$

όπου M είναι ο αριθμός των pixels του μπλοκ ($M = 16$ στα δεδομένα παραδείγματα).

1.6.3. Παράδειγμα:

Ένας 4×4 πίνακας προερχόμενος από μια μεγαλύτερη εικόνα παρουσιάζεται παρακάτω (το ίδιο παράδειγμα, όπως πριν).

184	180	174	163
183	181	176	163
181	177	176	162
180	176	173	162

Χρησιμοποιώντας ένα δίκτυο Hopfield για την ταξινόμηση αυτών των 16 τιμών pixels, όπως περιγράφεται παραπάνω, παίρνουμε:

1	1	1	0
1	1	1	0
1	1	1	0
1	1	0	0

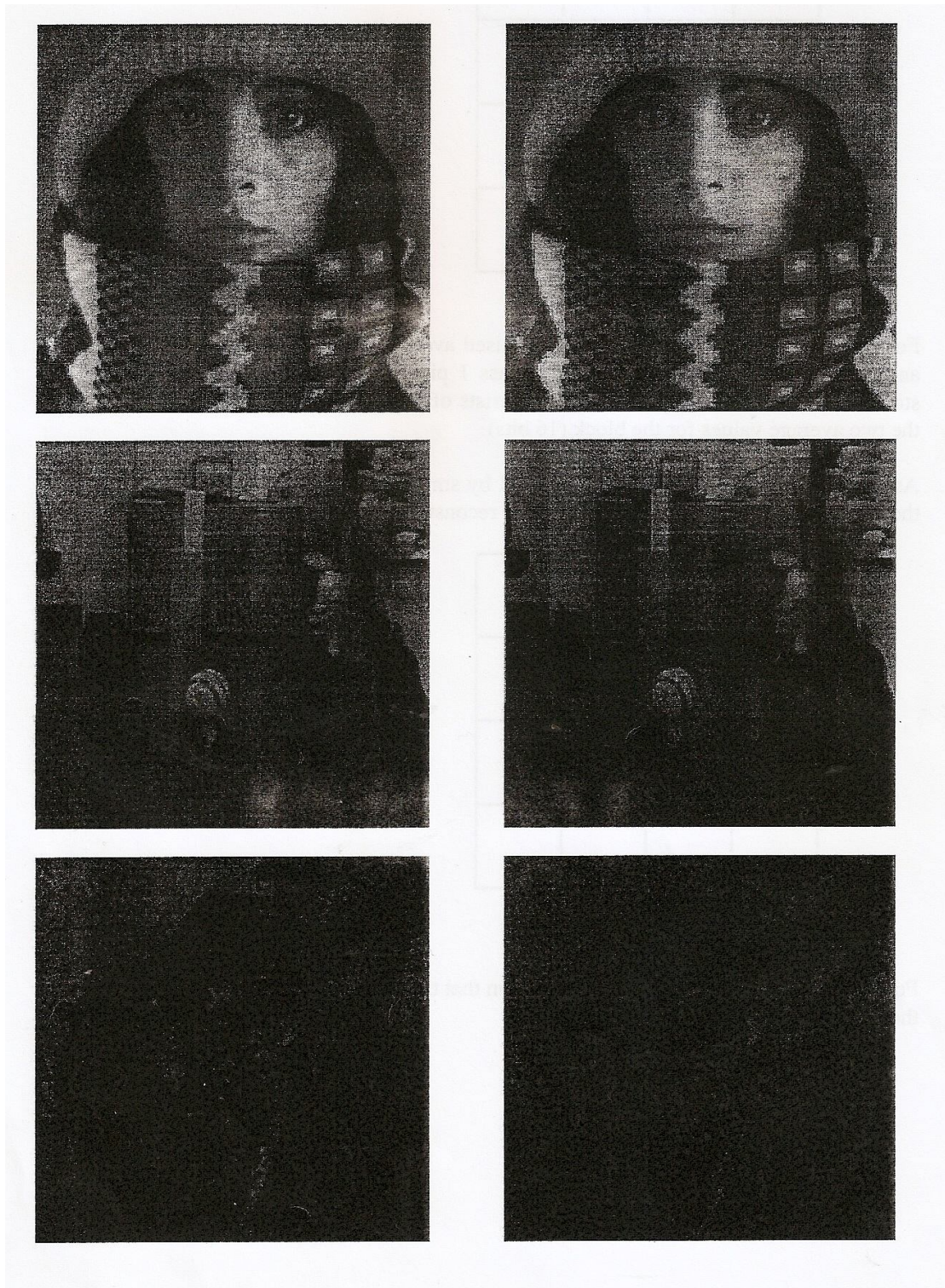
Για αυτό το παράδειγμα, διαπιστώνεται ότι η προσδιοριζόμενη ποσοτικά μέση τιμή της κατηγορίας 0 pixels είναι 164, και η προσδιοριζόμενη ποσοτικά μέση τιμή της κατηγορίας 1 pixels είναι 178. Και πάλι, ο μεταδιδόμενος ή αποθηκευμένος πίνακας (δηλ. τα συμπιεσμένα δεδομένα) αποτελούνται από το bitmap (16 bits), καθώς και τις δύο μέσες τιμές για τον πίνακα (16 bit).

Όπως και πριν, η ανοικοδόμηση πίνακα επιτυγχάνεται με την απλή αντικατάσταση κάθε bit στην εικόνα κατά την αντίστοιχη μέση τιμή, δίνοντας το ανακατασκευασμένο πίνακα που φαίνεται παρακάτω.

Για τον πίνακα στο παραπάνω παράδειγμα, είναι φανερό ότι η MSE (μέση τετραγωνική απόκλιση) είναι 12,9, σε σύγκριση με 15,3 για την αρχική μέθοδο.

Μια ολόκληρη εικόνα μπορεί να είναι συμπιεσμένη με την ίδια αναλογία συμπίεσης, συμπιέζοντας κάθε 4x4 πίνακες pixels ανεξάρτητα. Μερικά τυπικά αποτελέσματα φαίνονται παρακάτω.

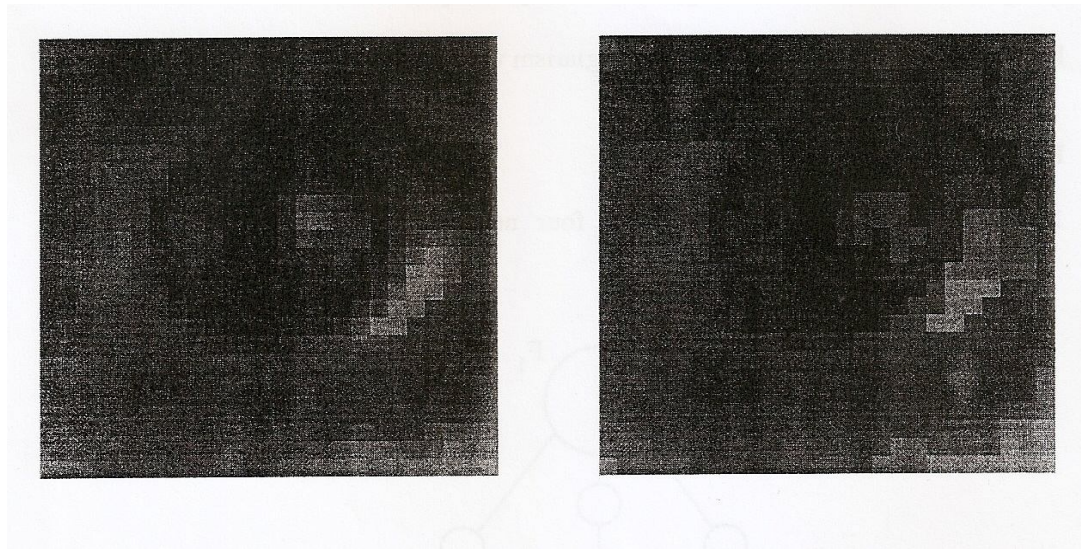
ΑΡΧΙΚΗ ΕΙΚΟΝΑ (256 x 256) ΑΝΑΣΥΓΚΡΟΤΗΜΕΝΗ ΕΙΚΟΝΑ (256 x 256)



Στα παραπάνω παραδείγματα, οι ανακατασκευασμένες εικόνες είναι παρόμοιες σε εμφάνιση με τις αντίστοιχες αρχικές εικόνες, ενώ η συμπίεσμένη έκδοση απαιτεί μόνο το ένα τέταρτο του χώρου στο δίσκο για την αποθήκευση.

Οι διαφορές που εισήγαγε το σχήμα συμπίεσης τις καταστεί σαφέστερες στις εικόνες σε μεγαλύτερη μεγέθυνση. Για παράδειγμα, ένα μικρό τμήμα της πρώτης εικόνας (25 x 25 pixels) παρουσιάζεται παρακάτω. Δεδομένου ότι τα pixel έχουν διευρυνθεί, οι διαφορές είναι σαφέστερες.

ΑΡΧΙΚΗ ΕΙΚΟΝΑ (256 x 256) ΑΝΑΣΥΓΚΡΟΤΗΜΕΝΗ ΕΙΚΟΝΑ (256 x 256)

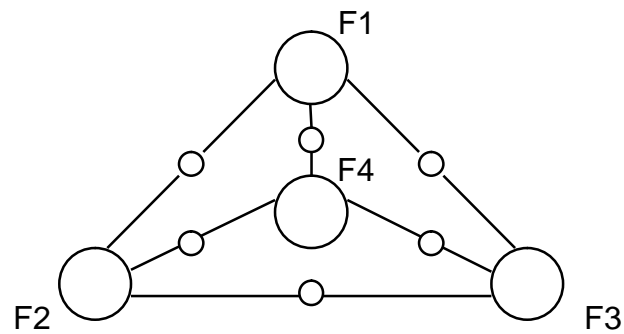


Παραλλαγές πάνω σε αυτή την τεχνική υπάρχουν, σύμφωνα με την οποία μεγαλύτεροι πίνακες μπορεί να χρησιμοποιηθούν σε περιοχές της εικόνας που περιέχουν μικρή λεπτομέρεια. Έτσι θα μπορούμε να χρησιμοποιούμε αρχικά 8 x 8 (pixels) για παράδειγμα, σε περιοχές όμως που έχουν ένα μεγάλο ποσό λεπτομέρειας μεγάλα λάθη θα εισαχθούν. Οι περιοχές αυτές μπορεί στη συνέχεια να επανεξεταστούν αποσυμπιεσμένα χρησιμοποιώντας 4 x 4 πίνακα. Σημειώστε ότι αν το μέγεθος του είναι αυξημένο, η αναλογία συμπίεσης αυξάνει το κόστος ενός υψηλότερου MSE (μέσης τετραγωνικής απόκλισης).

Για παράδειγμα, η αναλογία συμπίεσης για ένα ενιαίο μπλοκ 8 x 8 θα είναι:

1.6.4 Νευρωνικά δίκτυα Hopfield (παράδειγμα εκμάθησης)

1) Ένα νευρωνικό δίκτυο Hopfield με τέσσερις νευρώνες, όπως φαίνεται έχει τα βάρη και τα κατώφλια που ορίζονται κατωτέρω.



$$W_{11} = W_{22} = W_{33} = W_{44} = 0$$

$$T_1 = 0.70$$

$$W_{12} = W_{21} = 0.70$$

$$T_2 = -0.35$$

$$W_{13} = W_{31} = 0.35$$

$$T_3 = -0.38$$

$$W_{14} = W_{41} = 1.05$$

$$T_4 = 0$$

$$W_{23} = W_{32} = -0.38$$

$$W_{24} = W_{42} = -1.40$$

$$W_{34} = W_{43} = -1.40$$

Κεφάλαιο 2^ο: Η δομή της ανάδρασης για την εξέλιξη των νευρωνικών δικτύων με εγγενής συμπεριφορές μάθησης



Η Αμφίδρομη δομή της εξέλιξης των νευρωνικών δικτύων με εγγενής συμπεριφορές μάθησης

Περίληψη:

Το παρόν έγγραφο περιγράφει την ενότητα της μεταβλητής ανάδρασης: έναν απλό τρόπο υλοποίησης πίνακα που επιτρέπει την εξέλιξη των νευρωνικών δικτύων με εγγενή επίβλεψη ικανότητας μάθησης. Αυτό διευρύνει το φάσμα των δικτύων που μπορεί να εξελιχθούν αποτελεσματικά σε σύγκριση με προηγούμενες προσεγγίσεις και επίσης δίνει τη δυνατότητα στα δίκτυα να είναι αναστρέψιμα, δηλαδή το δίκτυο έχει εξελιχθεί για ένα συγκεκριμένο τομέα προβλημάτων και εκπαιδεύεται σε ένα συγκεκριμένο σύνολο δεδομένων. Το δίκτυο μπορεί στη συνέχεια να τρέξει προς τα πίσω για να παρατηρήσει τι είδους χαρτογράφηση έχει μάθει, ή για χρήση σε προβλήματα ελέγχου. Μια επίδειξη δίνεται από το είδος της αυτό-κατάρτισης δικτύου που θα μπορούσε να εξελιχθεί.



2.1. Εισαγωγή:

Παρά τη σημαντική έρευνα που εξελίσσεται στα νευρωνικά δίκτυα, υπάρχει σχετικά μικρή δουλειά που δημοσιεύθηκε στις εξελισσόμενες συμπεριφορές εκμάθησης για τα νευρωνικά δίκτυα. Σε αυτόν τον τομέα μπορούμε να προσδιορίσουμε τέσσερα είδη προσέγγισης:

- Να γράψετε ένα σύνολο κανόνων εκμάθησης μεταβλητών παραμέτρων για μια δεδομένη αρχιτεκτονική δικτύου, να εξελιχθεί η επιλογή των κανόνων και των παραμέτρων τους.

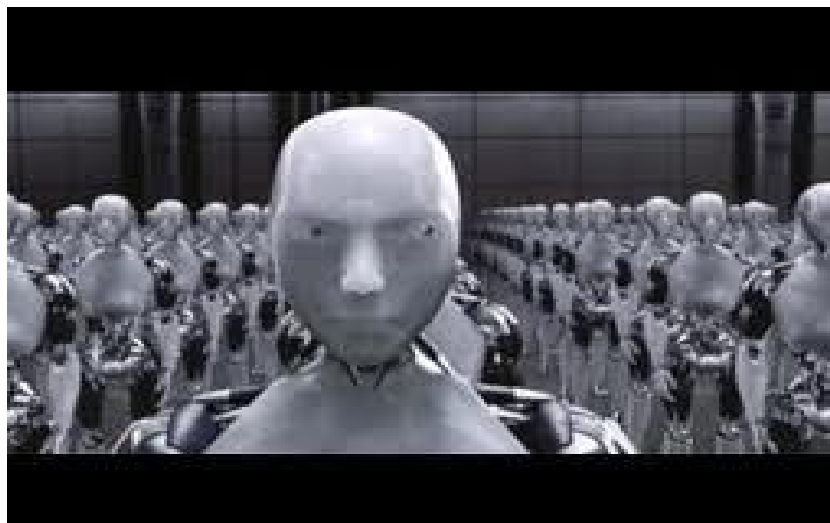
- Εξελίξτε κάποιο αυθαίρετο είδος του δικτύου, και ελπίζετε ότι θα μάθει.

- Να διευκρινιστεί το βήμα εκμάθησης μιας αρχιτεκτονικής, να επιτραπεί το μέρος της ανάδρασης (υπεύθυνη για την εκμάθηση), να εξελιχθεί με κάποιο αυθαίρετο τρόπο.

- Εξελίξτε κάποιο αυθαίρετο είδος του δικτύου, αλλά με κάποιους ειδικούς περιορισμούς που ωθούν τον τύπο εκμάθησης συμπεριφοράς να κατασκευαστούν σε αυτό.

Παραδείγματα της πρώτης κατηγορίας είναι το έργο του Chalmers [1] και επίσης του Bengio [2], ενώ ένα παράδειγμα της δεύτερης κατηγορίας μπορεί να βρεθεί στο [3]. Ο Συγγραφέας έχει ασχοληθεί με την τρίτη κατηγορία [4], ενώ το παρόν έγγραφο εισάγει την έννοια ενός συνόλου μεταβλητών για τη διευκόλυνση της τέταρτη έγκριση τύπου.

2.2. Υπόβαθρο:



Υπάρχουν δύο διαφορετικές μέθοδοι νευρωνικών δικτύων κωδικοποίησης στα χρωμοσώματα:

Άμεσες και έμμεσες. Σε άμεση περίπτωση, μπορούμε να εντοπίσουμε μια άμεση αντιστοιχία ανάμεσα σε κάθε τμήμα του χρωμοσώματος και κάθε μέρος του δικτύου. Η έμμεση μέθοδος κωδικοποίησης είναι εάν δεν υπάρχει αυτή η αντιστοιχία. Οι περισσότερες έμμεσες κωδικοποιήσεις που έχουν παρουσιαστεί στη βιβλιογραφία έχουν την ιδιότητα ότι τα μικρά χρωμοσώματα μπορούν να δημιουργήσουν μεγάλα δίκτυα.

Οι εξελικτικοί μέθοδοι έχουν χρησιμοποιηθεί πιο εκτεταμένα σε συνδυασμό με άμεση κωδικοποίηση και υπάρχουν εκατοντάδες άρθρα στη βιβλιογραφία που έχουν υιοθετήσει την προσέγγιση αυτή. Μερικοί έχουν χρησιμοποιήσει μια σταθερή αρχιτεκτονική και εξελίχθηκε μόνο τα βάρη αυτής της αρχιτεκτονικής [3, 5], άλλοι έχουν εξελίξει τόσο τη δομή όσο και τα βάρη [6, 7]. Οι ενδιαφέρουσες νευρωνικά λύσεις έχουν εξελιχθεί σε προβλήματα στην πρόβλεψη χρονοσειρών [8, 9], παίζοντας Tic-Tac-toe [5] και στην εξελισσόμενη δυναμική των δικτύων τα οποία μαθαίνουν την παραγωγή απλών αλληλουχιών [3]. Αυτή η τελευταία προσέγγιση έχει ιδιαίτερη σημασία για τις εργασίες που αναφέρονται εδώ, δεδομένου ότι μια μαθησιακή συμπεριφορά έχει εξελιχθεί. Ωστόσο, λόγω της χρήσης άμεσης κωδικοποίησης, ο τύπος εκμάθησης συμπεριφοράς δεν μπορεί να μεταφερθεί σε άλλα προβλήματα διαφορετικού μεγέθους. Επιπλέον, ενώ η

άμεση μέθοδος κωδικοποίησης έχει την αρετή της απλότητας κλιμακώνεται κακώς, και ως εκ τούτου είναι κατάλληλη μόνο στην εξέλιξη σχετικά μικρών δικτύων.

Για το λόγο αυτό, έχει υπάρξει κάποια ενδιαφέρουσα εργασία σχετικά με την έμμεση κωδικοποίηση. Μερικές από τις πρώτες εργασίες στον τομέα αυτό αναφέρθηκαν από τον Kitano το 1990 [10] και έκτοτε ακολούθησε με ανώτερες γλώσσες γενιάς του δικτύου και γραμματικές που σχεδιάστηκαν από τον Gruau [11, 12], Boers και Kuiper [13], Muhlenbein [14,15] και Sharman et al [16]. Όλα αυτά όμως, χρησιμοποιήθηκαν στο GA (που λειτουργεί σε χορδές ή γραφικές παραστάσεις στη νευρική γλώσσα περιγραφής ή των χρωμοσωμάτων), για να καταλήξει σε ένα συρματωμένο νευρωνικό δίκτυο, ή να χρησιμοποιηθεί το GA για να καταλήξουμε σε μια καλή τοπολογία (ή μια καλή τοπολογία και ένα σύνολο βαρών), το οποίο στη συνέχεια εκπαιδεύεται από ανάδρομη αναπαραγωγή σφάλματος ή προσομοιωμένης πυράκτωσης.

Ο Kitano [17] έχει επίσης αναπτύξει ένα ενιαίο πλαίσιο εντός του οποίου η δομή του δικτύου και τα βάρη έχουν τη δυνατότητα να εξελίσσονται, και συνέχισε με μια πιο λεπτομερή βιολογικά προσομοίωση [18]. Ακόμη και στην [17] ωστόσο, παρά τους ισχυρισμούς ότι όλες οι λεπτομέρειες του δικτύου εξελίσσονται, αυτό δεν είναι ακριβώς αλήθεια, δεδομένου ότι η μάθηση του αλγόριθμου καθορίζεται εκ των προτέρων στην ανάδρομη αναπαραγωγή σφάλματος αν και η τοπολογία του δικτύου και τα αρχικά βάρη έχουν εξελιχθεί.

Σε αντίθεση με αυτό, ο συντάκτης [19, 20, 4], έχει δείξει πώς θα ήταν δυνατόν να εξελιχθεί ο αλγόριθμος εκμάθησης στο ίδιο ενοποιημένο πλαίσιο εντός του οποίου τα άλλα στοιχεία του δικτύου έχουν εξελιχθεί. Επίσης, με τις σχετικές εργασίες άξια ενδιαφέροντος είναι η Strongly Typed Genetic Programming της Μοντάνα [21], το οποίο χρησιμοποιεί για να εξελιχθούν προγράμματα που ρητά λειτουργούν στο διάνυσμα του πίνακα και στον κατάλογο των δομών δεδομένων, καθώς και απλούστερους τύπους δεδομένων. Στο πλαίσιο αυτό εξελίσσει (μεταξύ άλλων) τον αλγόριθμο για την ενημέρωση των εισόδων παρακολούθησης και εκτίμησης σε ένα φίλτρο Kalman.

Τέλος, σχετικά με το θέμα της κωδικοποίησης, παρουσιάζουμε κάποιες ιδέες όπου τεχνητά χρωμοσώματα που χρησιμοποιούνται για νευρωνική εξέλιξη αρχιτεκτονικής πρέπει να αποβλέπουν σε αυτά:

- Το νευροδομικό στοιχείο θα πρέπει να είναι αυτορυθμιζόμενο. Αυτό θα επιτρέψει την εξέλιξη να προχωρήσει μέσω της ενοποίησης των υφιστάμενων δομικών στοιχείων με νέους τρόπους, καθώς και την ανάπτυξη νέων δομικών στοιχείων.
- Το νευροδομικό στοιχείο θα πρέπει να είναι αναγνώσιμο από τον άνθρωπο - αφού πέρασε μεγάλος χρόνος εξέλιξης για μια νέα λύση σε ένα ενδιαφέρον πρόβλημα, θα ήταν κρίμα να μπορούσαμε να το εκτιμήσουμε μόνο κατά τη σύνδεση ζεύξης-πίνακα.
- Το νευροδομικό στοιχείο θα πρέπει να παραμετροποιηθεί. Ως εκ τούτου, έχοντας εξελίξει ένα νέο είδος αρχιτεκτονικής για μια κατηγορία προβλήματος, θα πρέπει να είμαστε σε θέση να την τροποποιήσουμε παραμετρικά για την αντιμετώπιση των προβλημάτων που σχετίζονται με διαφορετικό μέγεθος.

Η στοιχειοδότηση είναι ύψιστης σημασίας για έναν καλό μηχανολογικό σχεδιασμό. Τα συστήματα παρουσιάζουν πολύπλοκη και χρήσιμη συμπεριφορά η οποία επιτυγχάνεται συνήθως μέσω της αλληλεπίδρασης με απλούστερα στοιχεία. Πολλά από τα προηγούμενα έμμεσα νευρωνικά συστήματα κωδικοποίησης έχουν στηριχθεί σε στοιχειοδομική κατασκευή, αλλά πέτυχε μόνο σε πολύ περιορισμένη έννοια. Σκεφτείτε για παράδειγμα το σχέδιο κωδικοποίησης του Kitano [10]. Αυτό βασίζεται στο πλαίσιο χωρίς επανεγγραφή των πινάκων. Ενώ διαθέτει μια πολύπλοκη στοιχειοδότηση στο γονότυπο, τα περισσότερα από αυτά τα στοιχεία δεν μεταφέρονται στο φαινότυπο, το οποίο είναι το κρίσιμο μέρος όπου το δομικό στοιχείο απαιτείται. Κατά το σχεδιασμό χρωμοσωμάτων για την νευρωνική εξέλιξη του δικτύου το καλύτερο μέρος για να αναζητήσει έμπνευση κάποιος είναι στη μηχανική υλικού ή λογισμικού, όπου η αποτελεσματική χρήση της στοιχειοδότησης είναι υψίστης σημασίας. Δεν είναι αρκετό να επιτρέψει χρήσιμες υποδομές να επαναληφθούν, θα πρέπει επίσης να είναι κατάλληλα διασυνδεδεμένες μεταξύ τους.

2.3. Η μεταβλητή της ανάδρασης:



Το προτεινόμενο δομικό στοιχείο είναι με μεταβλητή ανάδρασης. Αυτό αποτελείται από δύο συγκεκριμένες συμπεριφορές: το εμπρόσθιο τμήμα και το πίσω κομμάτι. Οι λεπτομέρειες της κάθε μεταβλητής ανάδρασης θα μπορούσαν να εξελιχθούν, ή θα μπορούσαν να είναι προκαθορισμένες. Εδώ, δείχνουμε πώς να τις καθορίσετε χειροκίνητα.

Η μεταβλητή ανάδρασης υλοποιεί δύο συμπεριφορές ή λειτουργίες - μία συμπεριφορά προς τα εμπρός και μία προς τα πίσω συμπεριφορά. Κάθε συμπεριφορά επιτυγχάνεται με την απλή ενημέρωση ενός συνόλου κελιών σε μια συγκεκριμένη σειρά. Συμβατικούς αριθμούς κελιών υπολογίζει η συνάρτηση ταυτότητας, και ενεργεί ως τοπική αποθήκευση. Στα Σχήματα 1-5 είναι τα κύτταρα που φαίνονται με λευκούς κύκλους (εμπρός) και γκρι κύκλους (πίσω).

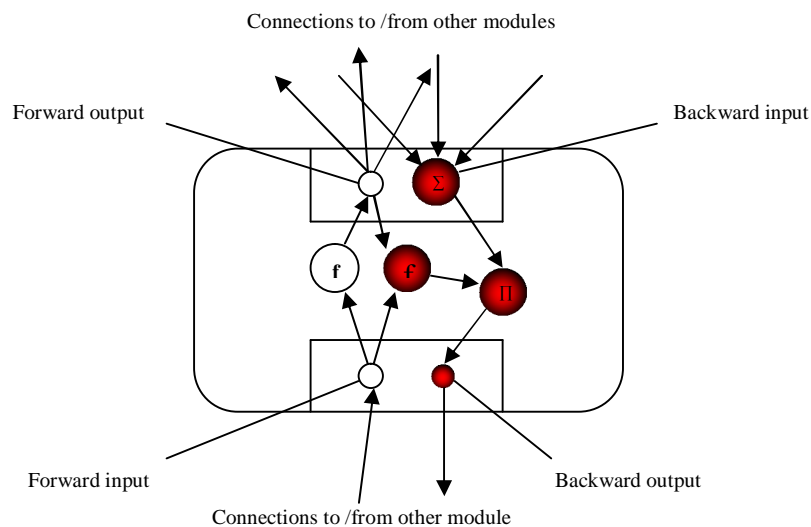
Ένα πλήρες δίκτυο με μεταβλητή ανάδρασης στη συνέχεια προσομοιώνεται σε λειτουργία προς τα εμπρός, με την εκτέλεση της δράσης κάθε μεταβλητής, από κάτω προς τα πάνω, όπως σκιαγραφείται στο Σχήμα 7. Παρομοίως, η προς τα πίσω λειτουργία κάθε μεταβλητής προσομοιώνεται εκτελώντας τις προς τα πίσω ενέργειες κάθε μεταβλητής από την κορυφή προς τα κάτω.

Οι μεταβλητές ανάδρασης μπορούν να είναι σχεδιασμένες για διάφορες συναρτήσεις, εδώ δίνουμε μόνο ένα μικρό δείγμα και ένα παράδειγμα εφαρμογής.

2.3.1. Ένας γενικός κόμβος ανάδρασης

Η γενική μορφή της μεταβλητής ανάδρασης παρουσιάζεται στο Σχήμα 1. Υπολογίζει συνάρτηση f για τη συμπεριφορά προς τα εμπρός. Για την προς τα πίσω συμπεριφορά υπολογίζει την παράγωγο f' της συνάρτησης σε σχέση με την είσοδο (ή την μερική παράγωγο wrt στη συγκεκριμένη είσοδο), και την πολλαπλασιάζει με το συσσωρευμένο σφάλμα της προς τα πίσω λειτουργίας (στο κελί Σ) και το περνά πίσω σαν έξοδο.

Σημειώστε ότι όλες οι συνδέσεις από και προς την εμπρός-πίσω μεταβλητή ομαδοποιούνται σε υποδοχές, όπου κάθε υποδοχή περιέχει δύο κελιά σύνδεσης, ένα για είσοδο και ένα για έξοδο. Η επάνω υποδοχή περιέχει την έξοδο προς τα εμπρός και την πίσω είσοδο, ενώ η κάτω υποδοχή περιλαμβάνει την είσοδο προς τα εμπρός και την πίσω έξοδο.



Εικ.1. Ένας γενικός κόμβος ανάδρασης. Οι λευκοί κύκλοι είναι οι κόμβοι που αξιολογήθηκαν κατά τη διάρκεια της λειτουργίας προς τα εμπρός, από κάτω προς τα πάνω. Οι σκιασμένοι κόμβοι που αξιολογήθηκαν στην πίσω λειτουργία περνούν από πάνω προς την κάτω σειρά.

2.3.2. Μεταβλητές ημίτονου και συνημίτονου:

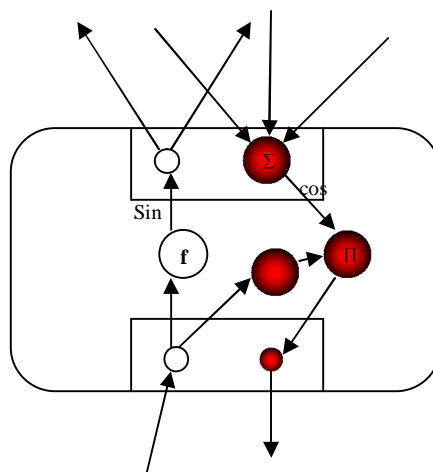
Οι μεταβλητές λειτουργίας ημίτονου και συνημίτονου δεν χρησιμοποιούνται συχνά σε τεχνητά νευρωνικά δίκτυα, αν και για πολλές εφαρμογές μπορούν να διαδραματίσουν χρήσιμο ρόλο. Στο Σχήμα.2. παρουσιάζεται η μεταβλητή ανάδρασης για την λειτουργία ημιτόνου. Η μεταβλητή ανάδρασης για συνημίτονο μπορεί να ληφθεί από την εναλλαγή των κόμβων ημίτονου και συνημίτονου και αναιρώνοντας την πίσω έξοδο.

Σχήμα.2. Από το μπροστινό πέρασμα ο κόμβος υπολογίζει το ημίτονο της μονής εισόδου και περνάει την έξοδο για (ενδεχομένως) πολλούς κόμβους. Από το πίσω πέρασμα, συσσωρεύεται ένα σήμα λάθους από όλους τους κόμβους με το οποίο η έξοδος συνδέεται. Πολλαπλασιάζει αυτό επί την παράγωγο της εξόδου σε σχέση με την είσοδο προς τα εμπρός και περνά πίσω στον κόμβο ότι έλαβε στοιχεία από την προς τα εμπρός είσοδο.

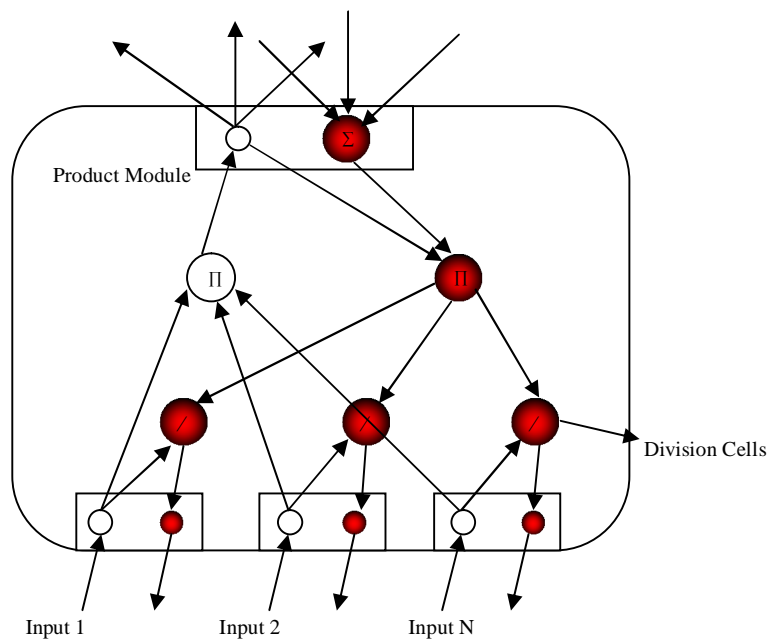
2.3.3. Παραγωγή και άθροιση μεταβλητών:

Ο παραγόμενος κόμβος είναι ένα παράδειγμα ενός κόμβου που μπορεί να πάρει πολλές εισόδους προς τα εμπρός. Δεδομένου ότι η μερική παράγωγος του σφάλματος δικτύου σε σχέση με κάθε είσοδο είναι διαφορετική, στην περίπτωση αυτού του τύπου κόμβου θα πρέπει η αντίστοιχη προς τα πίσω έξοδος να είναι διαφορετική.

Ο αθροιστικός κόμβος (Σχήμα 4), λαμβάνει επίσης πολλαπλές εισόδους, αλλά στην περίπτωση αυτή η πίσω είσοδος χορηγείται άμεσα σε όλες τις πίσω εξόδους.



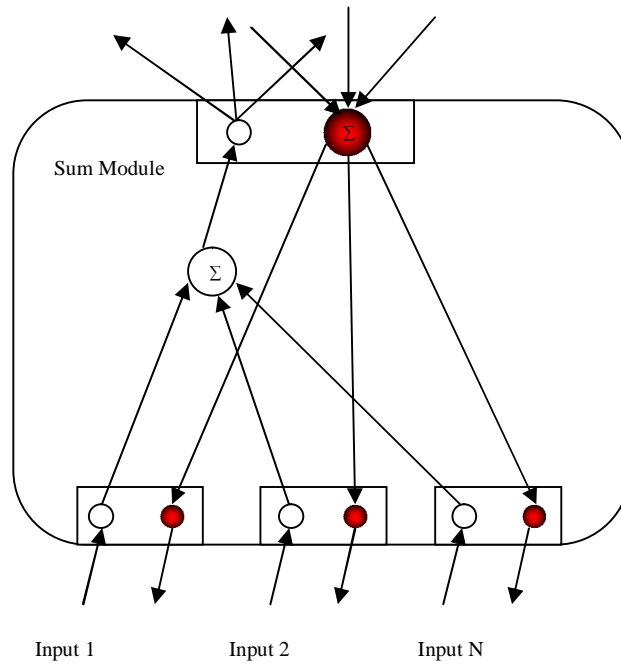
A sin forward-backward node



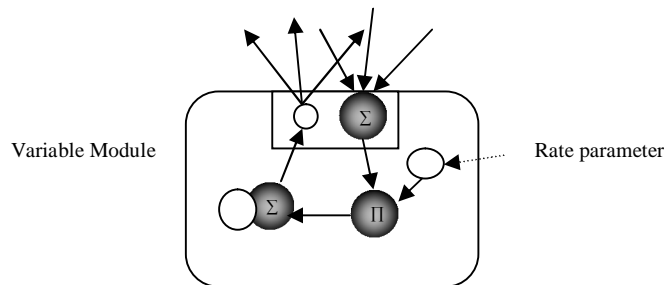
Εικ. 3. Ένα πιο περίπλοκο παράδειγμα: μια παραγωγή μεταβλητών, με πολλαπλές εισόδους. Στην περίπτωση αυτή, πρέπει να περάσει μια διαφορετική μερική παράγωγο πίσω από κάθε είσοδο.

2.3.4 Μεταβλητές καταστάσεις:

Η μεταβλητή κατάσταση αυτή παρέχει ένα τρόπο μοντελοποίησης των σημαντικών παραμέτρων της συγκεκριμένης νευρωνικής αρχιτεκτονικής στο ίδιο πλαίσιο των μεταβλητών ανάδρασης. Η μεταβλητή κατάσταση δεν κάνει τίποτα για την συμπεριφορά προς τα εμπρός. Προς τα πίσω δε, η συμπεριφορά της είναι η ενημέρωση προς την αντίθετη κατεύθυνση της κλίσης του συσσωρευμένου σφάλματος στην προς τα πίσω είσοδο της. Οι μεταβλητές καταστάσεις δεν έχουν είσοδο προς τα εμπρός ή έξοδο προς τα πίσω. Αυτό που έχουν είναι μια παράμετρος ρυθμού η οποία χρησιμοποιείται ως εκπαιδευτικός ρυθμός και χρησιμεύει για να μειώσει ή να αυξήσει τη στάθμη της προσαρμογής στην τιμή που είναι αποθηκευμένη στον κόμβο. Αυτό φαίνεται στο Σχήμα 5.



Εικ. 4. Η άθροιση μεταβλητών : Στην περίπτωση αυτή η κάθε παράγωγος θα τροφοδοτείται στην κάθε έξοδο προς τα πίσω.



Εικ. 5. Οι μεταβλητές καταστάσεις: Η παράμετρος του ρυθμού είναι μερικές αρνητικές σταθερές, που θα καθοριστούν από τις δοκιμές και τα σφάλματα ή από την εξέλιξη.

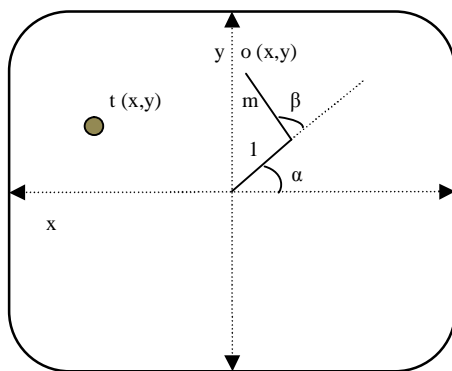
2.4. Παράδειγμα: Μοντελοποίηση ενός απλού ρομποτικού βραχίονα:



Το παράδειγμα αυτό δείχνει τα πιθανά οφέλη της προσέγγισης. Το σχήμα 6 απεικονίζει ένα απλό βραχίονα ρομπότ με δύο κινητά τμήματα. Η άκρη του βραχίονα, κινείται με τη ρύθμιση των γωνιών α και β . Ας υποθέσουμε ότι τα μήκη των l και m της κάθε ενότητας καθορίζονται στην πραγματικότητα, αλλά είναι άγνωστα στο δίκτυο (Σχήμα 7). Στη συνέχεια, το σύστημα μπορεί να εκπαιδευτεί ως εξής για τον υπολογισμό του μήκους. Για ένα σύνολο δειγμάτων εκπαίδευσης, από τις γωνίες α και β που έχουν συσταθεί, και τη θέση του άκρου σε συντεταγμένες x και y οι οποίες είναι ζευγαρωμένες με αυτές τις εξόδους σχηματίζουν στόχο για τις εισόδους. Για κάθε δείγμα εκπαίδευσης, το σφάλμα στην έξοδο λόγω της τρέχουσας εκτίμησης των l και m διαδίδεται πίσω μέσω του δικτύου, με τη διαδικασία της κάθε μεταβλητής που εκτελεί προς τα πίσω δράση. Αυτή είναι η εξέλιξη της διαδικασίας κατάρτισης μέχρι οι τιμές των l και m να συγκλίνουν στις πραγματικές τιμές των τμημάτων του ρομποτικού βραχίονα.

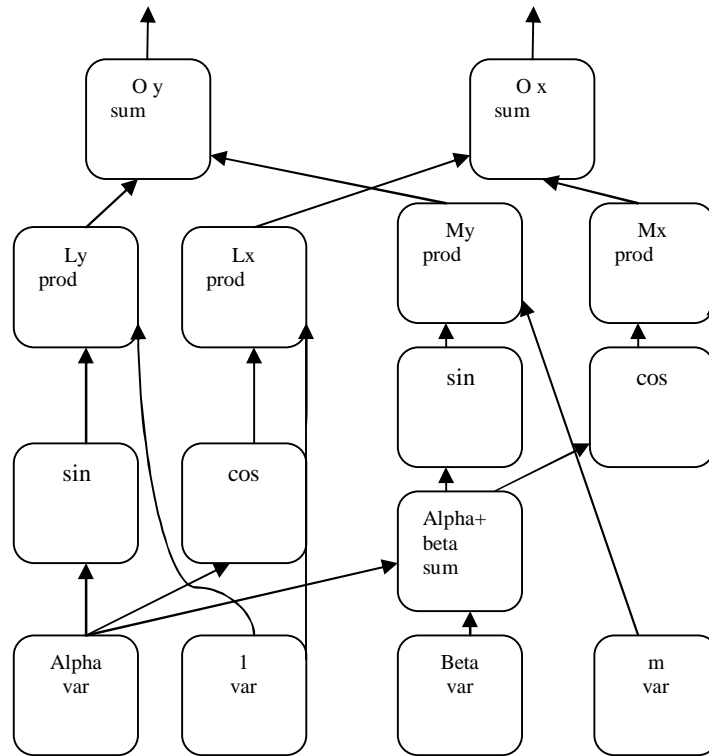
Η χαρτογράφηση από (α, β) συντεταγμένες σε (x, y) συντεταγμένες πρόκειται για λειτουργικό (πολλά-προς-ένα) που μπορεί να εκπροσωπείται από την ανάδραση του δικτύου. Ας υποθέσουμε τώρα, όπως θα ήταν πιο χρήσιμο στην περίπτωση, από την χαρτογράφηση των (x, y) συντεταγμένων να

βρεθούν οι γωνίες α και β που μπορούν να θέσουν το άκρο σε ένα συγκεκριμένο σημείο. Δεδομένου ότι πρόκειται για (ένα-προς-πολλά) χαρτογράφηση, δεν μπορούμε να εκπροσωπηθούμε από την κανονική διαβιβαστική δράση δικτύου (μπροστινή). Αντί αυτού, χρησιμοποιούμε το δίκτυο της πίσω λειτουργίας. Για ένα επιθυμητό σημείο στόχου (T_x, T_y) το σφάλμα μεταξύ του στόχου και του ισχύον σημείου εξόδου (O_x, O_y) τροφοδοτείται στις πίσω εισόδους των κόμβων O_x και O_y , αντίστοιχα. Στη συνέχεια, το δίκτυο λειτουργεί αντίστροφα για να ρυθμιστούν οι γωνίες α και β , έπειτα διαβιβάζει μπροστά για να υπολογιστεί το τρέχον λάθος, κατόπιν κινείται αντίστροφα και πάλι, και αυτή η διαδικασία επαναλαμβάνεται έως ότου το σφάλμα είναι αρκετά μικρό. (Σχήμα.8) Απεικονίζει τη συμπεριφορά του συστήματος. Κάθε ξαφνική αύξηση κατά λάθος αντιστοιχεί σε μια κίνηση του στόχου. Σημειώστε ότι μετά από κάθε κίνηση, το σύστημα ανταποκρίνεται για να μετακινηθεί το άκρο προς το στόχο αρκετά γρήγορα.

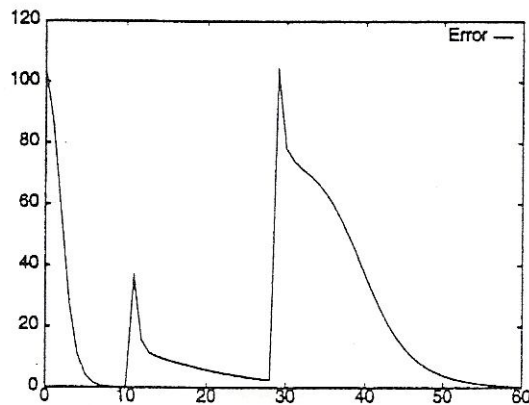


Εικ. 6. Ένας απλός ρομποτικός βραχίονας 2-τμημάτων.

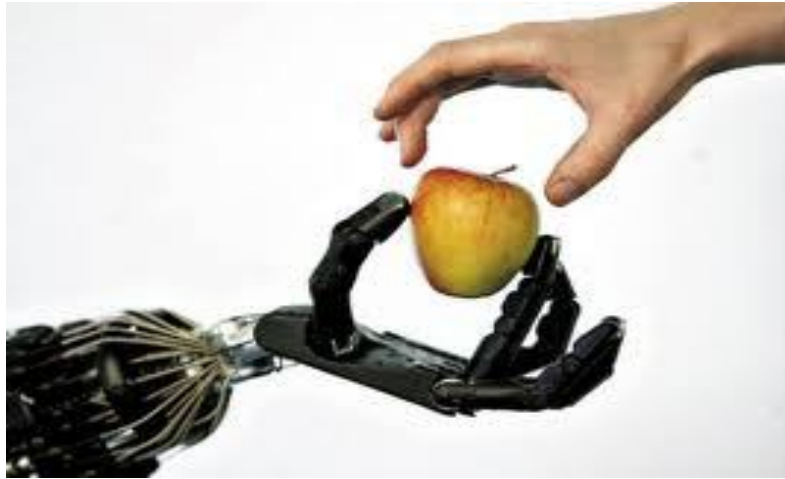
Σχήμα.7 Απεικονίζει τους υπολογισμούς προς τα εμπρός για να υπολογιστεί η θέση ενός ρομποτικού βραχίονα δύο τμημάτων, μήκους βραχίονα l και m και γωνίες βραχίονα α και β . Το β βρίσκεται σε σχέση με το α .



Εικ. 7. Ένα δίκτυο μπροστινών-πίσω μονάδων για τον καθορισμό του ρομποτικού βραχίονα από το σχήμα 6. Σημειώστε ότι για κάθε σύνδεση προς τα εμπρός στο διάγραμμα, υπάρχει επίσης μια προς τα πίσω σύνδεση που δεν φαίνεται να κρατήσει το διάγραμμα απλούστερο.



Εικ. 8. Πλοκή σφάλματος (Ευκλείδεια απόσταση μεταξύ του στόχου και το άκρο του ρομποτικού βραχίονα) ως προς το χρόνο. Κάθε απότομη αύξηση του σφάλματος αντιστοιχεί στο στόχο που μετακινείται.



2.5. Συζήτηση και συμπεράσματα:

Η εργασία αυτή έχει περιγράψει μια νέα και απλή μέθοδο για τον καθορισμό αντιστρέψιμων δικτύων ανάδρασης με «εγγενείς ικανότητες μάθησης». Κάθε κατάσταση στο δίκτυο είναι τώρα μια μεταβλητή που έχει ένα μέρος προς τα εμπρός και ένα προς τα πίσω. Το μπροστινό μέρος υπολογίζει ορισμένες διαφορίσιμες συναρτήσεις της εισόδου του, ενώ στο πίσω μέρος συσσωρεύεται το λάθος από τους διαδοχικούς κόμβους και το πολλαπλασιάζει με την παράγωγο της συνάρτησης σε σχέση με τους κόμβους των μπροστινών εισόδων. Από τη μία πλευρά, αυτό μπορεί να θεωρηθεί ως απλώς μια νέα δήλωση του γενικευμένου κανόνα Δέλτα [22]. Από την άλλη πλευρά όμως, αυτός ο απλός τρόπος προβολής, φαίνεται να είναι ένα σημαντικό βήμα για τους ερευνητές που ενδιαφέρονται στην εξέλιξη νευρωνικών δικτύων. Από την προηγούμενη εργασία έχει περιοριστεί είτε ο τύπος του δικτύου που έχει εξελιχθεί σε κλασική, πολλαπλών στρώσεων, αντίληψη ανάδρασης (συνήθως με ένα μη τυποποιημένο πρότυπο σύνδεσης), είτε άλλοι τύποι δικτύου (όπως ανώτερης τάξης δίκτυα) [14]), αλλά με τον περιορισμό ότι οι ελεύθεροι παράμετροι του δικτύου (π.χ. τα βάρη) θα βρεθούν μέσα από μια διαδικασία τυχαίας αναρρίχησης [14] ή προσομοιωμένης ανόπτησης [16]. Σε γενικές γραμμές, οι κλίσεις των μεθόδων συγκλίνουν πολύ πιο γρήγορα από τις μεθόδους διαταραχών, εξ ου και η προοπτική που προσφέρει αυτό το έγγραφο έχοντας σημαντικές επιπτώσεις.

Υπάρχουν δύο τρόποι όπου αυτό το πλαίσιο μπορεί να εφαρμοστεί σε εξελισσόμενες νευρωνικά αρχιτεκτονικές. Πρώτον κάθε άμεση ή έμμεση κωδικοποίηση που θα διευκρινίζει συνήθως ένα δίκτυο από απλά κελιά όπου αυτά θα μπορούσαν κάλλιστα να χρησιμοποιηθούν για τον καθορισμό ενός δικτύου μεταβλητών ανάδρασης - με άμεσο όφελος ότι το δίκτυο θα έχει τη δική του ενσωματωμένη του στάση απέναντι στη μάθηση και θα είναι δυνητικά αντιστρέψιμο. Η δεύτερη πιθανότητα είναι να εξελιχθούν τα στοιχεία της κάθε μεταβλητής, αντί να προσδιοριστούν από πριν. Κάθε μεταβλητή ανάδρασης που παρουσιάζεται εδώ προσδιορίζονται ως απλά κελιά, αλλά σαφώς, μια εναλλακτική λύση θα ήταν να προσδιοριστούν από την άποψη των σύνθετων μεταβλητών, όπου κάθε μία από αυτές θα μπορούσαν επίσης να καθοριστούν είτε από πλευράς απλών κελιών ή δικών τους συγκροτημένων μεταβλητών. Αυτό το είδος της πολυεπίπεδης ιεραρχικής παρεμπόδισης είναι πάγια πρακτική στον σχεδιασμό του υλικού και λογισμικού και έχει τη δυνατότητα σημαντικής επίπτωσης στην τεχνητή εξέλιξη των νευρωνικών δικτύων. Οι μεταβλητές ανάδρασης που παρουσιάζονται εδώ είναι χρήσιμο βήμα προς αυτήν την κατεύθυνση.

Βιβλιογραφία:

1. D. Chalmers, "The evolutions of learning: an experiment in genetic connectionism," in Proceedings of the 1990 Connectionist Models Summer School (D. Touretzky, J. Elman, T. Sejnowski, and G. Hinton, eds.), San Francisco: Morgan Kaufman, (1990).
2. S. Bengio, Y. Bengio, and J. Cloutier, "Use of genetic programming for the search of a new learning rule for neural networks," in Proceedings of IEEE International Conference on Evolutionary Computation, pp. 324 - 327, Orlando: IEEE, (1994).
3. B. Yamauchi and It. Beer, "Sequential behaviour and learning in evolved dynamical neural networks," Adaptive Behaviour, vol. 2, pp. 219 - 246, (1994).
4. S. Lucas, "Evolving neural network learning behaviours with set-based chromosomes," in Proceedings of European Symposium on Artificial Neural Networks (ESANN '96), pp. 291 - 296, Brussels: D acto, 1996 .
5. D. Fogel, "Using evolutionary programming to create networks that are capable of playing tic-tac-toe," in Proceedings of IEEE International Conference on Neural Networks, pp. 875 - 880, San Francisco: IEEE, (1993).

6. D. Dasgupta and D. McGregor, "Designing application specific neural networks using the structured genetic algorithm," in Proceedings of COGANN-92 - IEEE International Workshop on Combinations of Genetic Algorithms and Neural Networks, pp. 87 - 96, Baltimore: IEEE, (1992).
7. L. Marti, "Genetically generated neural networks ii: searching for an optimal representation," in Proceedings of the International Joint Conference on Neural Networks (Baltimore '92), pp. I, 221 - 226, San Diego, CA: IEEE, (1992).
8. J. McDonell and D. Waagen, "Neural network structure design by evolutionary programming," in Proceedings of the Third Annual Conference on Evolutionary Programming (D. Fogel and W. Atmar, eds.), pp. 79 - 89, Evolutionary Programming Society, (1993).
9. J. McDonell, W. Page, and D. Waagen, "Neural network construction using evolutionary search," in Proceedings of the Third Annual Conference on Evolutionary Programming (A. Sebald and L. Fogel, eds.), pp. 9 - 16, World Scientific, (1994).
10. H. Kitano, "Designing neural networks using genetic algorithm with graph generation system," Complex Systems, vol. 4, pp. 461 - 476, (1990).
11. F. Gruau, "Cellular encoding of genetic neural networks," Laboratoire de l'Informatique du Parallelisme Technical Report 92-21, Ecole Normale Superieure de Lyon, (1992).
- F. Gruau, "Automatic definition of modular neural networks," Adaptive Behaviour, vol. 3, pp. 151 - 183, (1994).
13. E. Boers and H. Kuiper, "Biological metaphors and the design of modular artificial neural networks," Masters thesis, Department of Computer Science and Experimental and Theoretical Psychology, Leiden University, the Netherlands, (1993).
14. H. Muhlenbein and B. Zhang, "Synthesis of sigma-pi neural networks by the breeder genetic programming," in Proceedings of IEEE International Conference on Evolutionary Computation, pp. 318 - 323, Orlando: IEEE, (1994).
15. B. Zhang and H. Muehlenbein, "Balancing accuracy and parsimony in genetic programming," Evolutionary Computation, vol.3, pp. 17 - 381 (1995). -
16. K. Sharman, A. Esparcia-Alcazar, and Y. Li, "Evolving signal processing algorithms by genetic programming," in Proceedings of IEE 1st International Conference on Genetic