



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ  
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΕΥΦΥΩΝ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Εμπόριο στην Αγορά Συναλλάγματος  
με Αναδρομική Ενισχυτική Μάθηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνος Α. Βρακόπουλος

Επιβλέπων: Ανδρέας-Γεώργιος Σταφυλοπάτης  
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος 2013





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ  
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΕΥΦΥΩΝ ΥΠΟΛΟΓΙΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

Εμπόριο στην Αγορά Συναλλάγματος  
με Αναδρομική Ενισχυτική Μάθηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνος Α. Βρακόπουλος

Επιβλέπων: Ανδρέας-Γεώργιος Σταφυλοπάτης  
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 17η Ιουλίου 2013.

---

Ανδρέας-Γεώργιος Σταφυλοπάτης Καθηγητής Ε.Μ.Π.	Στέφανος Κόλλιας Καθηγητής Ε.Μ.Π.	Γεώργιος Στάμου Επικ.Καθηγητής Ε.Μ.Π.
--	--------------------------------------	--

Αθήνα, Ιούλιος 2013

.....  
**Κωνσταντίνος Α.Βρακόπουλος**  
Διπλωματούχος  
Ηλεκτρολόγος Μηχανικός  
και Μηχανικός Υπολογιστών

Copyright © Κωνσταντίνος Α.Βρακόπουλος, 2013

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

## Περίληψη

Μια ενδιαφέρουσα περιοχή της σύγχρονης επιστήμης είναι η προσέγγιση της Διαχείρισης Ιδιοκτησίας μέσω Υπολογιστικής Νοημοσύνης. Τα οικονομικά συστήματα είναι μια καλή επιλογή για την Μηχανική Μάθηση να εφαρμοσθεί μιας που είναι εξαιρετικά πολύπλοκα και η πολυπλοκότητα είναι αυτό που η Μάθηση, θεωρητικά, τίθεται να χειριστεί. Επιπλέον, εκτός από ενδιαφέρουσες τέτοιες εφαρμογές είναι και χρήσιμες αφού το πόσο καλά αποδίδουν μετρίεται με την πιο ρευστή έννοια της απόδοσης · τα χρήματα. Σε αυτήν την εργασία εξετάζεται το εμπόριο στην αγορά συναλλάγματος με χρήση ενός αλγόριθμου Μηχανικής Μάθησης, της Αναδρομικής Ενισχυτικής Μάθησης. Μετά την παρουσίαση του αυτός ο αλγόριθμος χρησιμοποιείται για να διαμορφωθεί μια Υπολογιστική Εμπορική Στρατηγική στην οποία γίνεται ρύθμιση παραμέτρων και αξιολογείται πάνω σε πραγματικά δεδομένα της αγοράς. Παραθέτονται τα συμπεράσματα της αξιολόγησης και ιδέες για σχετική μελλοντική έρευνα. Τέλος, γίνεται αναφορά στο πρωτόκολλο *FIX* το οποίο είναι το κυρίαρχα διαδεδомένο πρότυπο για συναλλαγές οικονομικής πληροφορίας μέσω ηλεκτρονικής επικοινωνίας, και περιγράφεται η χρήση μιας λογισμικής διεπαφής που υλοποιεί αυτό το πρωτόκολλο για να ενωθεί η Υπολογιστική Εμπορική Στρατηγική στην πραγματική αγορά συναλλάγματος έτσι ώστε να πάρει μέρος στις εξελίξεις του κόσμου των σύγχρονων αγορών · στην άνοδο των μηχανών.

**Λέξεις Κλειδιά :** διαχείριση ιδιοκτησίας, υπολογιστική νοημοσύνη, μηχανική μάθηση, αγορά συναλλάγματος, αναδρομική ενισχυτική μάθηση, υπολογιστική εμπορική στρατηγική, αναδρομικά νευρωνικά δίκτυα, πρωτόκολλο *FIX*



## Abstract

The Computational Intelligence approach to Asset Management is an interesting area of modern science. Financial systems are a good option for Machine Learning to be applied since they are uniquely complex and complexity is what Learning is supposed in theory to be dealing with. Moreover, such applications are of high utility, because the measurements of how good they are doing are taken by the most liquid notion of performance · money. This diploma thesis is investigating the foreign exchange market trading via a Machine Learning algorithm, Recurrent Reinforcement Learning. Following its presentation, this algorithm is intergrated into a Computational Trading Strategy which is fine-tuned and evaluated by real market data. Conclusions of the evaluation and ideas for further research are pointed. Finally, a reference is made to protocol FIX, which is the widespread standard for financial information exchange through electronic communication, and the use of a software interface is described, targeting the coupling of the Computational Trading Strategy with the real Forex Market so that it participates in the progression of the modern markets' world · the rise of the machines.

**KeyWords:**asset management, computational intelligence, machine learning, forex market, recurrent reinforcement learning, computational trading strategy, recurrent neural networks, protocol *FIX*





## Ευχαριστίες

Ευχαριστώ τον καθηγητή κ.Σταφυλοπάτη για την ανάθεση και την επίβλεψη αυτής της εργασίας και τον Δρ.Γιώργο Σιόλα από το εργαστήριο Ευφυών Υπολογιστικών Συστημάτων για την βοήθεια σε κρίσιμα διαστήματα. Επιπλέον, πολλές ευχαριστίες πηγαίνουν στην οικογένεια και τους φίλους μου για όλη την υποστήριξη.



# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>15</b>
<b>2</b>	<b>Νευρωνικά Δίκτυα</b>	<b>19</b>
2.1	Στατικά Νευρωνικά Δίκτυα . . . . .	20
2.2	Αναδρομικά Νευρωνικά Δίκτυα . . . . .	24
<b>3</b>	<b>Ενισχυτική Μάθηση</b>	<b>27</b>
3.1	Διατύπωση του προβλήματος . . . . .	27
3.2	Έμμεσος τρόπος επίλυσης . . . . .	29
3.3	Άμεσος τρόπος επίλυσης . . . . .	30
3.4	Σύγκριση έμμεσου και άμεσου τρόπου . . . . .	32
<b>4</b>	<b>Η Αγορά Συναλλάγματος</b>	<b>34</b>
4.1	Είσοδος . . . . .	38
4.2	Σύμβολα και Τιμές . . . . .	38
4.3	Συναλλαγές και Κέρδη . . . . .	39
<b>5</b>	<b>Υπολογιστικές Εμπορικές Στρατηγικές</b>	<b>47</b>
<b>6</b>	<b>Αναδρομική Ενισχυτική Μάθηση</b>	<b>51</b>
6.1	<i>Batch</i> μάθηση . . . . .	54
6.2	<i>Online</i> μάθηση . . . . .	56
6.3	Μάθηση με <i>bid/ask</i> τιμές . . . . .	57
6.4	Τεχνητές τιμές . . . . .	58
<b>7</b>	<b>Εμπόριο στην Αγορά Συναλλαγματος και Πειραματικά Αποτελέσματα</b>	<b>63</b>
7.1	Δεδομένα και Μέθοδος . . . . .	64
7.2	Ρύθμιση Παραμέτρων . . . . .	67
7.3	Αξιολόγηση . . . . .	70
7.4	Συμπεράσματα και Μελλοντική Έρευνα . . . . .	83



# Κατάλογος Σχημάτων

1.1	Διαχείριση Ιδιοκτησίας . . . . .	17
1.2	Κέρδος μέσω διαδοχικών αγοροπωλησιών . . . . .	18
2.1	NN στρωμάτων εμπρόσθιας τροφοδότησης . . . . .	20
2.2	Σύνθετη συνάρτηση . . . . .	21
2.3	Οπισθοδιάδοση . . . . .	22
2.4	Συνάψεις και Νευρώνες . . . . .	23
2.5	NN στρωμάτων αναδρομικής τροφοδότησης . . . . .	24
2.6	Αποδιπλωμένο ANN . . . . .	25
3.1	Σχήμα Ενισχυτικής Μάθησης . . . . .	28
4.1	Παράδειγμα κυματισμού της τιμής του συμβόλου <i>EurCzk</i> , έτος 2012 . . . . .	37
4.2	Η δομή της σύγχρονης Αγοράς Συναλλάγματος . . . . .	38
4.3	<i>bid/ask</i> αγοροπωλησίες . . . . .	39
4.4	Αγορές με υψηλό και χαμηλό <i>Movement/Spread</i> λόγο . . . . .	46
5.1	Εμπορική Στρατηγική με Επιβλεπόμενη Μάθηση . . . . .	48
5.2	Εμπορική Στρατηγική με Πρόβλεψη Τιμών . . . . .	49
5.3	Εμπορική Στρατηγική με Άμεση Ενισχυτική Μάθηση . . . . .	49
6.1	Διάγραμμα <i>Batch</i> μάθησης . . . . .	54
6.2	Δείγμα αξιολόγησης της <i>RRL</i> μεθόδου πάνω σε <i>bid/ask</i> τιμές διακροτήματος . . . . .	59
6.3	Δείγμα αξιολόγησης της <i>RRL</i> μεθόδου πάνω σε <i>bid/ask</i> τιμές διακροτήματος με υπέρθεση θορύβου . . . . .	60
6.4	Δείγμα αξιολόγησης της <i>RRL</i> μεθόδου πάνω σε <i>bid/ask</i> τιμές θορύβου . . . . .	61
6.5	Δείγμα αξιολόγησης της <i>RRL</i> μεθόδου πάνω σε <i>bid/ask</i> τιμές από έξοδο στοχαστικού συστήματος . . . . .	62

7.1	Τιμές συναλλάγματος για τις αγορές <i>UsdMyr,UsdSek,GbpDem</i> και <i>DemFrf</i> . . . . .	65
7.2	Δείγμα αξιολόγησης της <i>RRL</i> μεθόδου πάνω σε <i>bid/ask</i> τιμές της αγοράς <i>DemFim</i> . . . . .	66
7.3	Ρύθμιση παραμέτρων <i>CadUsd</i> , ANN με 1 στρώμα . . . . .	72
7.4	Ρύθμιση παραμέτρων <i>UsdDem</i> , ANN με 1 στρώμα . . . . .	73
7.5	Ρύθμιση παραμέτρων <i>UsdBef</i> , ANN με 1 στρώμα . . . . .	74
7.6	Ρύθμιση παραμέτρων <i>UsdZar</i> , ANN με 1 στρώμα . . . . .	75
7.7	Ρύθμιση παραμέτρων <i>GbpDem</i> , ANN με 2 στρώματα . . . . .	76
7.8	Ρύθμιση παραμέτρων <i>UsdChf</i> , ANN με 2 στρώματα . . . . .	77
7.9	Ρύθμιση παραμέτρων <i>DemJpy</i> , ANN με 2 στρώματα . . . . .	78
7.10	Εφαρμογή και σύγκριση της ρυθμισμένης ΕΣ με ANN 1 στρώματος στην αγορά <i>GbpUsd</i> . . . . .	80
7.11	Εφαρμογή και σύγκριση της ρυθμισμένης ΕΣ με ANN 1 στρώματος στην αγορά <i>UsdFrf</i> . . . . .	81
7.12	Εφαρμογή και σύγκριση της ρυθμισμένης ΕΣ με ANN 2 στρώματων στην αγορά <i>UsdSgd</i> . . . . .	82
7.13	Εμπορικός Εγκέφαλος . . . . .	85
8.1	Ακολουθιακό Διάγραμμα <i>UML</i> για την συνολική δομή της εφαρμογής . . . . .	87
8.2	Ακολουθιακό Διάγραμμα <i>UML</i> για αγορά και πώληση . . . . .	88
8.3	Ακολουθιακό Διάγραμμα <i>UML</i> για την ανανέωση των <i>bid/ask</i> δεδομένων . . . . .	89

## Κατάλογος Πινάκων

4.1	Λίστα με τα πιο ανταλασσόμενα νομίσματα . . . . .	36
4.2	Λίστα των οργανισμών με την μεγαλύτερη συμμετοχή . . . . .	36
4.3	Παράδειγμα ακολουθίας αποφάσεων και ανταμοιβών . . . . .	45
7.1	Ρυθμισμένες παράμετροι για Εμπορική Στρατηγική με 1 στρώμα	70
7.2	Τελικό στάδιο ρύθμισης παραμέτρων για εμπορική στρατηγική με 1 στρώμα . . . . .	71
7.3	Ρυθμισμένες παράμετροι για Εμπορική Στρατηγική με 2 στρώματα	71
7.4	Τελικό στάδιο ρύθμισης παραμέτρων για εμπορική στρατηγική με 2 στρώματα . . . . .	79
7.5	Συνολική αξιολόγηση και σύγκριση της <i>RRL</i> στρατηγικής . . .	83
7.6	Μετρήσεις χαρακτηριστικών του εμπορίου στις αγορές ελέγχου	84

# Κεφάλαιο 1

## Εισαγωγή

Στο σχήμα 1.1 φαίνεται μια αναλογία ανάμεσα στα πεδία της Φυσικής, της Επιστήμης Υπολογιστών και της Οικονομίας έτσι ώστε όταν είμαστε στον κόσμο των οικονομικών αγορών να θεωρούμε τα αντικείμενα ιδιοκτησίας (*asset*) ως τα βασικά αντικείμενα και το εμπόριο (*trading*) ως το βασικό σύνολο κινήσεων που μπορεί να κάνει κάποιος για να διαχειριστεί τα αντικείμενα του. Έτσι όπως στην Φυσική και στην Επιστήμη Υπολογιστών πολλές ερωτήσεις ανάγονται στο πως να εφαρμοστεί ενέργεια ή μέθοδοι ώστε να έρχονται η ύλη ή οι δομές δεδομένων σε ευχάριστες καταστάσεις, έτσι και το βασικό ερώτημα στο κόσμο των οικονομικών αγορών είναι πως να εφαρμοστεί εμπόριο ώστε να φέρνουμε την ιδιοκτησία μας σε ευχάριστες καταστάσεις. Η επιστήμη που αντιμετωπίζει αυτό το ερώτημα αναφέρεται ως Διαχείριση Ιδιοκτησίας (*asset management*).

Στο δεύτερο μισό του 20ού αιώνα παράλληλα με την ανάπτυξη των υπολογιστών η μοντελοποίηση και η προσομοίωση έγινε βασικός τρόπος για να συμβεί ο έλεγχος φυσικών συστημάτων, ανακαλώντας όμως και το πρώτο μισό του ρητού η *Οικονομία είναι λιγότερο ακριβής απ'τις Φυσικές επιστημές αλλά περισσότερο ακριβής από κάθε άλλη κοινωνική επιστήμη* δεν συνέβη το ίδιο και για οικονομικά συστήματα. Τις δύο τελευταίες δεκαετίες ωστόσο, η θεωρία Μηχανικής Μάθησης αναπτύχθηκε και η μάθηση με Οπισθοδιάδοση ανανέωσε την θεωρία των Νευρωνικών Δικτύων κάνοντας έτσι δυνατό τον σχεδιασμό ευφυών υπολογιστικών συστημάτων που εκτελούν πολύπλοκες εργασίες αφού πρώτα έχουν μάθει πως να το κάνουν. Η αναμονή ότι η Διαχείριση Ιδιοκτησίας μπορεί να γίνει αυτόματα και αποδοτικά μέσω υπολογιστών έδωσε πλέον την θέση της στην έρευνα έρχοντας να καλύψει την ζήτηση των επενδυτικών εταιρειών, *hedge funds* κτλπ. Δύο απ'τις πρώτες δημοσιεύσεις Μηχανικής Μάθησης για οικονομικά συστήματα είναι η [1] και η [8].

Κατά την εκκίνηση μιας εργασίας υπολογιστικών τεχνικών πάνω στην Διαχείριση Ιδιοκτησίας είναι σκόπιμο να θεωρηθεί μόνο ένα υποσύνολο του συνολικού προβλήματος και αυτό επειδή είναι δύσκολο να μπει όλο το πρόβλημα



σε ένα τυπικό πλαίσιο τέτοιο ώστε ταυτόχρονα να είναι και υπολογιστικά διαχειρίσιμο. Η παρέλευση της παραπάνω δυσκολίας είναι εν πολλοίς ένας άπ'τους στόχους της Οικονομικής Επιστήμης. Για παράδειγμα μπορούμε να θεωρήσουμε το εμπόριο μόνο σε μια συγκεκριμένη αγορά όπως η αγορά ομολόγων, μετοχών, συναλλάγματος ή άλλες με μια υπολογιστική στρατηγική που βασίζεται σε κάποια απο τις διάφορες περιπτώσεις της Μηχανικής Μάθησης, όπως η Επιβλεπόμενη Μάθηση ή η Ενισχυτική Μάθηση.

Η Ενισχυτική Μάθηση είναι ένα γενικό θεωρητικό πλαίσιο όπου το σύστημα που μαθαίνει βρίσκεται μέσα σε ένα άγνωστο περιβάλλον και προσπαθεί να μάθει να συμπεριφέρεται έτσι ώστε να μεγιστοποιήσει ένα δεδομένο μέτρο απόδοσης. Σε αυτό το τυπικό πλαίσιο θέτονται διάφορες συναρτήσεις οι οποίες μπορούν να αναπαρασταθούν με Νευρωνικά Δίκτυα καθώς αυτά είναι καθολικοί προσεγγιστές συναρτήσεων. Η λύση προβλημάτων Ενισχυτικής Μάθησης χρησιμοποιώντας Νευρωνικά Δίκτυα είναι μια ιδέα Μηχανικής Μάθησης με έντονη ερευνητική δράση[25][26][30][31][32].

Σε αυτήν την εργασία εξετάζεται η δυνατότητα για κερδοφόρο εμπόριο στην Αγορά Συναλλάγματος μέσω Ενισχυτικής Μάθησης με Αναδρομικά Νευρωνικά Δίκτυα. Περιληπτικά, οι βασικοί τροποι με τους οποίους μπορεί να συμβεί κέρδος φαίνονται στο σχήμα 1.2 όπου γίνονται διαδοχικά πώληση και αγορά όταν η αξία μειώνεται ή αγορά και πώληση όταν η αξία αυξάνεται. Η Αγορά Συναλλάγματος θεωρείται η πλέον προσοδοφόρα επιλογή για υπολογιστικές και τεχνικές μελέτες επειδή είναι η πιο ρευστή αγορά του κόσμου, ο όγκος της είναι μεγάλος και οι παράγοντες που την επηρεάζουν είναι γεωγραφικά και πολιτικά αποκεντρωμένοι. Όλα αυτά συμβάλλουν ώστε η κατάσταση της να μην εξαρτάται έντονα από κάποιους συγκεκριμένους και λίγους παράγοντες αλλά περισσότερο απ'την λεγόμενη ψυχολογία της αγοράς (*market psychology*) η οποία μπορεί να είναι προβλέψιμη. Στην ορολογία των αγορών *μαύρος κύκνος* σημαίνει ένα μεγάλης κλίμακας γεγονός που αλλάζει την κατάσταση της αγοράς έτσι ώστε να καταρρίπτει τις προβλέψεις, οπότε θα μπορούσε να πει κανείς ότι η Αγορά Συναλλάγματος συγκρινόμενη με άλλες αγορές, έχει λιγότερους μεγάλους και περισσότερους μικρούς μαύρους κυκνους, έτσι ώστε αν κάποιος κοιτάζει τις κινήσεις των ψαριών προσπαθώντας να ψαρέψει, να υπάρχουν πολύ λιγότερες πιθανότητες να περιπλεύσει κάποιος μεγάλος κύκνος και να χαλάσει τις προβλέψεις του.

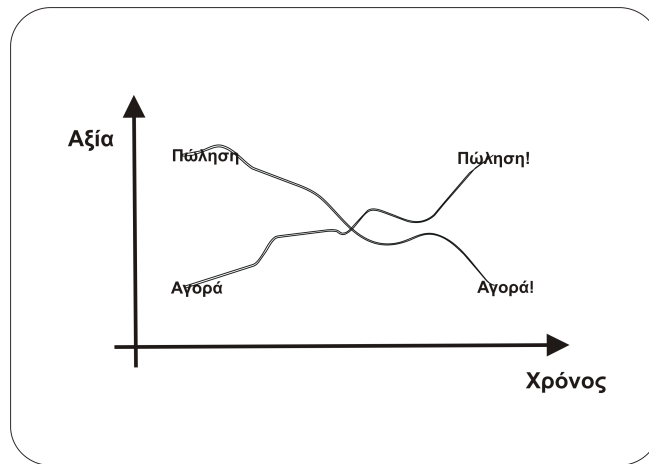
Αναφορικά με τις σχετικές εργασίες, δυο δημοσιεύσεις που συγκρίνουν τρόπους Ενισχυτικής Μάθησης για εμπόριο είναι οι [7] και [2]. Η μέθοδος Αναδρομική Ενισχυτική Μάθηση που θα χρησιμοποιηθεί και σε αυτήν την εργασία βασίζεται στην εργασία των *Moody et al* [1][2][3]. Η ίδια μέθοδος εξετάζεται πάνω σε διάφορες αγορές στα [9][5] και ειδικά στην αγορά συναλλάγματος στα [4][10][6].



Σχήμα 1.1: Διαχείριση Ιδιοκτησίας

- Στο κεφάλαιο 2 παρουσιάζονται τα Νευρωνικά Δίκτυα ως συστήματα μάθησης και η μέθοδος Οπισθοδιάδοση (*backpropagation*) που υπολογίζει παραγώγους σύνθετων συναρτήσεων, και στην οποία βασίζονται πολλοί αλγόριθμοι μάθησης
- στο κεφάλαιο 3 εισάγεται το γενικό πλαίσιο προβλημάτων που θέτει η Ενισχυτική Μάθηση και συγκρίνονται δύο διαφορετικοί τρόποι λύσης τέτοιων προβλημάτων
- στο κεφάλαιο 4 δίνονται κάποιες βασικές ιδέες των οικονομικών αγορών και ειδικά της Αγοράς Συναλλάγματος, όπως οι βασικοί τρόποι συναλλαγών και ο τρόπος υπολογισμού του κέρδους από αυτές
- στο κεφάλαιο 5 γίνεται αναφορά σε διάφορες εμπορικές στρατηγικές που βασίζονται στην Μηχανική Μάθηση
- στο κεφάλαιο 6 αναπτύσσεται αναλυτικά, η Αναδρομική Ενισχυτική Μάθηση, η οποία είναι η βασική μέθοδος της εργασίας
- στο κεφάλαιο 7 παρουσιάζεται η εμπορική στρατηγική που θα χρησιμοποιηθεί, η οποία περιλαμβάνει ως διαδικασία μάθησης την Αναδρομική Ενισχυτική Μάθηση, και τα αποτελέσματα που επιτυγχάνονται πάνω στο σύνολο δεδομένων *HFDF* – 96 της εταιρείας *Olsen*.

Εκτός από την δυνατότητα δημιουργίας μιας Υπολογιστικής Εμπορικής Στρατηγικής, η εξέλιξη στις επικοινωνίες, στην τεχνολογία και στις διαδικασίες πίστωσης καθιστούν την Αγορά Συναλλάγματος διαθέσιμη σε οποιονδήποτε έχει μια σύνδεση στο διαδίκτυο. Έχοντας αναπτύξει κανείς μια τέτοια στρατηγική το επόμενο βήμα είναι να μπορέσει να την εφαρμόσει μέσω κάποιου διαδικτυακού σταθμού.



Σχήμα 1.2: Κέρδος μέσω διαδοχικών αγοροπωλησιών

-στο κεφάλαιο 8 γίνεται αναφορά στο πρωτόκολλο *FIX* (*Financial Information eXchange*) το οποίο διευθετεί τέτοιες διαδικασίες και είναι ιδιαίτερα χρήσιμο για αλγοριθμικό Εμπόριο Υψηλής Συχνότητας. Επιπλέον γίνεται αναφορά σε διεπαφές λογισμικού (*APIs*) που το υλοποιούν και παρουσιάζεται με λεπτομέρειες η χρήση μιας απ' αυτών.

## Κεφάλαιο 2

### Νευρωνικά Δίκτυα

Τα συστήματα Μηχανικής Μάθησης περιέχουν ένα σύνολο από ρυθμίσιμες παραμέτρους και ως μάθηση εννοούμε την εξέλιξη αυτών των παραμέτρων ώστε ένα μέτρο, η *συνάρτηση μάθησης*, να βελτιστοποιείται. Μετά την μάθηση, το σύστημα μπαίνει σε λειτουργία ελέγχου και αξιολογείται με βάση ένα εν γένει άλλο μέτρο απόδοσης, την *συνάρτηση αποδοσης*. Ο σχεδιασμός ενός συστήματος που μαθαίνει παίρνει ως δεδομένο την *συνάρτηση απόδοσης* και περιλαμβάνει τον ορισμό της *συνάρτησης μάθησης*, της *δομής* του συστήματος και του αλγόριθμου που ρυθμίζει τις παραμέτρους [11][12][13][14].

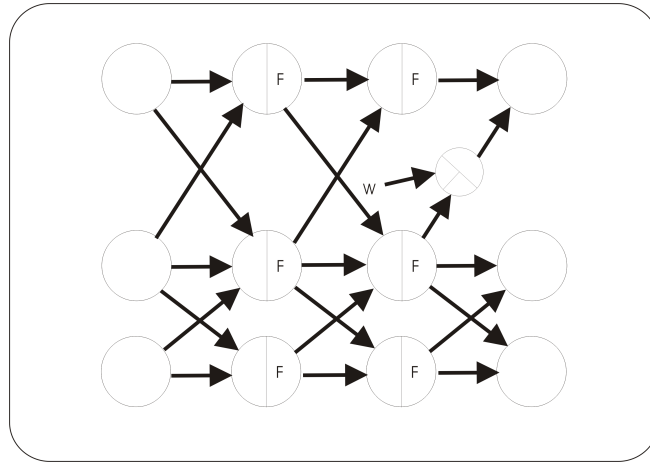
Πολλά προβλήματα μάθησης προκύπτουν από τρία παρακάτω πλαίσια:

*Επιβλεπόμενη Μάθηση*: το σύστημα παίρνει ένα σύνολο από παραδείγματα εισόδου-εξόδου μιας άγνωστης *συνάρτησης* και καλείται να ρυθμίσει τις παραμέτρους του έτσι ώστε να προβλέπει σωστά την έξοδο για εντός και εκτός παραδειγμάτων εισόδους

*Μη-Επιβλεπόμενη Μάθηση*: είναι διαθέσιμο ένα μεγάλο πλήθος από εισόδους και ο στόχος είναι να βρεθεί κάποια κρυφή δομή αυτών των δεδομένων

*Ενισχυτική Μάθηση*: το σύστημα βρίσκεται σε ένα άγνωστο εξελισσόμενο περιβάλλον, και καλείται να συμπεριφέρεται με τέτοιο τρόπο ώστε να δέχεται πολλές ανταμοιβές.

Αναφορικά με τον σχεδιασμό ενός συστήματος που μαθαίνει, υπάρχουν δύο ανταγωνιστικές τάσεις σε αυτά που χρειάζονται. Αν και είναι καλό μέσω της μάθησης να συλλαμβάνει πολύπλοκες συμπεριφορές, πρέπει ταυτόχρονα η δομή του να εξαρτάται με απλό τρόπο από τις ρυθμιζόμενες παραμέτρους έτσι ώστε να μπορούν να αναπτυχθούν αποδοτικοί αλγόριθμοι μάθησης, θεωρήματα σύγκλισης κτλπ. Δηλαδή, κατά έναν τρόπο, το σύστημα πρέπει να είναι σύνθετο αλλά να είναι συντιθεμένο με απλούς κανόνες. Ένας τρόπος να γίνει αυτό είναι να οριστεί η δομή του μέσω ενός Γραφήματος (*graph*), και έτσι ανακύπτει μια μεγάλη κλάση συστημάτων μάθησης, τα Γραφικά Μοντέλα (*Graphical Models*).



Σχήμα 2.1: NN στρωμάτων εμπρόσθιας τροφοδότησης

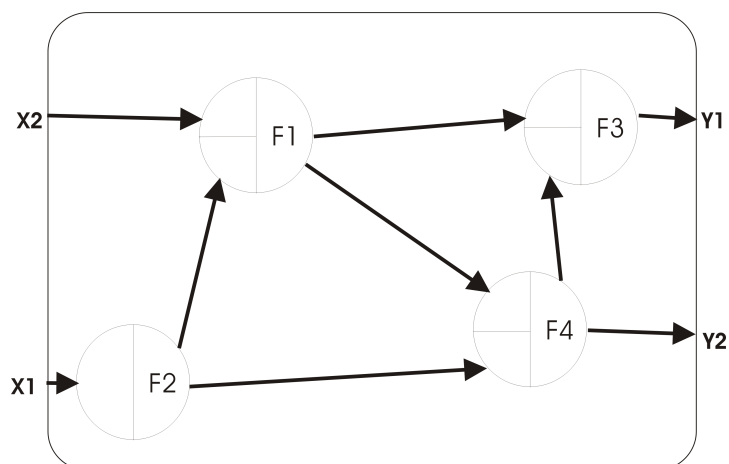
Μια περίπτωση τέτοιων μοντέλων είναι τα Νευρωνικά Δίκτυα στα οποία εστιάζουμε σε αυτό το κεφάλαιο [15][16].

## 2.1 Στατικά Νευρωνικά Δίκτυα

### Δομή

Η δομή ενός στατικού νευρωνικού δικτύου (ΣΝΝ) περιγράφεται από μια πλειάδα  $(N, E, w, f, I, O)$ , όπου  $N$  και  $E$  είναι οι κόμβοι και οι ακμές ενός κατευθυνόμενου γραφήματος χωρίς κύκλους,  $w$  είναι ένα σύνολο από παραμέτρους η κάθε μια εκ των οποίων αντιστοιχεί σε μια ακμή από σύνολο  $E$ ,  $f$  είναι ένα σύνολο από σχετικά απλές συναρτήσεις η κάθε μια εκ των οποίων αντιστοιχεί σε έναν κόμβο από σύνολο  $N$ , και τα  $I, O$  είναι υποσύνολα του  $N$  τα οποία λειτουργούν ως κόμβοι εισόδου και κόμβοι εξόδου αντίστοιχα. Στο σχήμα 2.1 φαίνεται ένα παράδειγμα ΣΝΝ πολλών στρωμάτων. Το πρώτο στρώμα νευρώνων είναι στρώμα εισόδου και απλώς προκαλεί την προώθηση των εισόδων. Τα υπόλοιπα στρώματα είναι στρώματα υπολογισμού και σε κάθε νευρώνα τους υπάρχει μια επιπλέον ακμή καταφλίου ή οποία παίρνει πάντοτε είσοδο 1. Το πέρασμα μιας τιμής από μια ακμή σημαίνει πολλαπλασιασμό αυτής με την αντίστοιχη παράμετρο  $w$  και κάθε νευρώνας προσθέτει τις εισόδους του και περνάει το άθροισμα μέσα από μια συνάρτηση ενεργοποίησης  $f$  την οποία θεωρούμε κοινή για όλους τους νευρώνες του δικτύου.

Συμβολίζουμε την έξοδο κάθε στρώματος  $i$  με το διάνυσμα  $\underline{x}_{(i)}$  και την επέκταση αυτού, έτσι ώστε να περιλαμβάνει και την είσοδο 1 των ακμών καταφλίου, με  $\hat{\underline{x}}_{(i)}$ . Επιπλέον, ο πίνακας ακμών του στρώματος  $i$  συμβολίζεται με  $W_{(i)}$  και ο



Σχήμα 2.2: Σύνθετη συνάρτηση

αντίστοιχος επεκταμένος πίνακας, με μια επιπλέον γραμμή ώστε να περιλαμβάνει και τις ακμές κατωφλίου, με  $\hat{W}_{(i)}$ .

Με αυτούς τους συμβολισμούς, ο υπολογισμός που κάνει το δίκτυο με είσοδο  $\underline{x}_{(0)}$  είναι

$$\underline{x}_{(1)} = f(\hat{W}_{(0)}^T \cdot \hat{\underline{x}}_{(0)}) \quad (2.1)$$

και γενικά με αναδρομή ως το τελευταίο στρώμα του οποίου η έξοδος  $\underline{y}$  είναι η συνολική έξοδος του δικτύου

$$\underline{x}_{(k)} = f(\hat{W}_{(k-1)}^T \cdot \hat{\underline{x}}_{(k-1)}) \quad (2.2)$$

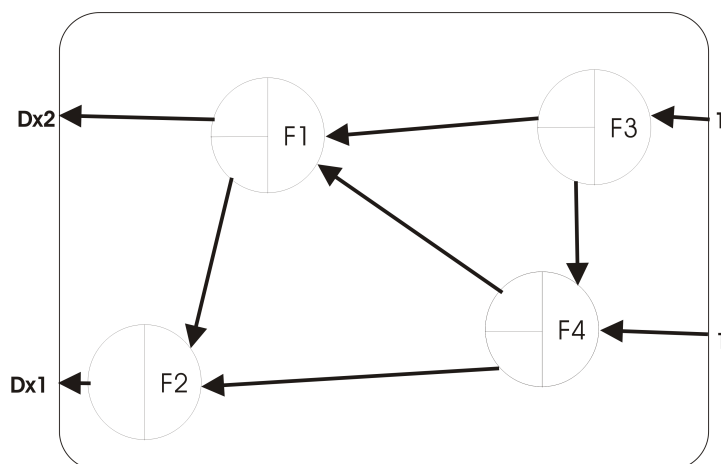
$$k = 1..M, \quad \underline{y} = \underline{x}_{(M)}$$

## Μάθηση

Ο στόχος της μάθησης είναι η ρύθμιση των παραμέτρων του συστήματος έτσι ώστε αυτό να λειτουργεί αποδοτικά και αυτό επιτυγχάνεται με βελτιστοποίηση της συνάρτησης μάθησης. Όταν χρησιμοποιείται μέθοδος κλίσης για αυτήν την βελτιστοποίηση είναι χρήσιμο να μπορούμε να υπολογίσουμε τις παραγώγους των εξόδων του δικτύου ως προς αυτές τις παραμέτρους. Στην συνέχεια θα δούμε έναν συστηματικό τρόπο υπολογισμού αυτών των παραγώγων, την Οπισθοδιάδοση (*backpropagation*) [19][20].

## Οπισθοδιάδοση

Στο σχήμα 2.2 φαίνεται ο τρόπος υπολογισμού μια σύνθετης συνάρτησης που είναι ορισμένη μέσω ενός γραφήματος. Αρχικά γίνεται ο εμπρόσθιος υπολογισμός των εξόδων  $y_1$   $y_2$ . Οι ακμές, αρχίζοντας από τις βασικές εισόδους  $x_1$   $x_2$



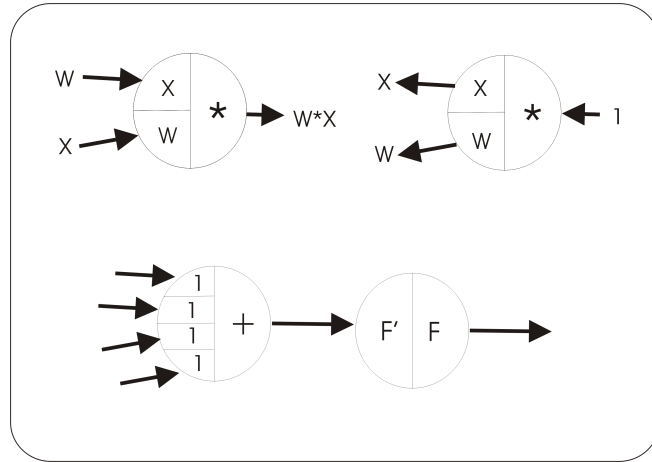
Σχήμα 2.3: Οπισθοδιάδοση

προωθούν τις τιμές δίνοντας ορίσματα στις συναρτήσεις των κόμβων. Σε κάθε κόμβο υπολογίζονται τα αποτελέσματα της συνάρτησης και της παραγώγου της για τα εισερχόμενα ορίσματα, οι οποίες αποθηκεύονται στο δεξί και αριστερό κομμάτι του κόμβου αντίστοιχα. Οι τιμές του δεξιού κομματιού απλώς προωθούνται μπροστά μέχρι να υπολογιστούν οι έξοδοι. Στην συνέχεια, για να υπολογίσουμε την παράγωγο κάποιας εξόδου ως προς κάποια είσοδο, βάζουμε την τιμή 1 στην θέση της εξόδου και την οπισθοδιαδίδουμε μέχρι την είσοδο, δηλαδή περνάμε από ένα μονοπάτι που ενώνει την είσοδο και την έξοδο και κάθε κόμβος του μονοπατιού προκαλεί πολλαπλασιασμό της εισερχόμενης τιμής με την αποθηκευμένη στο αριστερό μέρος του τιμής. Η παράγωγος που ψάχνουμε είναι το άθροισμα των τιμών από όλα τα μονοπάτια που προκύπτουν. Για παράδειγμα στο σχήμα 2.3 η παράγωγος της  $y_1$  ως προς την  $x_2$  είναι

$$\frac{\partial y_1}{\partial x_2} = \frac{\partial F_3}{\partial F_1} \cdot \frac{\partial F_1}{\partial x_2} + \frac{\partial F_3}{\partial F_4} \cdot \frac{\partial F_4}{\partial F_1} \cdot \frac{\partial F_1}{\partial x_2} \quad (2.3)$$

Αυτή η τεχνική απλώς εφαρμόζει τον κανόνα αλυσίδας της παραγωγίσης συναρτήσεων, αλλά αποδεικνύεται ιδιαίτερα χρήσιμη όταν η σύνθετη συνάρτηση είναι αρκετά πολύπλοκη. Στο σχήμα 2.4 φαίνεται η επίδραση μια συνάρτησης πολλαπλασιασμού σε μονοπάτια εμπρόσθιου υπολογισμού και οπισθοδιάδοσης. Στην συνέχεια του κεφαλαίου κάθε ακμή νευρωνικού δικτύου θα θεωρείται ως ένας τέτοιος κόμβος πολλαπλασιασμού. Στο ίδιο σχήμα φαίνεται και ότι οι νευρώνες προσθέτουν τις εισόδους τους και εφαρμόζουν μια συνάρτηση ενεργοποίησης στο άθροισμα.

Ως ένα παράδειγμα χρήσης της Οπισθοδιάδοσης θεωρούμε ένα ΣΝΝ Μ στρωμάτων όπως αυτό για το οποίο είδαμε τους τύπους υπολογισμού στην προηγούμενη παράγραφο. Γενικά ο συμβολισμός που θα χρησιμοποιηθεί για



Σχήμα 2.4: Συνάψεις και Νευρώνες

τις δομές γραμμικής άλγεβρας είναι ότι τα σύμβολα με μια γραμμή από κάτω όπως το  $\underline{a}$  σημαίνουν διάνυσμα στήλη, σύμβολα όπως το  $a^T$  σημαίνουν διάνυσμα γραμμή, και τα κεφαλαία γράμματα  $A$  σημαίνουν πίνακα του οποίου η  $n$ -οστή γραμμή είναι η  $A(n, :)$ . Ο στόχος μας είναι να βρούμε την παράγωγο κάθε εξόδου του τελευταίου στρώματος ως προς κάθε παράμετρο.

Έστω ότι έχουμε επιλέξει μια έξοδο, έχουμε βάλει την τιμη 1 στην θέση της και οπισθοδιαδίδουμε. Συμβολίζουμε με  $\underline{\delta}_{(i)}$  τις τιμές που υπάρχουν μετά το στρώμα  $(i + 1)$ . Επομένως, με βάση το σχήμα 2.4, ο πίνακας των ζητούμενων παραγώγων είναι

$$\Delta \hat{W}_{(i)} = \underline{\hat{x}}_{(i)} \cdot \delta_{(i)}^T \quad (2.4)$$

Θεωρώντας γνωστό το  $\underline{\delta}_{(i+1)}$  το οποίο είναι οι τιμές που βρίσκονται ακριβώς μετά από το στρώμα  $(i + 2)$  και βασιζόμενοι και πάλι στο σχήμα 2.4, οι τιμές που θα φτάσουν μπροστά από το στρώμα  $(i + 1)$  είναι το διάνυσμα

$$W_{(i+1)} \cdot \underline{\delta}_{(i+1)}$$

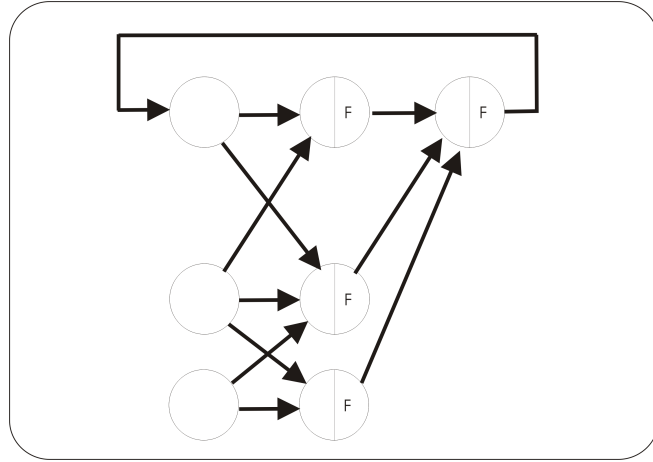
Τελικά περνώντας και το στρώμα  $(i + 1)$  αυτό το διάνυσμα πολλαπλασιάζεται με τον πίνακα

$$f'(diag(\hat{W}_{(i)}^T \cdot \underline{\hat{x}}_{(i)}))$$

αφού το διάνυσμα που διαγωνιοποιείται είναι οι εισοδοι αυτού του στρώματος κατά τον εμπρόσθιο υπολογισμό, και έτσι ο διαγωνιοποιημένος πίνακας περιέχει όλες τις τιμές που είναι αποθηκευμένες στα αριστερά μέρη των κόμβων. Επομένως ισχύει η αναδρομή

$$\underline{\delta}_{(i)} = f'(diag(\hat{W}_{(i)}^T \cdot \underline{\hat{x}}_{(i)})) \cdot W_{(i+1)} \cdot \underline{\delta}_{(i+1)} \quad (2.5)$$





Σχήμα 2.5: NN στρωμάτων αναδρομικής τροφοδότησης

$$i = 0 \dots M - 2$$

Το  $\underline{d}_{(M-1)}$ , δηλαδή οι τιμές που βρίσκονται ακριβώς μετά το τελευταίο στρώμα εξαρτάται απ'την έξοδο για την οποία ψάχνουμε τις παραγωγούς. Αν είναι η  $n$ -οστη έξοδος, και συμβολίζοντας με  $\underline{n}$  το μοναδιαίο διάνυσμα με όλες τις θέσεις μηδενικά εκτός απ'την  $n$ -οστή, τότε ισχύει

$$\underline{d}_{(M-1)} = \underline{n} \cdot f'(\hat{W}_{(M-1)}^T(n, :) \cdot \hat{\underline{x}}_{(M-1)}) \quad (2.6)$$

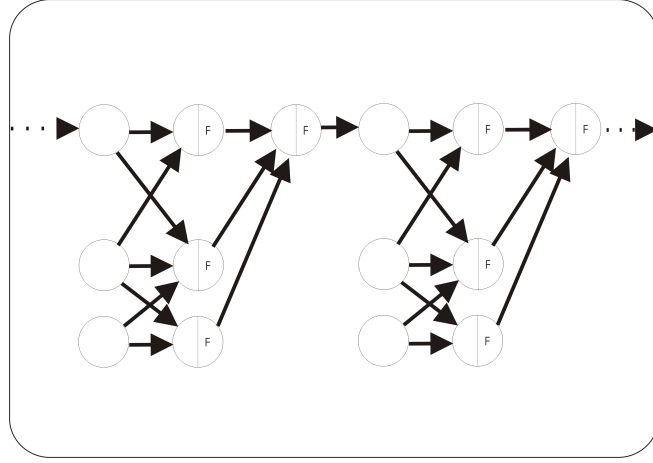
## 2.2 Αναδρομικά Νευρωνικά Δίκτυα

### Δομή

Τα Αναδρομικά Νευρωνικά Δίκτυα (ANN) έχουν την ίδια δομή  $(N, E, w, f, I, O)$  με τα στατικά με την διαφορά ότι υπάρχουν και κύκλοι στο κατευθυνόμενο γράφημα  $(N, E)$ . Επομένως, τα ANN έχουν ακολουθιακή κατάσταση και μπορούν να μοντελοποιούν συμπεριφορά εξελισσόμενων συστημάτων [17][18]. Ένα παράδειγμα ANN με δυο στρώματα φαίνεται στο σχήμα 2.5.

### Μάθηση

Η Οπισθοδιάδοση μπορεί να χρησιμοποιηθεί και για την μάθηση ANN. Το πρώτο βήμα είναι να θεωρήσουμε μια αποδιπλωμένη (*unfolded*) μορφή του δικτύου όπως στο σχήμα 2.6. Αυτή η μορφή μοιάζει με ΣNN πολλών στρωμάτων με την διαφορά ότι οι παράμετροι άνα  $K$  στρώματα, όπου  $K$  είναι ο αριθμός



Σχήμα 2.6: Αποδιπλωμένο ANN

στρωμάτων του ANN, ταυτίζονται και επιπλέον μας ενδιαφέρουν οι παράγωγοι των νευρώνων ανά κάθε K στρώματα.

Αναφορικά με το σχήμα 2.6 οι υπολογισμοί των παραγώγων για μια οποιαδήποτε παράμετρο  $w$  μπορούν να γίνουν με την αναδρομή

$$\frac{dy_t}{dw} = \frac{\partial y_t}{\partial w} + \frac{\partial y_t}{\partial y_{t-1}} \cdot \frac{dy_{t-1}}{dw} \quad (2.7)$$

οπού με  $y$  συμβολίζεται η έξοδος του ANN.

Συμβολίζοντας με  $\hat{\theta}$  και  $\hat{W}$  τις παραμέτρους των δύο στρωμάτων του ANN, με  $\hat{x}_t$  την είσοδο και με  $\hat{z}_t$  την έξοδο του πρώτου στρώματος, βρίσκουμε με οπισθοδιάδοση

$$\frac{\partial y_t}{\partial \hat{\theta}} = \hat{z}_t \cdot f'(\hat{z}_t^T \cdot \hat{\theta}) \quad (2.8)$$

$$\Delta \hat{W} = \hat{x}_{(t-1)} \cdot (f'(\text{diag}(\hat{W}^T \cdot \hat{x}_{(t-1)})) \cdot \hat{\theta})^T \cdot f'(\hat{z}_t^T \cdot \hat{\theta}) \quad (2.9)$$

και

$$\frac{\partial y_t}{\partial y_{t-1}} = W(1, :) \cdot (f'(\text{diag}(\hat{W}^T \cdot \hat{x}_{(t-1)})) \cdot \hat{\theta}) \cdot f'(\hat{z}_t^T \cdot \hat{\theta}) \quad (2.10)$$

Συνοψίζοντας, οι παράγωγοι των εξόδων των νευρωνικών δικτύων ως προς τις ρυθμίσιμες παραμέτρους χρειάζονται σε διάφορες εργασίες μάθησης όταν χρησιμοποιείται βελτιστοποίηση με μέθοδο κλίσης. Για παράδειγμα, στην Επιβλεπόμενη Μάθηση, δίνεται μια ακολουθία εισόδου και μια ακολουθία στόχος  $(\psi)_t$  και η απόδοση της εκπαίδευσης  $U$  μπορεί να εξαρτάται από τετράγωνο της διαφοράς των ακολουθιών εξόδου και στόχου, δηλαδή

$$U = \sum_{t=1}^T (y_t - \psi_t)^2 \quad (2.11)$$

και η παράγωγος που μας ενδιαφέρει τελικά είναι η

$$\frac{\partial U}{\partial w} = \sum_{t=1}^T \frac{\partial (y_t - \psi_t)^2}{\partial y_t} \cdot \frac{\partial y_t}{\partial w} \quad (2.12)$$

Στην άμεση Ενισχυτική Μάθηση που χρησιμοποιεί ANN, η συνάρτηση μάθησης  $U$  μπορεί να σχετίζεται με τις ρυθμιζόμενες παραμέτρους μέσω μιας αλυσίδας που περιλαμβάνει τις ανταμοιβές και τις εξόδους του ANN, επομένως στο επεκταμένο γράφημα που δείχνει αυτές τις εξαρτήσεις μπορούμε να κάνουμε οπισθοδιάδοση απ'την τελική έξοδο  $U$  μέχρι κάθε παράμετρο για να βρούμε κάθε παράγωγο  $\frac{\partial U}{\partial w}$ . Η οπισθοδιάδοση αρχικά φτάνει ως τις εξόδους του ANN, κάτι που γενικά είναι σχετικά απλό, και μετά αρχίζει η οπισθοδιάδοση απ'τις εξόδους του ANN προς τις ρυθμιζόμενες παραμέτρους με τις μεθόδους που είδαμε σε αυτό το κεφάλαιο.

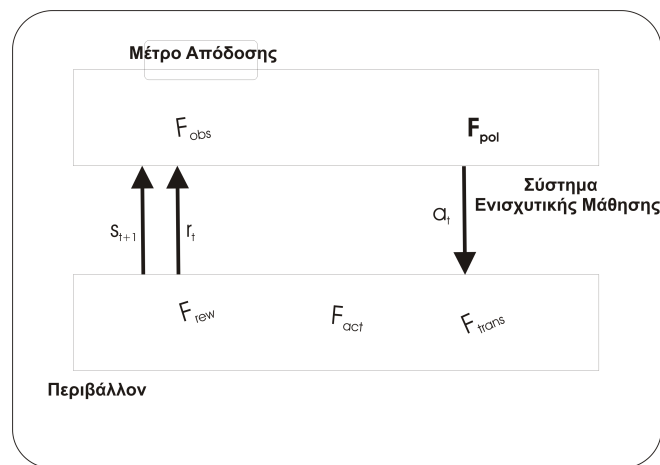
# Κεφάλαιο 3

## Ενισχυτική Μάθηση

### 3.1 Διατύπωση του προβλήματος

Η Ενισχυτική Μάθηση είναι ένα γενικό θεωρητικό πλαίσιο μάθησης όπου το σύστημα που μαθαίνει βρίσκεται μέσα σε ένα άγνωστο περιβάλλον και προσπαθεί να μάθει να συμπεριφέρεται έτσι ώστε να μεγιστοποιήσει ένα δεδομένο μέτρο απόδοσης. Τυπικά υπάρχουν τα εξής:

- σύνολο  $S$  από καταστάσεις στις οποίες μπορεί να βρίσκεται το περιβάλλον
- σύνολο  $O$  από παρατηρήσεις, δηλαδή πληροφορίες που αντιλαμβάνεται το σύστημα
- συνάρτηση,  $F_{obs} : S \rightarrow O$ , που περιγράφει την αντίληψη του συστήματος ανάλογα με την κατάσταση του περιβάλλοντος
- σύνολο  $A$  από κινήσεις τις οποίες μπορεί να κάνει το σύστημα
- συνάρτηση,  $F_{trans} : S \times A \rightarrow S$ , που περιγράφει τις μεταβάσεις του περιβάλλοντος
- συνάρτηση,  $F_{act} : S \rightarrow \mathcal{P}(A)$ , που περιγράφει τις επιτρεπτές κινήσεις ανάλογα με την κατάσταση
- αριθμός  $R \in \mathcal{R}$  που παριστάνει την ανταμοιβή
- συνάρτηση,  $F_{rew} : S \times A \times S \rightarrow \mathcal{R}$ , που περιγράφει την ανταμοιβή μιας μετάβασης
- μέτρο απόδοσης  $U$ , σχετικό με τις ανταμοιβές



Σχήμα 3.1: Σχήμα Ενισχυτικής Μάθησης

Την χρονική στιγμή  $t$ , η κατάσταση του περιβάλλοντος συμβολίζεται με  $s_t$ , και το σύστημα έχει να επιλέξει μια εκ των κινήσεων  $F_{act}(s_t)$ . Η επιλογή συμβολίζεται με  $a_t$ . Στην συνέχεια, το περιβάλλον θα μεταβεί στην κατάσταση  $s_{t+1} = F_{trans}(s_t, a_t)$ , και το σύστημα δέχεται την ανταμοιβή  $R_t = F_{rew}(s_t, a_t, s_{t+1})$ . Καθώς περνάει ο χρόνος, ο στόχος είναι η μεγιστοποίηση του μέτρου απόδοσης το οποίο συνήθως είναι μια χρονική συσσώρευση ανταμοιβής.

Για να επιτευχθεί αυτός ο στόχος πρέπει να βρεθεί μια κατάλληλη Πολιτική δηλαδή

- συνάρτηση  $F_{pol} : O \rightarrow A$ , που περιγράφει την κίνηση που επιλέγει το σύστημα ανάλογα με την παρατήρηση που αντιλαμβάνεται

Όλες αυτές οι συναρτήσεις μπορεί να είναι στοχαστικές και η κατάσταση  $s_t$  μπορεί να είναι πληροφορία που σηματοδοτεί δυναμικό γεγονός δηλαδή να περιλαμβάνει δεδομένα που ήταν διαθέσιμα και σε προηγούμενες χρονικές στιγμές.

Επιπλέον μπορούν να γίνουν διάφορες πιο γενικές διατυπώσεις του προβλήματος, όπου εκτός από στοχαστικές οι συναρτήσεις μπορεί να είναι χρονικά μεταβαλλόμενες ή και άγνωστες, αν και κάτι τέτοιο μπορεί να κάνει την εύρεση λύσης από πολύ δύσκολη έως αδύνατη. Επομένως όσον αφορά προβλήματα των οποίων το περιβάλλον αντιστοιχεί σε ένα περιβάλλον του πραγματικού κόσμου, το πρώτο βήμα σε μια χρήσιμη εφαρμογή της Ενισχυτικής Μάθησης είναι η ανακάλυψη ενός καλού σημείου μεταξύ υπολογιστικής πολυπλοκότητας και αφαιρετικής έντασης.

Έχοντας επιλέξει μια αφαίρεση, το επόμενο βασικό ερώτημα είναι αν υπάρχει κάποια Πολιτική που να αυξάνει το μέτρο απόδοσης σε ικανοποιητικά επίπεδα. Αυτό το ερώτημα ενσωματώνει την εκ των προτέρων άγνοια τόσο για επαρκή

αιτιοκρατικότητα του πραγματικού περιβάλλοντος, όσο και για θέματα απόδοσης της υπολογιστικής διαδικασίας λύσης πάνω στην επιλεγμένη αφαίρεση.

Στην συνέχεια θα αναπτυχθούν οι δύο βασικοί τρόποι επίλυσης του προβλήματος, ο έμμεσος και ο άμεσος.

## 3.2 Έμμεσος τρόπος επίλυσης

Για να λυθεί ένα πρόβλημα Ενισχυτικής Μάθησης με τον έμμεσο τρόπο επίλυσης μια απ' τις βασικές προϋποθέσεις είναι ότι το περιβάλλον μοντελοποιείται σαν μια Μαρκοβιανή Διαδικασία Αποφάσεων (*Markov Decision Process – MDP*). Η αξιολογήση μιας Πολιτικής με βάση αυτήν την *MDP* γίνεται με τον τύπο

$$\rho(F_{pol}) = E\{U|F_{pol}\} \quad (3.1)$$

δηλαδή με την αναμενόμενη απόδοση όταν ακολουθείται η Πολιτική  $F_{pol}$ . Η τυχαία μεταβλητή  $U$  που συμβολίζει το μέτρο απόδοσης, είναι συνήθως μια χρονική συσσώρευση των ανταμοιβών

$$U = \sum_{t=1}^{end} R_t \quad (3.2)$$

ή αν ο χρονικός ορίζοντας είναι άπειρος χρησιμοποιείται ένας αποσβήνων παράγοντας  $\gamma$ , μεταξύ 0 και 1

$$U = \sum_{t=1}^{\infty} \gamma^k R_t \quad (3.3)$$

Η διαδικασία που ακολουθείται για να βελτιστοποιηθεί ο χώρος Πολιτικών είναι ο Προσεγγιστικός Δυναμικός Προγραμματισμός. Ο πιο βασικός ορισμός αυτής της μεθόδου είναι η συνάρτηση *value function*

$$V^{\pi}(s) = E_{\pi}\{U_t|s_t\} \quad (3.4)$$

που αναπαριστά το αναμενόμενο μελλοντικό κέρδος όταν η τρέχουσα κατάσταση είναι η  $s_t$  και ακολουθώντας την πολιτική  $F_{pol} = \pi$ . Με βάση αυτόν τον ορισμό το πρόβλημα ανάγεται στο να βρεθεί μια Πολιτική η οποία βελτιστοποιεί την *value function* για τυχαία αρχική κατάσταση  $s_{init}$ . Ένας δεύτερος χρήσιμος ορισμός είναι η συνάρτηση  $Q$  δηλαδή το αναμενόμενο μελλοντικό κέρδος όταν η τρέχουσα κατάσταση είναι η  $s_t$ , συμβαίνει η κίνηση  $\alpha_t$  και ακολουθείται η πολιτική  $\pi$  από εκεί και πέρα

$$Q^{\pi}(s, \alpha) = E_{\pi}\{U_t|s_t = s, \alpha_t = \alpha\} \quad (3.5)$$

Αναφορικά με το *MDP* περιβάλλον χρησιμοποιούμε τον συμβολισμό

$$\mathcal{P}_{ss'}^\alpha = Pr\{s_{t+1} = s' | s_t, \alpha_t\} \quad (3.6)$$

για την πιθανότητα να περάσει το περιβάλλον από την κατάσταση  $s$  στην  $s'$  μέσω της κίνησης  $\alpha$  και

$$\mathcal{R}_{ss'}^\alpha = E\{r_{t+1} | s_t, \alpha_t, s_{t+1}\} \quad (3.7)$$

για την αναμενόμενη ανταμοιβή που δίνει αυτό το πέρασμα. Με λίγες πράξεις προκύπτουν οι παρακάτω εξισώσεις *Bellman* στις οποίες βασίζεται η λύση του προβλήματος μέσω Δυναμικού Προγραμματισμού. Παραδείγματα τέτοιων μεθόδων είναι οι *temporal difference learning* μέθοδοι και η μέθοδος *Q-learning*.

$$V^\pi(s) = E_\pi\{U_t | s_t = s\} = \quad (3.8)$$

$$= \sum_\alpha \pi(s, \alpha) \sum_{s'} \mathcal{P}_{ss'}^\alpha [\mathcal{R}_{ss'}^\alpha + \gamma E_\pi\{\sum_{k=0}^{\infty} \gamma^k R_{t+k+2} | s_{t+1} = s'\}]$$

$$V^*(s) = \max_{\alpha \in \mathcal{A}(s)} Q^{\pi^*}(s, \alpha) = \quad (3.9)$$

$$= \max_\alpha E_{\pi^*}\{R_t | s_t = s, \alpha_t = \alpha\} = \max_{\alpha \in \mathcal{A}(s)} \sum_{s'} \mathcal{P}_{ss'}^\alpha [\mathcal{R}_{ss'}^\alpha + V^*(s')]$$

Στην βιβλιογραφία [21][22][23][33] υπάρχει λεπτομερής ανάπτυξη του έμμεσου τρόπου επίλυσης. Η Αναδρομική Ενισχυτική Μάθηση που θα παρουσιασθεί στο κεφάλαιο 6 και θα χρησιμοποιηθεί στο κεφάλαιο 7, χρησιμοποιεί τον άμεσο τρόπο επίλυσης που ακολουθεί στην επόμενη παράγραφο.

### 3.3 Άμεσος τρόπος επίλυσης

Ο άμεσος τρόπος επίλυσης είναι η κατ'ευθείαν αναγωγή του προβλήματος σε βελτιστοποίηση στον χώρο των Πολιτικών. Η Πολιτική εκφράζεται μέσω ενός πεπερασμένου συνόλου από παραμέτρους και έτσι δημιουργείται ο χώρος των Πολιτικών στον οποίον γίνεται αναζήτηση. Γενικά, υπάρχουν δύο επιλογές για μια τέτοια αναζήτηση ανάλογα με το αν η μέθοδος εξαρτάται από την Κλίση της Πολιτικής (policy gradient) ή όχι.

Συμβολίζοντας το σύνολο παραμέτρων με  $\theta$ , και την αντίστοιχη Πολιτική με  $F_{pol}(\theta)$ , βασιζόμαστε στην συνάρτηση απόδοσης του προβλήματος  $F_{rew}$  ώστε να ορίσουμε μια συνάρτηση απόδοσης  $U(F_{pol}(\theta))$  που αξιολογεί μια ολόκληρη Πολιτική. Η αναζήτηση μπει να γίνει με gradient descent αν η  $U(\theta)$  που προκύπτει είναι αρκετά ομαλή και είναι γνωστή η αναλυτική έκφραση της κλίσης της. Παραδείγματα τέτοιων αλγορίθμων είναι ο REINFORCE και η

Αναδρομική Ενισχυτική Μάθηση στην οποία εστιάζουμε σε αυτήν την εργασία.

Σημειώνουμε τις εναλλακτικές λύσεις πέρα από το gradient descent και γενικά τις μεθόδους κλίσης αναφέροντας ότι για να γίνει βελτιστοποίηση μια συνάρτησης, δεν είναι αναγκαίο να χρησιμοποιηθεί πληροφορία που σχετίζεται με την κλίση της. Έχουν χρησιμοποιηθεί με επιτυχία διάφορες μέθοδοι χωρίς κλίση, μερικές από τις οποίες βασίζονται σε γνωστές φυσικές και βιολογικές διεργασίες όπως η Προσομοιωμένη Ανόπτηση και η Εξέλιξη.

Συγκρίνοντας αυτά τα δυο είδη βελτιστοποίησης, οι μέθοδοι κλίσης έχουν το μειονέκτημα ότι μπορεί να κολλήσουν σε τοπικώς βέλτιστα σημεία. Αν η συνάρτηση προς βελτιστοποίηση είναι αρκετά ομαλή και δεν έχει πολλά τοπικά βέλτιστα, συνήθως λειτουργούν αποδοτικά. Οι μέθοδοι χωρίς κλίση δεν έχουν αυτό το πρόβλημα της τοπικής αναζήτησης ωστόσο είναι δύσκολο να αναπτυχθούν σχετικά θεωρήματα που εξασφαλίζουν κάποιου είδους σύγκλιση στην λύση ή έστω δίνουν κάποια πιθανότητα να βρεθεί το ολικό βέλτιστο. Η ίδια η μορφή της συνάρτησης προς βελτιστοποίηση ορίζει ποιά μέθοδος είναι καταλληλότερη, ωστόσο στα περισσότερα προβλήματα αυτή η μορφή είναι άγνωστη.

## Μάθηση με *gradient descent*

Συμβολίζουμε με  $ss_t$  το υποσύνολο της κατάστασης  $s_t$  που δεν εξαρτάται από την τρέχουσα Πολιτική. Όπως η Επιβλεπόμενη Μάθηση οδηγείται με το εκπαιδευτικό σύνολο εισόδων-στόχων, έτσι και η Ενισχυτική Μάθηση με τον άμεσο τρόπο μπορούμε να θεωρήσουμε ότι οδηγείται από την ακολουθία καταστάσεων  $ss_t$ . Η ακολουθία αυτή είναι σταθερή σε έναν δεδομένο χρονικό ορίζοντα όπου αναζητείται η βέλτιστη Πολιτική. Θεωρούμε ότι η συνάρτηση μάθησης  $U(\theta; (ss)_t)$ , ή απλά  $U(\theta)$ , αναφέρεται στην απόδοση του συστήματος πάνω σε αυτήν την σταθερή ακολουθία. Συμβολίζοντας την κλίση της  $U$  ως προς  $\theta$  με  $G(\theta)$ , η βασική *gradient descent* μάθηση παίρνει την μορφή

$$\theta_{k+1,1} = \theta_{k,1} + \rho \cdot G(\theta_{k,1}) \quad (3.10)$$

Ο συμβολισμός  $\theta_{k,t}$  σημαίνει τις παραμέτρους όταν έχουν συμβεί  $k$  εποχές και το σημείο ανάμεσα στις καταστάσεις  $ss_{t-1}$  και  $ss_t$  είναι αυτό που οδηγεί την μάθηση και μια εποχή λήγει όταν όλη ακολουθία έχει χρησιμοποιηθεί. Η παραπάνω εξίσωση ονομάζεται *batch* μάθηση επειδή όλη η ακολουθία  $(ss)_t$  οδηγεί την διαδικασία με μαζικό τρόπο σε αντίθεση με την *online* μάθηση

$$\theta_{k,t+1} = \theta_{k,t} + \rho \cdot g(ss_t, \theta_{k,t}) \quad (3.11)$$

$$t = 1 \dots N, \quad \theta_{k+1,1} = \theta_{k,N+1}$$



όπου  $g$  είναι η κλίση της απόδοσης της Πολιτικής στο σημείο  $ss_t$  και  $N$  το μήκος της  $(ss)_t$ . Η  $g$  θεωρείται *online* ανάλογο της  $G$  όταν ερμηνεύεται ως κλίση της τοπικής συνεισφοράς στην  $U$ . Στην περίπτωση που ισχύει

$$G(\underline{\theta}) = \sum_{t=1}^N g(ss_t, \underline{\theta}) \quad (3.12)$$

, ενδεχομένως επειδή η μαζική απόδοση είναι άθροισμα των αποδόσεων σε όλα τα σημεία όπως για παράδειγμα στην Επιβλεπόμενη Μάθηση με τετραγωνικό σφάλμα, τότε η *batch* μάθηση εφαρμόζοντας αναδρομή παίρνει την παρακάτω μορφή που είναι παρόμοια με την *online*

$$\underline{\theta}_{k,t+1} = \underline{\theta}_{k,t} + \rho \cdot g(ss_t, \underline{\theta}_{k,t}) \quad (3.13)$$

$$t = 1 \dots N, \quad \underline{\theta}_{k+1,1} = \underline{\theta}_{k,N+1}$$

Ο δείκτης  $k$  των εποχών τρέχει ανάμεσα στις τιμές 1 και  $n_e$  όπου η  $n_e$  δηλαδή ο συνολικός αριθμός εποχών είναι, όπως και ο ρυθμός μάθησης  $\rho$ , μια βασική παράμετρος της διαδικασίας μάθησης .

### 3.4 Σύγκριση έμμεσου και άμεσου τρόπου

Με την διατύπωση της Ενισχυτικής Μάθησης στην παράγραφο 3.1 φαίνεται ότι μπορούν να υπάρχουν Πολιτικές που εκμεταλλεύονται καθυστερημένες ανταμοιβές. Δηλαδή το σύστημα μπορεί να μάθει να κάνει επιλογές με βάση ποιά κίνηση αναμένεται να δώσει την μεγαλύτερη ανταμοιβή όχι άμεσα αλλά σε κάποιο μακρύ χρονικό ορίζοντα. Όταν οι καθυστερημένες ανταμοιβές παίζουν σημαντικό ρόλο σε ένα περιβάλλον, είναι πολύ χρήσιμο να υπάρχει ένα αξιόπιστο μοντέλο του περιβάλλοντος στο οποίο θα βασιστεί ο σχεδιασμός της Πολιτικής. Ο έμμεσος τρόπος της παραγράφου 3.2 προσεγγίζει το πρόβλημα με έμφαση σε αυτές τις καθυστερημένες ανταμοιβές θεωρώντας το περιβάλλον ως *MDP* και προχωρώντας στην λύση με Προσεγγιστικό Δυναμικό Προγραμματισμό.

Στην περίπτωση που σε ένα περιβάλλον οι ευκαιρίες για καθυστερημένες ανταμοιβές δεν είναι έντονες σε σχέση με τις άμεσες ανταμοιβές, η αμεσότητα της παραγράφου 3.3 ενδέχεται να είναι πιο αποδοτική. Ειδικά όταν ο χώρος των Πολιτικών είναι σχετικά μικρός και μπορεί να δομηθεί έτσι ώστε να είναι εφικτή η εύρεση καλών σημείων του τότε η βελτιστοποίηση στον χώρο των Πολιτικών δείχνει ο πιο προσοδοφόρος τρόπος αντιμετώπισης του προβλήματος. Για το εμπόριο σε μια συγκεκριμένη αγορά όπως θα φανεί και στα επόμενα κεφάλαια, ο χώρος των Πολιτικών πράγματι είναι σχετικά μικρός επειδή οι επιτρεπόμενες κινήσεις είναι μόνο δύο, η Πώληση και η Αγορά, και επιπλέον όταν οι εμπορικές κινήσεις είναι *spot* και όχι σχετιζόμενες με τον μελλοντικό

χρόνο και με περίπλοκες πληροφορίες, όπως τα *swaps*, *options* κτλπ τότε οι άμεσες ανταμοιβές κυριαρχούν.

## Κεφάλαιο 4

# Η Αγορά Συναλλάγματος

Οι οικονομικές αγορές είναι ο φυσικός χώρος της Διαχείρισης Ιδιοκτησίας. Η λειτουργία τους δίνει την δυνατότητα στα μέλη τους να ανταλλάσουν διάφορα αντικείμενα αξίας σε τιμές που καθορίζονται από τον νόμο της ζήτησης και της προσφοράς και έτσι προκαλεί αύξηση του συλλογικού συμφέροντος. Συγκεκριμένα οι αγορές μετοχών και ομολόγων αυξάνουν το κεφάλαιο, οι αγορές χρήματος και οι αγορές παραγώγων συμβάλλουν στην μεταφορά της ρευστότητας και του ρίσκου αντίστοιχα, και τέλος η αγορά συναλλάγματος επιτρέπει το διεθνές εμπόριο.

Μια απ' τις βασικές μονάδες μιας αγοράς είναι ένας δημιουργός αγοράς. Η τυπική λειτουργία ενός δημιουργού είναι να δέχεται μέλη της αγοράς και να σημειώνει την διάθεση τους να πουλήσουν ή να αγοράσουν διάφορα αντικείμενα. Για κάποιο συγκεκριμένο αντικείμενο κάποιο μέλος μπορεί να θέσει την τιμή ζήτησης (*ask price*), δηλαδή την τιμή στην οποία διατίθεται να αγοράσει αυτό το αντικείμενο, και την τιμή προσφοράς (*bid price*) δηλαδή την τιμή στην οποία διατίθεται να πουλήσει αυτό το αντικείμενο. Ο δημιουργός σημειώνει στο τετράδιο παραγγελιών του όλες αυτές τις τιμές που σχετίζονται με όλα τα εισερχόμενα μέλη, και λαμβάνοντας υπόψιν τα περιεχόμενα του τετραδίου ενώνει σε ζευγάρια τους ενδιαφερόμενους για πώληση και για αγορά έτσι ώστε οι διαθέσεις τους να ταιριάζουν. Με αυτόν τον τρόπο δημιουργείται ρευστότητα για την αγορά με αντάλλαγμα το κέρδος απ' τις διαφορές των τιμών ζήτησης και προσφοράς για ανταλλαγές που συμβαίνουν.

Η Αγορά Συναλλάγματος είναι μια παγκόσμια αποκεντρωμένη αγορά που επιτρέπει την ανταλλαγή νομισμάτων διαφορετικών χωρών. Ο βασικός λόγος που προκαλεί αυτές τις ανταλλαγές είναι το διεθνές εμπόριο, δηλαδή όταν κάποιος πολίτης ή μια εταιρεία μιας χώρας θέλει να αγοράσει προϊόντα από κάποια ξένη χώρα πρέπει πρώτα να περάσει πρώτα απ' την αγορά συναλλάγματος για να αγοράσει νόμισμα της ξένης χώρας. Μετά τον Β' παγκόσμιο πόλεμο και μέχρι το 1970, η τιμή μεταξύ των νομισμάτων των περισσότερων μεγάλων

χωρών ήταν δεσμευμένη με βάση το σύστημα *Bretton Woods*. Από το 1973, ωστόσο, σταδιακά όλες οι τιμές τέθηκαν να κυμαίνονται ελεύθερα, δίνοντας έτσι την βάση για έναν δεύτερο λόγο συμμετοχής στην αγορά συναλλάγματος · τις κερδοσκοπικές επενδύσεις. Στο διάγραμμα 4.1 φαίνεται ένα παράδειγμα κυματισμού της τιμής του ζεύγους Ευρώ και Τσέχικης Κορόνας (*EurCzk*) σε όλο το έτος 2012.

Τα ειδικά χαρακτηριστικά της αγορά συναλλάγματος είναι ότι παραμένει ανοιχτή εικοσιτέσσερις ώρες το εικοσιτετράωρο, οι παράγοντες επιρροής της είναι γεωγραφικά και πολιτικά αποκεντρωμένοι, και ότι είναι η πιο ρευστή αγορά του κόσμου ενώ ο όγκος της είναι πολύ μεγάλος. Έτσι θεωρείται ως η καλύτερη προσέγγιση στην ιδέα της αγοράς με τέλειο ανταγωνισμό.

Σύμφωνα με την Τράπεζα Διεθνών Διακανονισμών [27], το 2010 η ημερήσια συναλλαγή ήταν περίπου 4 τρισεκατομύρια δολάρια με το παγκόσμιο προϊόν να είναι στα 63 τρις. Από αυτά τα 4 τρις

- 1.49 ήταν σε συναλλαγές στιγμής (*spot exchange*) δηλαδή σε άμεσες συναλλαγές που αποφασίζονται και εκτελούνται την ίδια στιγμή
- 0.475 σε *forwards* τα οποία είναι συμβόλαια για συναλλαγή σε κάποια αποφασισμένη μελλοντική ημερομηνία
- 1.765 σε *swaps* στα οποία γίνεται συναλλαγή για ένα ορισμένο χρονικό διάστημα με την συμφωνία να γίνει αντιστροφή της συναλλαγής σε μεταγενέστερη ημερομηνία.
- 0.207 σε *options* τα οποία είναι *forwards* με επιπλέον χαρακτηριστικό ότι η μελλοντική συναλλαγή θα γίνει υπό προϋποθέσεις











Στον πίνακα 4.1 φαίνεται η ημερήσια συμμετοχή όλων των μεγάλων νομισμάτων στο έτος 2010, σύμφωνα με την παραπάνω έρευνα, και στον πίνακα 4.2 οι δέκα κορυφαίοι συμμετέχοντες στην αγορά το έτος 2013 έως τον Μάιο.

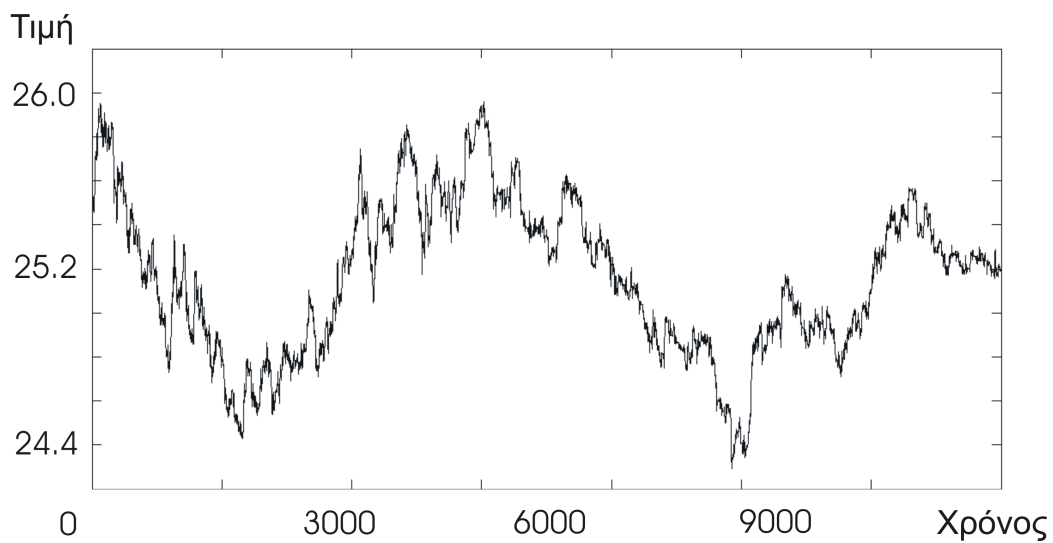
Οι κερδοσκοπικές επενδύσεις, ως ένας βασικός λόγος που οδηγεί κάποιον να συμμετέχει στην Αγορά Συναλλάγματος, μπορούν να υλοποιούνται είτε χειροκίνητα είτε με κάποια αυτόματη αλγοριθμική εμπορική στρατηγική μέσω του υπολογιστή. Η δημοσίευση *The 4 trillion question: what explains FX growth since the 2007 survey?* [28] αναφέρει ότι το ποσοστό της αυτόματης συμμετοχής ανήλθε από 2% το 2004, σε 49% το 2010 ένα μεγάλο ποσοστό απ'την οποία θεωρείται ότι υλοποιεί το λεγόμενο Εμπόριο Υψηλής Συχνότητας (*HFT*) όπου οι αποφάσεις παίρνονται σε κλάσματα της ώρας ή ακόμη και σε επίπεδο μικρότερο του λεπτού της ώρας.

Πίνακας 4.1: Τα πιο ανταλασσόμενα νομίσματα για το σύμφωνα με την Τράπεζα Διεθνών Διακανονισμών, 2010(Απρίλιος)

	νόμισμα	σύμβολο	ημερήσιο ποσοστό
1	 Αμερικάνικο Δολλάριο	<i>Usd</i>	84.9%
2	 Ευρώ	<i>Eur</i>	39.1%
3	 Γιαπωνέζικο Γιεν	<i>Jpy</i>	19.0%
4	 Αγγλική Λίρα	<i>Gbp</i>	12.9%
5	 Δολλάριο Αυστραλίας	<i>Aud</i>	7.6%
6	 Ελβετικό Φράγκο	<i>Chf</i>	6.4%
7	 Δολλάριο Καναδά	<i>Cad</i>	5.3%
8	 Δολλάριο Χονγκ Κονγκ	<i>Hkd</i>	2.4%
9	 Σουηδική Κορόνα	<i>Sek</i>	2.2%
10	 Δολλάριο Νέας Ζηλανδίας	<i>Nzd</i>	1.6%
11	 Γουόν Νότιας Κορέας	<i>Krw</i>	1.5%
12	 Δολλάριο Σιγκαπούρης	<i>Sgd</i>	1.4%
13	 Νορβηγική Κορόνα	<i>Nok</i>	1.3%
14	 Μεξικάνικο Πέσο	<i>Mxn</i>	1.3%
15	 Ρούπι Ινδίας	<i>Inr</i>	0.9%
	Άλλα νομίσματα		12.2%
	Σύνολο		200.0%

Πίνακας 4.2: Οι οργανισμοί με την μεγαλύτερη συμμετοχή, 2013(Μάιος)

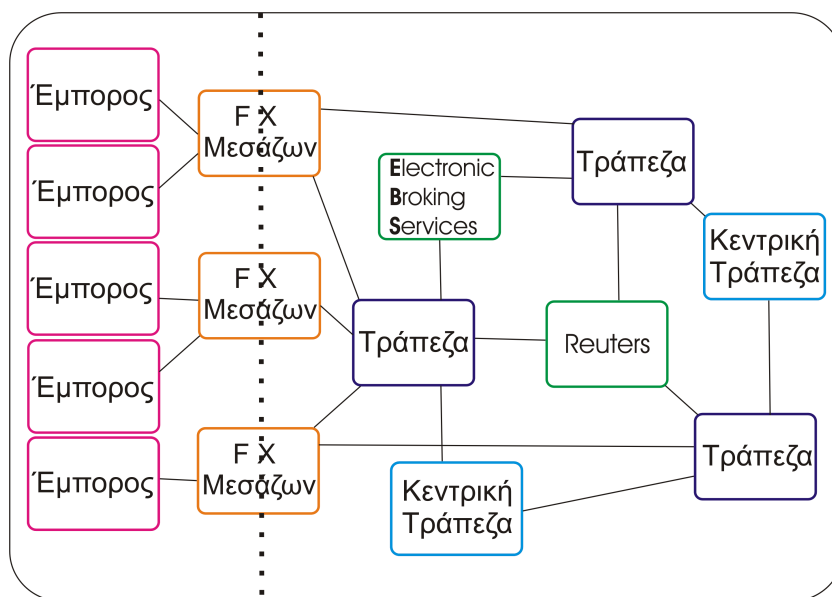
	όνομα	ποσοστό
1	 <i>Deutsche Bank</i>	15.18%
2	 <i>Citi</i>	14.90%
3	 <i>Barclays Investment Bank</i>	10.24%
4	 <i>UBS AG</i>	10.11%
5	 <i>HSBC</i>	6.93%
6	 <i>JPMorgan</i>	6.07%
7	 <i>Royal Bank of Scotland</i>	5.62%
8	 <i>Credit Suisse</i>	3.70%
9	 <i>Morgan Stanley</i>	3.15%
10	 <i>Bank of America Merrill Lynch</i>	3.08%



Σχήμα 4.1: Παράδειγμα κυματισμού της τιμής του συμβόλου *EurCzk*, Ευρώ και Τσέχικη Κορόνα, έτος 2012

Στο διάγραμμα 5.1 φαίνονται το διατραπεζικό επίπεδο και το επίπεδο όπου βρίσκονται οι μικρού όγκου έμποροι, τα οποία αποτελούν τα δύο μέρη της σύγχρονης Αγοράς Συναλλάγματος. Η βασική πληροφορία είναι οι τιμές των συμβόλων  $XxxYyy$  με τα  $Xxx$  και  $Yyy$  να αναπαριστούν την μονάδα χρήματος δύο διαφορετικών χωρών. Οι δημιουργοί αγοράς παρέχουν ρευστότητα στους μικρούς έμπορους ενώνοντάς τους με το διατραπεζικό επίπεδο. Ο γενικός μηχανισμός που επηρεάζει τις τιμές είναι η ζήτηση και η προσφορά και οδηγείται απ'την διαχείριση ρίσκου και γενικότερα απ'τις διαδικασίες διαχείρισης ιδιοκτησίας που εφαρμόζουν οι διάφοροι συμμετέχοντες στην αγορά [24].

Στην συνέχεια του κεφαλαίου και στα υπόλοιπα κεφάλαια, θα βλέπουμε την Αγορά Συναλλάγματος απ'την οπτική ενός μικρού όγκου εμπόρου και θα αντιμετωπίσουμε τον μεσάζοντα που μας ενώνει με το διατραπεζικό επίπεδο ως μελανή δεξαμενή ρευστότητας (*dark pool of liquidity*) δηλαδή ως δημιουργό αγοράς που απλά μας παρέχει ρευστότητα παρουσιάζοντας μας *bid/ask* τιμές και χωρίς να δημοσιοποιεί τον μηχανισμό στο οποίο βασίζονται αυτές οι τιμές. Πρακτικά αυτό σημαίνει ότι δεν θα παίρνουμε υπόψιν τον χειρισμό που κάνει ο μεσάζοντας στο βιβλίο παραγγελιών (*order's book*) του, ούτε πληροφορίες σχετικά με το διατραπεζικό επίπεδο.



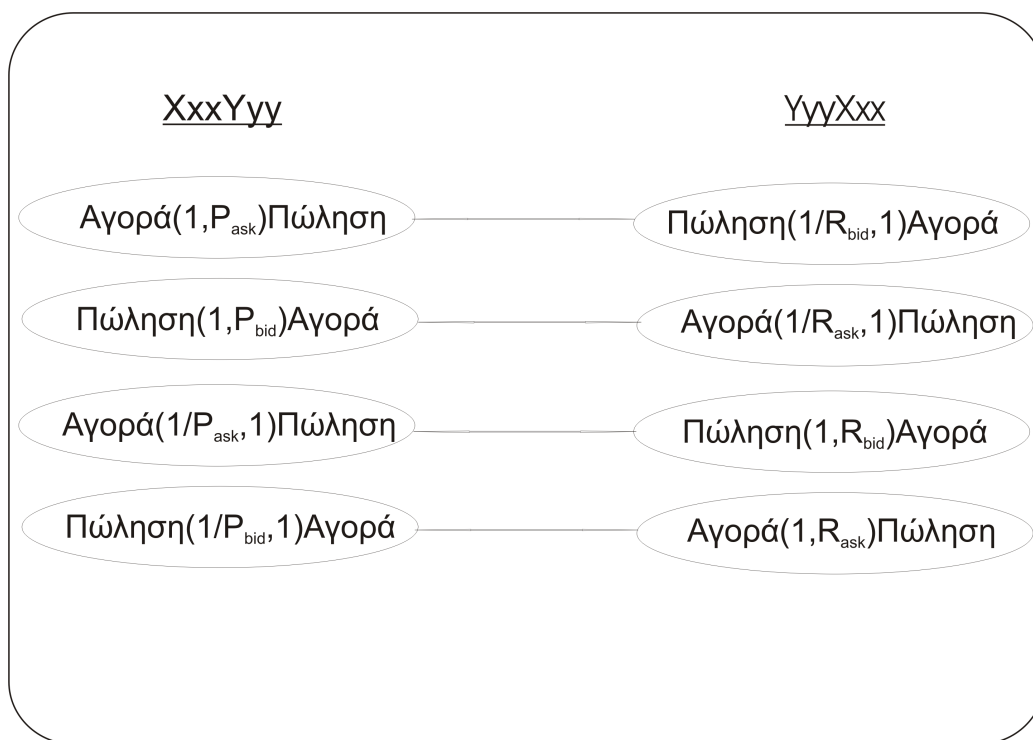
Σχήμα 4.2: Η δομή της σύγχρονης Αγοράς Συναλλάγματος

## 4.1 Είσοδος

Η πρώτη επιλογή που έχουμε όταν μπαίνουμε στο περιβάλλον είναι η χώρα της οποίας την μονάδα χρήματος θα χρησιμοποιήσουμε ως βάση. Συμβολίζουμε αυτήν την παράμετρο με  $HOME$ . Το σύνολο των χρημάτων που έχουμε, συμβολίζεται με  $Bal$  (*Balance*), θα είναι σε  $HOME$  μονάδες και αρχικά είναι 0. Για να επιτραπεί η είσοδος πρέπει να παραχωρήσουμε ένα ποσό χρημάτων  $Bal_{init}$ . Τώρα είναι  $Bal = Bal_{init}$ .

## 4.2 Σύμβολα και Τιμές

Απτόν δημιουργό αγοράς μας παρέχεται η τιμή ζήτησης και η τιμή προσφοράς για κάθε σύμβολο  $XxxYyy$ . Η διαφορά μεταξύ αυτών των δύο τιμών ονομάζεται *spread*. Η χρηματική μονάδα  $Xxx$  ονομάζεται Βάση του συμβόλου και η  $Yyy$  Λόγος (*quote*). Οι λέξεις προσφορά και ζήτηση αναφέρονται στην Βάση, δηλαδή η τιμή προσφοράς του συμβόλου  $EurUsd$  σημαίνει ότι αν πουλήσουμε 1 Ευρώ, παίρνουμε  $Pr_{bid}$  Δολλάρια, και η Τιμή ζήτησης σημαίνει ότι αν αγοράσουμε 1 Ευρω, δίνουμε  $Pr_{ask}$  Δολλάρια. Για δύο δεδομένες μονάδες χρήματος  $Xxx$  και  $Yyy$  χρησιμοποιείται μόνο ένα σύμβολο, για παράδειγμα το  $XxxYyy$  και όχι το  $YyyXxx$ , ωστόσο γράφοντας ως  $\{Pr_{bid}, Pr_{ask}\}$  τις τιμές του  $XxxYyy$  οι αντίστοιχες τιμές για το  $YyyXxx$  είναι  $\{\frac{1}{Pr_{ask}}, \frac{1}{Pr_{bid}}\}$ . Εξαντλητικά, κάθε *bid/ask* αγοραπωλησία των νομισμάτων  $Xxx$  και  $Yyy$  φαίνεται στο



Σχήμα 4.3: *bid/ask* αγοροπωλησίες,  $P$  είναι η τιμή του  $XxxYyy$  και  $R$  η τιμή του  $YyyXxx$

σχήμα 4.3.

### 4.3 Συναλλαγές και Κέρδη

Οι τύποι συναλλαγών που ονομάζονται παράγωγα (*derivatives*) όπως τα *swaps*, τα *futures* και τα *options* έχουν το εξής χαρακτηριστικό: η συμμετοχή σε μια τέτοια συναλλαγή προϋποθέτει ότι θεωρούμε τρέχουσα κατάσταση της αγοράς αρκετά περισσότερα από τις πρόσφατες τιμές. Αντιθέτως, μια μικρή κατάσταση, που διευκολύνει την τεχνική ανάλυση, με το μόνο περιεχόμενο τις πρόσφατες τιμές σηματοδοτεί αρκετή πληροφορία για να συμμετέχει κανείς σε συναλλαγές στιγμής (*spot*). Στην συνέχεια θα μας απασχολήσουν μόνο συναλλαγές στιγμής και θα αναλύσουμε τον τρόπο που γίνεται μια τέτοια συναλλαγή και πως υπολογίζεται το κέρδος της.

Έστω ότι θέλουμε να κάνουμε εμπόριο με ένα σύμβολο,  $XxxYyy$ . Η πρώτη επιλογή μας είναι ο αριθμός μονάδων τον οποίον συμβολίζουμε με  $\mu$ . Στην συνέχεια πρέπει να αποφασίσουμε αν πουλήσουμε το σύμβολο (*going short*) ή αν το αγοράσουμε (*going long*).



Αρχικά για απλότητα δεν θα κάνουμε διάκριση της τιμής σε *bid/ask* τιμές. Το άνοιγμα και η ολοκλήρωση μιας θέσης αγοράς του συμβόλου  $XxxYyy$  κατά  $\mu$  μονάδες σημαίνει ότι την χρονική στιγμή  $t_1$  αγοράζουμε  $\mu$   $Xxx$  μονάδες πουλώντας  $\mu \cdot Pr_{t1}$   $Yyy$  μονάδες και την μεταγενέστερη χρονική στιγμή  $t_2$  πουλάμε  $\mu$   $Xxx$  μονάδες αγοράζοντας  $\mu \cdot Pr_{t2}$   $Yyy$  μονάδες, το οποίο ως συμβολίσουμε με

$$(Xxx, Yyy)$$

$$t_1 : \text{αγορά}(\mu, \mu \cdot Pr_{t1}) \text{πώληση}$$

$$t_2 : \text{πώληση}(\mu, \mu \cdot Pr_{t2}) \text{αγορά}$$

επομένως το κέρδος από αυτήν την συνδιαλλαγή είναι  $\mu \cdot (Pr_{t2} - Pr_{t1})$  και είναι σε  $Yyy$  μονάδες. Η τιμή  $Pr$  αναφέρεται στο σύμβολο στο οποίο έχει γίνει αναφορά δηλαδή στο  $XxxYyy$  · θα χρησιμοποιηθεί και ο συμβολισμός  $Pr^{(XxxYyy)}$ .

Για να αρχίσει αυτή η συνδιαλλαγή πρέπει να υπάρχει διάθεση  $\mu \cdot Pr_{t1}$   $Yyy$  μονάδων προς πώληση την χρονική στιγμή  $t_1$ . Αυτές οι μονάδες δεν παρέχονται από το περιεχόμενο του λογαριασμού μας που έχουμε συμβολίσει με  $Bal$ . Ανταυτού, πριν το βήμα  $t_1$  γίνεται η λήψη τους από το σύστημα και μετά το βήμα  $t_2$  επιστρέφονται, γι'αυτό και θεωρούμε την απώλεια  $-\mu \cdot Pr_{t1}$  ως μέρος του τελικού κέρδους. Αυτός ο τρόπος συνδιαλλαγής λέγεται συνδιαλλαγή με περιθώριο (*margin transaction*) και επιτρέπει στην θέση να έχει όγκο πολλές τάξεις μεγαλύτερο από το περιεχόμενο τους λογαριασμού μας.

Ανάποδα τώρα, το άνοιγμα και η ολοκλήρωση μιας θέσης πώλησης του συμβόλου  $XxxYyy$  κατά  $\mu$  μονάδες σημαίνει ότι την χρονική στιγμή  $t_1$  πουλάμε  $\mu$   $Xxx$  μονάδες αγοράζοντας  $\mu \cdot Pr_{t1}$   $Yyy$  μονάδες και την μεταγενέστερη χρονική στιγμή  $t_2$  αγοράζουμε  $\mu$   $Xxx$  μονάδες πουλώντας  $\mu \cdot Pr_{t2}$   $Yyy$  μονάδες, δηλαδή

$$(Xxx, Yyy)$$

$$t_1 : \text{πώληση}(\mu, \mu \cdot Pr_{t1}) \text{αγορά}$$

$$t_2 : \text{αγορά}(\mu, \mu \cdot Pr_{t2}) \text{πώληση}$$

επομένως το κέρδος από αυτήν την συνδιαλλαγή είναι  $\mu \cdot (Pr_{t1} - Pr_{t2})$  και είναι σε  $Yyy$  μονάδες επίσης.

Για να αρχίσει η πώληση πρέπει να υπάρχει διάθεση  $\mu$   $Xxx$  μονάδων προς πώληση την χρονική στιγμή  $t_1$ , οι οποίες λαμβάνονται πριν αυτήν την χρονική στιγμή και επιστρέφονται μετά την  $t_2$  ακριβώς όπως το επιτρέπει η συνδιαλλαγή περιθωρίου.

Μετά το άνοιγμα μιας θέσης, είτε πώλησης είτε αγοράς, μπορεί να συμβαίνουν απώλειες ανάλογα με την εξέλιξη που θα έχει η τιμή. Κάθε χρονική στιγμή ο

πάροχος της συνδιαλλαγής περιθωρίου υπολογίζει το κέρδος αν έκλεινε εκείνη την στιγμή η θέση, και αν αυτο το κέρδος φτάσει το  $-Bal$  η θέση κλείνει αυτόματα και συμβαίνει το λεγόμενο *margin call*, δηλαδή η κλήση που ενημερώνει ότι η θέση έκλεισε αυτόματα λόγω ανεπάρκειας του λογαριασμού στις τρέχουσες απώλειες. Γενικά, κάποια στιγμή, η τιμή μπορεί να έχει μια απότομη αλλαγή έτσι ώστε παρά τους υπολογισμούς και το αυτόματο κλείσιμο, η θέση μπορεί να οδηγήσει τον λογαριασμό σε απώλειες αρνητικότερες του  $-Bal$ , Γιαυτό, στην περίπτωση που ο πάροχος της συνδιαλλαγής περιθωρίου εξασφαλίζει ότι δεν μπορούν να συμβούν απώλειες για κάποιον χρήστη πέρα από το τρέχον  $Bal$ , πρέπει για κάθε θέση που ανοίγει να δεσμεύεται ένα μέρος του  $Bal$  ως σωσίβιο για τέτοιες απότομες περιπτώσεις. Το ακριβές ποσό που δεσμεύεται συνήθως καθορίζεται από μια παράμετρο μόχλευσης (*leverage*) που προσφέρεται στον χρήστη κατά την είσοδο του.

Συμβολίζοντας την απόφαση για άνοιγμα θέσης αγοράς ή πώλησης με  $B$  με τιμές  $+1$  ή  $-1$  αντίστοιχα, ο τυπος για τον υπολογισμό του κέρδους το οποίο συμβολίζουμε με  $R_{t_2}$  γράφεται

$$R_{t_2} = \mu \cdot B_{t_1} (Pr_{t_2} - Pr_{t_1}) \quad (4.1)$$

Αψηφώντας την διάκριση της τιμής σε *bid/ask* τιμές ο παραπάνω υπολογισμός του κέρδους είναι ακριβής εκτός του ότι οι  $Yyy$  πρέπει να μετατραπούν σε  $HOME$  μονάδες. Αυτό συμβαίνει ακριβώς μετά την χρονική στιγμή  $t_2$  με βάση την τιμή του συμβόλου  $YyyHOME$  το οποίο συμβολίζουμε με  $Pr_{t_2}^{(YyyHOME)}$

( $Yyy, HOME$ )

Αν το πρόσημο του  $R_{t_2}$  είναι θετικό τότε,

$t_2 ++$  : πώληση( $R_{t_2}, R_{t_2} \cdot Pr_{t_2}^{(YyyHOME)}$ ) αγορά

ειδάλλως αν είναι αρνητικό

$t_2 ++$  : αγορά( $R_{t_2}, R_{t_2} \cdot Pr_{t_2}^{(YyyHOME)}$ ) πώληση

Αυτό είναι το βήμα που επηρεάζει τον λογαριασμό  $Bal$  ο οποίος είναι σε  $HOME$  μονάδες. Δηλαδή, στο τέλος της συνδιαλλαγής σε κάθε περίπτωση αγοράς ή πώλησης και κέρδους ή απώλειας, το τελικό αποτέλεσμα είναι

$$Bal \leftarrow Bal + \mu \cdot B_{t_1} \cdot (Pr_{t_2} - Pr_{t_1}) \cdot Pr_{t_2}^{(YyyHOME)} \quad (4.2)$$

Αν το σύμβολο μεταξύ  $HOME$  και  $Yyy$  που χρησιμοποιείται είναι το  $HOME$   $Yyy$  αντί του  $YyyHOME$ , τα παραπάνω απλώς μετατρέπονται σε :

( $HOME, Yyy$ )

Αν το πρόσημο του  $R_{t_2}$  είναι θετικό τότε

$t_2 ++$  : αγορά( $R_{t_2}/Pr_{t_2}^{(HOMEYyy)}, R_{t_2}$ ) πώληση

ειδάλλως αν είναι αρνητικό  
 $t_2 + + : \text{πώληση}(R_{t_2}/Pr_{t_2}^{(HOMEYyy)}, R_{t_2})\text{αγορά}$

και η ανανέωση του λογαριασμού μετατρέπεται σε

$$Bal \leftarrow Bal + \mu \cdot B_{t_1}(Pr_{t_2} - Pr_{t_1})/Pr_{t_2}^{(HOMEYyy)} \quad (4.3)$$

Για παράδειγμα με το Ευρώ ως *HOME* το κέρδος απο συναλλαγές στο σύμβολο ΕυρωΔολλαρίου *EurUsd* είναι

$$\mu \cdot B_{t_1}(Pr_{t_2} - Pr_{t_1})/Pr_{t_2} \quad (4.4)$$

όπου ως συνήθως το *Pr* αναφέρεται στο σύμβολο υπό συναλλαγή και στην προκειμένη περίπτωση η διαίρεση πρέπει να γίνει με το  $Pr_{t_2}^{(EurUsd)}$  δηλαδή με το  $Pr_{t_2}$ .

### Συναλλαγές και κέρδη με *bid/ask* τιμές

Αν οι χρονικές στιγμές  $t_2$  και  $t_1$  απέχουν πολύ μεταξύ τους, τότε η διαφορά  $|Pr_{t_2} - Pr_{t_1}|$  αναμένεται να είναι πολύ μεγαλύτερη του *spread*, και επομένως η διάκριση σε *bid/ask* τιμές δεν είναι τόσο σημαντική.

Ωστόσο, το αυτόματο εμπόριο με μια υπολογιστική εμπορική στρατηγική στοχεύει στο Υψηλής Συχνότητας Εμπόριο αφενός έτσι ώστε να εκμεταλλευτεί το πλεονέκτημα του να επεξεργάζεται δεδομένα γρήγορα αφετέρου επειδή ενδεχομένως να τίθεται η υπόθεση ότι το δυναμικό πρότυπο της αγοράς το οποίο καλείται να προβλεφθεί παρουσιάζεται με μεγάλη συχνότητα. Επομένως υπάρχει ενδιαφέρον τα  $t_2$  και  $t_1$  να είναι σε μικρή απόσταση και έτσι η διαφορά  $|Pr_{t_2} - Pr_{t_1}|$  θα είναι σε συγκρίσιμο επίπεδο σε σχέση με το *spread*, και η διάκριση της τιμής σε *bid/ask* τιμές θα είναι κρίσιμη. Ο υπολογισμός του κέρδους λαμβάνοντας υπόψιν αυτήν την διάκριση γίνεται όπως στην προηγούμενη παράγραφο με την διαφορά ότι όταν συμβαίνει πώληση πρέπει να εμπλέκεται η *bid* τιμή και όταν συμβαίνει αγορά πρέπει να εμπλέκεται η *ask* τιμή.

Ξανακάνοντας τους υπολογισμούς για την θέση αγοράς ( $B=+1$ ) με βάση το παραπάνω σκεπτικό και το σχήμα 4.3, προκύπτει

$$Bal \leftarrow Bal + \mu \cdot (+1) \cdot (Pr_{t_2,bid} - Pr_{t_1,ask}) \cdot Pr_{t_2,kkk}^{(YyyHOME)} \quad (4.5)$$

όπου το *kkk* είναι *bid* ή *ask* αντίστοιχα με το αν η διαφορά είναι θετική ή αρνητική. Για παράδειγμα όταν είναι θετική σημαίνει ότι έχει συμβεί κέρδος σε *Yyy* μονάδες και γίνεται προσφορά του *Yyy* για να πάρουμε *HOME*. Τα ίδια συμβαίνουν και στον αντίστοιχο του τύπου (4.3) με το *HOMEYyy* και την

διαίρεση αντί για πολλαπλασιασμό με την διαφορά ότι το  $kkk$  είναι  $bid$  ή  $ask$  αντίστοιχα με το αν η διαφορά είναι αρνητική ή θετική.

Για την θέση πώλησης ( $B=-1$ ) τα  $bid$   $ask$  της διαφοράς αντιστρέφονται :

$$Bal \leftarrow Bal + \mu \cdot (-1) \cdot (Pr_{t2,ask} - Pr_{t1,bid}) \cdot Pr_{t2,kkk}^{(YyyHOMe)} \quad (4.6)$$

Για το  $kkk$  ισχύει ότι και για την θέση πώλησης.

### Εμπόριο στην αγορά συναλλάγματος

Υποθέτουμε μια ακολουθία  $(Pr)_t$  από τιμές η οποία θα είναι η βάση για το εμπόριο. Η ακολουθία διαφορών  $(r)_t$  ορίζεται με τον τύπο

$$r_t = Pr_t - Pr_{t-1} \quad (4.7)$$

Θεωρούμε την ακολουθία αποφάσεων  $(F)_t$  με κάθε δείγμα να παίρνει τιμή  $+1$  ή  $-1$  αντίστοιχα με την διάθεση μας να είναι σε θέση αγοράς ή θέση πώλησης αντίστοιχα. Η πιο γενική επιλογή να επιτρέπεται και η τιμή  $0$  απορρίπτεται επειδή αφενός δυσχεραίνει την ανάλυση και αφετέρου είναι συμφέρον να μην διακοπτόμαστε απ'την αγορά. Θεωρούμε επίσης την ακολουθία  $(B)_t$  με τιμές  $+1, 0$  ή  $-1$  που σημαίνουν εναλλαγή θέσης από πώληση σε αγορά, διατήρηση θέσης και εναλλαγή θέσης από αγορά σε πώληση αντίστοιχα. Ισχύει

$$B_t = \frac{1}{2} \cdot (F_t - F_{t-1}) \quad (4.8)$$

Ο στόχος στην συνέχεια θα είναι να χρησιμοποιήσουμε τους τύπους κέρδους της προηγούμενης ενότητας που ισχύουν για αυθαίρετες χρονικές στιγμές  $t_1, t_2$  για να εξάγουμε τύπο τοπικού κέρδους ανάμεσα σε δύο δείγματα των παραπάνω ακολουθιών. Την ακολουθία αυτού του τοπικού κέρδους θα την συμβολίζουμε με  $(R)_t$ .

Αρχικά χωρίς διάκριση σε  $bid/ask$  τιμές, έστω ότι μια θέση μένει σταθερή από την στιγμή  $t$  έως να αλλάξει στην στιγμή  $t+n$ . Τότε απ'τον τύπο (4.1) το συνολικό κέρδος είναι

$$\mu \cdot F_t \cdot (Pr_{t+n} - Pr_t)$$

ή

$$\mu \cdot (F_t \cdot (Pr_{t+1} - Pr_t) + F_{t+1} \cdot (Pr_{t+2} - Pr_{t+1}) + \dots + F_{t+n-1} \cdot (Pr_{t+n} - Pr_{t+n-1}))$$

δηλαδή

$$\mu \cdot \sum_{i=2}^n F_{t+i-1} \cdot r_{t+i}$$

και με βάση αυτό τον τύπο μπορούμε να θεωρήσουμε την τοπική συνεισφορά στο κέρδος

$$R_t = \mu \cdot F_{t-1} \cdot r_t \quad (4.9)$$

την οποία καλούμε ανταμοιβή της στιγμής  $t$ .

Τώρα με διάκριση σε *bid/ask* τιμές, έστω ότι μια θέση αγοράς ανοίγει και μένει σταθερή από την στιγμή  $t$  έως να αλλάξει στην στιγμή  $t+n$ . Τότε από τον τύπο (4.5) το συνολικό κέρδος σε  $Ygy$  μονάδες είναι

$$\mu \cdot F_t (Pr_{t+n,bid} - Pr_{t,ask})$$

ή

$$\begin{aligned} \mu \cdot (- (Pr_{t,ask} - Pr_{t,bid}) + F_t \cdot (Pr_{t+1,bid} - Pr_{t,bid}) + F_{t+1} \cdot (Pr_{t+2,bid} - Pr_{t+1,bid}) + \dots \\ + F_{t+n-1} \cdot (Pr_{t+n,bid} - Pr_{t+n-1,bid})) \end{aligned}$$

δηλαδή, συμβολίζοντας την ακολουθία των *spreads* με  $(s_t)$ ,

$$\mu \cdot (-s_t + \sum_{i=2}^n F_{t+i-1} \cdot r_{t+i,bid})$$

Ακολουθώντας το ίδιο σκεπτικό για μια θέση πώλησης που ανοίγει και μένει στάθερη από την στιγμή  $t_1$  έως να αλλάξει στην στιγμή  $t+n$  το κέρδος προκύπτει

$$\mu \cdot (-s_t + \sum_{i=2}^n F_{t+i-1} \cdot r_{t+i,ask})$$

Επομένως, επειδή κάθε ακολουθία αποφάσεων μπορεί να διαμεριστεί σε ακολουθίες όπως αυτές οι δύο που αναλύθηκαν παραπάνω, η τοπική ανταμοιβή  $R_t$  σε εμπόριο με *bid/ask* τιμές είναι

$$\mu \cdot F_{t-1} \cdot r_{t,kkk}$$

με το *kkk* να είναι *bid* ή *ask* αντίστοιχα με το πρόσημο του  $F_{t-1}$ , και με την επιπλέον διαφορά ότι οποτεδήποτε γίνεται εναλλαγή θέσης πρέπει να πληρώνεται το τρέχον *spread*. Δηλαδή είναι

$$R_t = \mu \cdot (F_{t-1} \cdot r_{t,kkk} - s_t \cdot |B_t|) \quad (4.10)$$

Πίνακας 4.3: Παράδειγμα ακολουθίας αποφάσεων και ανταμοιβών

ανταμοιβή		$-s_1$			$-s_4$		$-s_6$		$-s_8$
ανταμοιβή			$r_{2,bid}$	$r_{3,bid}$	$r_{4,bid}$	$-r_{5,ask}$	$-r_{6,ask}$	$r_{7,bid}$	$r_{8,bid}$
$B_t$		+1	0	0	-1	0	+1	0	-1
$F_t$		+1	+1	+1	-1	-1	+1	+1	-1
$t$	0	1	2	3	4	5	6	7	8

ή αλλιώς

$$R_t = \mu \cdot (F_{t-1} \cdot r_{t,kkk} - \frac{s_t}{2} |F_t - F_{t-1}|) \quad (4.11)$$

Μια προσέγγιση αυτού του τύπου γίνεται αν τεθεί το  $r_{t,kkk}$  στην μέση τιμή των  $bid$  και  $ask$  διαφορών, δηλαδή

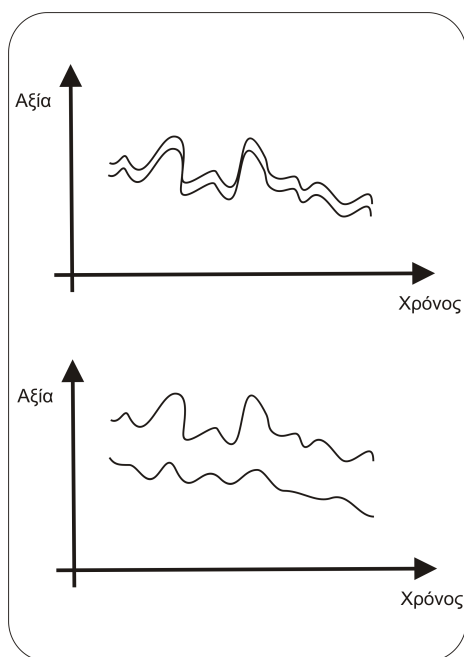
$$r_{t,kkk} = \frac{r_{t,bid} + r_{t,ask}}{2} \quad (4.12)$$

Οι παραπάνω τύποι είναι σημαντικοί επειδή θα αναπαριστούν την ανταμοιβή όταν στα επόμενα κεφάλαια η διαδικασία του εμπορίου τεθεί ως πρόβλημα Ενισχυτικής Μάθησης. Σημειώνεται εκ των υστέρων, ότι αυτές οι ανταμοιβές βασίστηκαν στο κέρδος σε  $Ygy$  μονάδες και όχι σε  $HOME$ , επειδή ο πολλαπλασιασμός με την τελική πληροφορία της στιγμής  $t_2$  στους τύπους (4.2) και (4.6) δεν μπορεί να διαμεριστεί σε τοπικές συνεισφορές. Ένα παράδειγμα εμπορίου φαίνεται στον πίνακα 4.3.

### *Movement/Spread* λόγος

Ανεξάρτητα από αν το εμπόριο γίνεται στην αγορά συναλλάγματος ή σε κάποια άλλη αγορά και από αν γίνεται με υπολογιστική μέθοδο ή όχι, η ακολουθία των  $bid/ask$  τιμών μπορεί να έχει κάποια χαρακτηριστικά που επηρεάζουν σημαντικά την δυνατότητα κέρδους. Για παράδειγμα στο πάνω διάγραμμα του σχήματος 4.4 βλέπουμε ότι η μέση τιμή των  $bid/ask$  τιμών εξελίσσεται με αρκετά δυναμικό τρόπο σε σχέση με το επίπεδο των  $spreads$ , κάτι που θέτει την ευκαιρία να συμβεί κερδοφόρο εμπόριο σε αντίθεση με το κάτω διάγραμμα όπου το επίπεδο των  $spreads$  είναι πολύ μεγάλο σε σχέση με τις αλλαγές στην μέση τιμή. Μια ποσοτικοποίηση των παραπάνω είναι ο *Movement/Spread* λόγος

$$\frac{M}{S}(\tau; Pr(1..T)) = \frac{1}{T - \tau} \sum_{t=1}^{T-\tau} \frac{|Pr_{t+\tau} - Pr_t|}{\bar{s}(t \dots t + \tau)} \quad (4.13)$$



Σχήμα 4.4: Αγορές με υψηλό και χαμηλό *Movement/Spread* λόγο

Η απόλυτη τιμή της διαφοράς κατά μήκος ενός παραθύρου μήκους  $\tau$  δια την μέση τιμή του *spread* σε αυτό το παράθυρο δείχνει το κατά πόσο θα μπορούσαν να υπάρξουν κερδοφόρες θέσεις μέσα σε αυτό. Αυτές οι τιμές προστίθενται σε όλη την αγορά μήκους  $T$  με το παράθυρο να ολισθαίνει, και παίρνοντας την μέση τιμή προκύπτει το τελικό αποτέλεσμα.

Ο *Movement/Spread* λόγος εκτός απ'την ίδια την αγορά, εξαρτάται και απ'τον τρόπο που γίνεται το εμπόριο, μέσω της παραμέτρου  $\tau$  που μπορεί να θεωρηθεί ως το μέσο μήκος των διαστημάτων που ορίζονται απ'τις διαδοχικές εναλλαγές θέσης, και μέσω της περιόδου με την οποία θεωρούμε τα δείγματα τιμών δηλαδή εν πολλοίς μέσω της συχνότητας εμπορίου.

## Κεφάλαιο 5

# Υπολογιστικές Εμπορικές Στρατηγικές

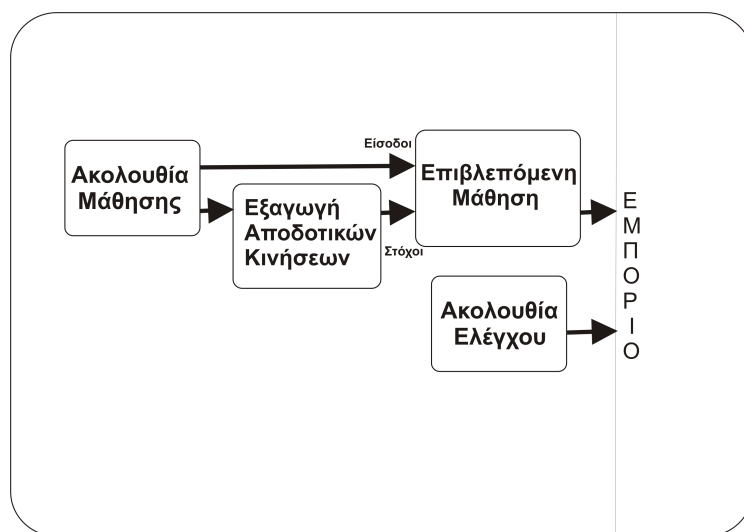
Σε αυτό το κεφάλαιο θα δούμε τρία γενικά σχήματα Μηχανικής Μάθησης στα οποία μπορούν να βασιστεί μια Υπολογιστική Εμπορική Στρατηγική και θα γίνουν αναφορές σε διάφορες σχετικές εργασίες.

Στο σχήμα 5.1 το εμπόριο γίνεται μέσω Επιβλεπόμενης Μάθησης ως εξής: Αρχικά η ακολουθία καταστάσεων εκπαίδευσης χρησιμοποιείται για να βρεθεί μια αποδοτική ακολουθία αποφάσεων, κάτι το οποίο μπορεί να γίνει μέσω μιας διαδικασίας βελτιστοποίησης. Στην συνέχεια και οι δυο ακολουθίες κατάστασης και αποφάσεων εισέρχονται σε ένα σύστημα Επιβλεπόμενης Μάθησης ως είσοδοι και στόχοι αντίστοιχα. Το αποτέλεσμα της μάθησης είναι η θέση ενός συνόλου παραμέτρων που θα χαρακτηρίζουν την Εμπορική Στρατηγική έτσι ώστε να μπορούν να λαμβάνονται αποδοτικές αποφάσεις σε άλλες άγνωστες ακολουθίες καταστάσεων ελέγχου.

Το σχήμα 5.2 δείχνει το εμπόριο μέσω Πρόβλεψης της Κατάστασης. Η δεδομένη ακολουθία καταστάσεων χρησιμοποιείται για να γίνει μάθηση ενός συστήματος έτσι ώστε να μπορεί να προβλέπει τις μελλοντικές καταστάσεις. Κατά την διάρκεια του ελέγχου το σύστημα προβλέπει τις μελλοντικές καταστάσεις για έναν μελλοντικό χρονικό ορίζοντα και με βάση αυτή την προβλεφθείσα ακολουθία βρίσκεται η ακολουθία αποφάσεων που θα εφαρμοστεί. Η εργασία να βρεθεί αυτή η ακολουθία αποφάσεων είναι ίδια με την ανακάλυψη της ακολουθίας στόχου στο σχήμα της Επιβλεπόμενης Μάθησης.

Στο σχήμα 5.3 φαίνεται το εμπόριο μέσω Άμεσης Ενισχυτικής Μάθησης το οποίο θα ακολουθήσουμε στα υπόλοιπα κεφάλαια. Εδώ η δεδομένη ακολουθία καταστάσεων χρησιμοποιείται ως ακολουθία καταστάσεων του περιβάλλοντος Ενισχυτικής Μάθησης και με βάση κάποιες συναρτήσεις Πολιτικής, ανταμοιβής και απόδοσης γίνεται θέση ενός συνόλου από παραμέτρους που χαρακτηρίζουν την Εμπορική Στρατηγική. Αυτό γίνεται με άμεση βελτιστοποίηση της Πολι-





Σχήμα 5.1: Εμπορική Στρατηγική με Επιβλεπόμενη Μάθηση

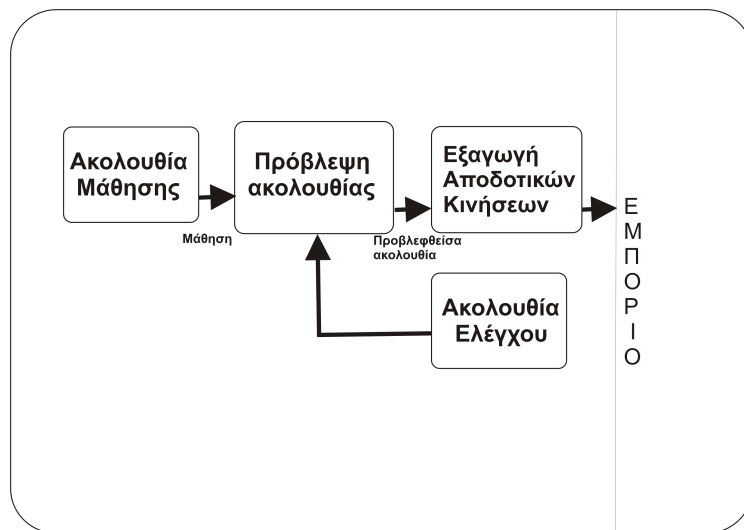
τικής ως προς τις ρυθμιζόμενες παραμέτρους όπως είδαμε στην παράγραφο 3.3.

Συγκρίνοντας αυτά τα τρία σχήματα, το σχήμα Άμεσης Ενισχυτικής Μάθησης μπορεί να θεωρηθεί ως μια απαλή (*soft*), με την έννοια του Soft Computing, έκδοση του σχήματος Επιβλεπόμενης Μάθησης αφού μια επιβλεπόμενη απόφαση πρέπει να θεωρηθεί αναγκαστικά τελείως σωστή είτε τελείως λάθος ανάλογα με τον σκληρό στόχο ενώ μια υπό ενίσχυση απόφαση παίρνει ενδιάμεσες τιμές μεταξύ σωστού και λάθους ανάλογα με την ανταμοιβή που προσφέρει. Επιπλέον σημειώνεται η αμεσότητα της Άμεσης Ενισχυτικής Μάθησης σε σχέση με το σχήμα Πρόβλεψης της Κατάστασης το οποίο έχει μοναδικό σκοπό να γίνει έμμεση πρόβλεψη της σωστής κίνησης. Η Άμεση Ενισχυτική Μάθηση αντιπαρέρχεται τις έμμεσες προβλέψεις προβλέποντας κατευθείαν την αποδοτική κίνηση.

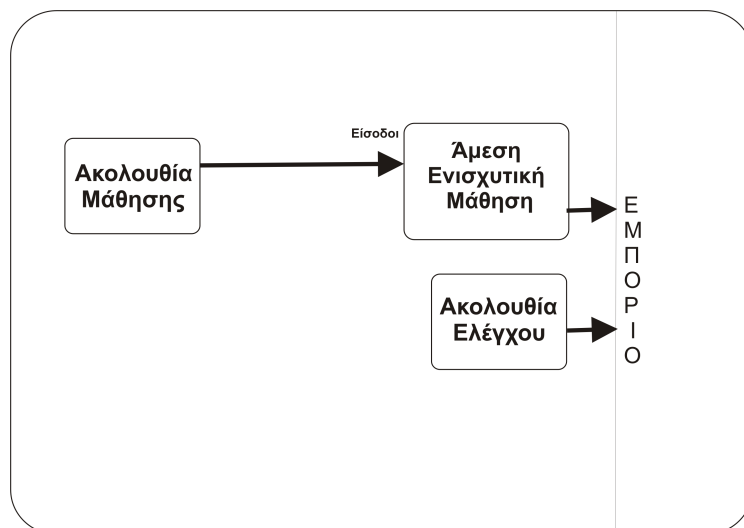
Για το σχήμα έμμεσης Ενισχυτικής Μάθησης, μιας που είναι περισσότερο γενικό και πολύπλοκο, χρησιμοποιούμε το βασικό σχήμα Ενισχυτικής Μάθησης του σχήματος 3.1.

Αν και τα παραπάνω σχήματα δεν εξαντλούν τους τρόπους που μπορεί να χρησιμοποιηθεί Μηχανική Μάθηση σε μια Εμπορική Στρατηγική, υπάρχουν πολλές εργασίες που εξετάζουν το εμπόριο με βάση αυτά. Μερικές διαφορές σε αυτές τις διάφορες προσεγγίσεις είναι οι μονάδες μάθησης που χρησιμοποιούνται, με παραδείγματα τα Κρυμμένα Μαρκοβιανά Μοντέλα και τα Νευρωνικά Δίκτυα, και οι διαδικασίες βελτιστοποίησης, με παραδείγματα την Εξέλιξη και την *gradient – descent* μέθοδο.

Παραδείγματα εργασιών για εμπόριο με βάση την Επιβλεπόμενη Μάθηση



Σχήμα 5.2: Εμπορική Στρατηγική με Πρόβλεψη Τιμών



Σχήμα 5.3: Εμπορική Στρατηγική με Άμεση Ενισχυτική Μάθηση

είναι η [35] και η [36] η οποία χρησιμοποιεί Support Vector Machines στην αγορά συναλλάγματος.

Στις [37][38] γίνεται πρόβλεψη τιμών της αγοράς μετοχών και στα [39][40][41] πρόβλεψη τιμών της αγοράς συναλλάγματος με Νευρωνικά Δίκτυα. Συγκεκριμένα, η [41] χρησιμοποιείται *RBF* νευρωνικό δίκτυο και στην [40] η βελτιστοποίηση γίνεται με Γενετικό Προγραμματισμό.

Στις [42][43][44] γίνεται εμπόριο με έμμεση Ενισχυτική Μάθηση και συγκεκριμένα στην [43] με *Q-learning* στην αγορά μετοχών, στην [42] με Γενετικό Προγραμματισμό στην αγορά μετοχών, και στην [44] με εξελικτικές μεθόδους στην αγορά συναλλάγματος.

Η Άμεση Ενισχυτική Μάθηση χρησιμοποιείται, για την αγορά μετοχών στις [5][9] και για την αγορά συναλλάγματος στις [4][6][10]. Στην [9], ειδικά, χρησιμοποιείται μια στοχαστική επέκταση της Αναδρομικής Ενισχυτικής Μάθησης.

Στις [7][2] γίνεται σύγκριση της έμμεσης και της άμεσης EM στην αγορά μετοχών. Στην [47] γίνεται σύγκριση ενισχυτικής μάθησης και συστημάτων πρόβλεψης τιμής στην αγορά μετοχών και στις [45][46] σύγκριση Εξελικτικών μεθόδων και Γενετικού Προγραμματισμού με Ενισχυτική Μάθηση στην αγορά συναλλάγματος. Τέλος ο έμμεσος και ο άμεσος τρόπος επίλυσης της Ενισχυτικής Μάθησης, με εκπρόσωπους την *Q-learning* και την *RRL* αντίστοιχα, συγκρίνονται στις [7][2][48] και στην [49].

Περνώντας εκτός χρήσης Νευρωνικών Δικτύων, η [51] είναι μια εργασία που χρησιμοποιεί Κρυμμένα Μαρκοβιανά Μοντέλα για πρόβλεψη τιμών, και εκτός Μηχανικής Μάθησης αλλά εντός Τεχνητής Νοημοσύνης η [50] χρησιμοποιεί έμπειρο σύστημα (*Rule – Based Expert System*) στην αγορά συναλλάγματος.

Στα επόμενα κεφάλαια χρησιμοποιούμε το σχήμα της άμεσης Ενισχυτικής Μάθησης και συγκεκριμένα τον αλγόριθμο *RRL* των *Moody et. al* [1][2][3]. Η εργασία του *Gold* [6] θεωρείται κεντρική στην αξιολόγηση της *RRL* όπου γίνεται ρύθμιση παραμέτρων μιας εμπορικής στρατηγικής που χρησιμοποιεί *online* μάθηση. Στο κεφάλαιο 7 θα παρουσιασθεί ρύθμιση παραμέτρων εμπορικής στρατηγικής με *batch RRL* μάθηση και τα εμπειρικά αποτελέσματα που προκύπτουν.

## Κεφάλαιο 6

# Αναδρομική Ενισχυτική Μάθηση

Θεωρούμε έναν διακριτό χρονικό ορίζοντα  $t=0 \dots N$ , και την τιμή  $Pr_t$  ενός οικονομικού αντικειμένου (*asset*). Οι διαφορές τις τιμές ορίζεται ως  $r_t = Pr_t - Pr_{t-1}$  για  $t=1 \dots T$ . Η απόφαση που μπορούμε να κάνουμε την χρονική στιγμή  $t$  συμβολίζεται με  $F_t$  και παίρνει τιμές  $+1$  ή  $-1$  οι οποίες ερμηνεύονται ως θέση αγοράς ή θέση πώλησης αντίστοιχα. Καθώς κυμαίνεται η τιμή, ο στόχος μας είναι να παίρνουμε τις σωστές αποφάσεις, ώστε να μεγιστοποιήσουμε μια ορισμένη συνάρτηση απόδοσης. Αυτή η συνάρτηση απόδοσης θα εξαρτάται από την ανταμοιβή κάθε χρονικής στιγμής του ορίζοντα η οποία είναι

$$R_t = \mu \cdot (F_{t-1} \cdot r_t - \delta_t \cdot |F_t - F_{t-1}|) \quad (6.1)$$

όπου  $\mu$  είναι η σταθερή ποσότητα του αντικειμένου που εμπορευόμαστε και  $\delta_t$  το κόστος συναλλαγής.

### Κατάσταση του περιβάλλοντος

Ανακαλώντας το κεφάλαιο 3, βλέπουμε ότι η παραπάνω διατύπωση ταιριάζει στο πλαίσιο της Ενισχυτικής Μάθησης. Θεωρώντας την κατάσταση του περιβάλλοντος να περιλαμβάνει τον τρέχον κόστος συναλλαγής, την προηγούμενη απόφαση του συστήματος  $F_{t-1}$  και ένα παράθυρο μήκους  $M$  με τις πιο πρόσφατες διαφορές

$$s_t = \{F_{t-1}, r_t, r_{t-1}, r_{t-2}, \dots, r_{t-M}, \delta_t\} \quad (6.2)$$

η συνάρτηση ανταμοιβής ορίζεται απ'τον παραπάνω τύπο του  $R_t$ . Το σύνολο των κινήσεων  $A$  είναι  $\{+1, -1\}$ , η συνάρτηση μετάβασης  $F_{trans}$  είναι τελείως άγνωστη, η συνάρτηση επιτρεπών πράξεων είναι σταθερή και ίση με το σύνολο

Α και η συνάρτηση αντίληψης  $F_{obs}$  μπορεί να θεωρηθεί ταυτοτική. Το πρόβλημα είναι να βρούμε μια Πολιτική  $F_{pol}$  η οποία έχει ικανοποιητική απόδοση.

Θα επιτεθούμε στο πρόβλημα με τον άμεσο τρόπο επίλυσης της Ενισχυτικής Μάθησης, δηλαδή αναγάγοντας το απευθείας σε μια βελτιστοποίηση στον χώρο των Πολιτικών. Το πρώτο βήμα είναι να χαρακτηρίσουμε την Πολιτική με ένα σύνολο απο παραμέτρους το οποίο αποκαλούμε  $\underline{\theta}$ . Δηλαδή είναι :

$$F_{pol} : \underline{\theta} \times S \rightarrow A$$

Στην συνέχεια πρέπει να ορίσουμε μια συνάρτηση μάθησης  $U(\underline{\theta}; (ss_1, ss_2, \dots, ss_N))$  η οποία αξιολογεί την λειτουργία των παραμέτρων, καθώς το σύστημα εκτίθεται στην δεδομένη ακολουθία από καταστάσεις. Το πρόβλημα της μάθησης τώρα είναι η βελτιστοποίηση αυτής της συνάρτησης η οποία μπορεί να γίνει με οποιαδήποτε μέθοδο βελτιστοποίησης.

Στις επόμενες παραγράφους επιλέγοντας συνάρτηση μάθησης  $U$ , συνάρτηση Πολιτικής  $F_{pol}$  και μέθοδο βελτιστοποίησης διαμορφώνεται ο αλγόριθμος της Αναδρομικής Ενισχυτικής Μάθησης (recurrent reinforcement learning - RRL).

## Συνάρτηση Πολιτικής

Η συνάρτηση Πολιτικής θα είναι του τύπου  $F_{pol} : \underline{\theta} \times S \rightarrow A$ , δηλαδή την χρονική στιγμή  $t$  θα παίρνει ως είσοδο το σύνολο των παραμέτρων  $\underline{\theta}$  και το σύνολο  $\{F_{t-1}, r_t, r_{t-1}, r_{t-2}, \dots, r_{t-M}, \delta_t\}$  και θα επιστρέφει την απόφαση  $F_t$ . Όπως φάνηκε απ' τον ορισμό της ανταμοιβής, η απόφαση  $F_{t-1}$  είναι κρίσιμη πληροφορία για να γίνει λήψη απόφασης την στιγμή  $t$ . Αυτό συμβαίνει επειδή η αλλαγή απόφασης επιφέρει κόστος συναλλαγής, γι' αυτό πρέπει να ξέρουμε σε τι απόφαση είμαστε στην χρονική στιγμή  $t - 1$  πριν πάρουμε μια απόφαση που τουλάχιστον αντιπαρέρχεται αυτό το κόστος. Επομένως χρειαζόμαστε μια αναδρομική συνάρτηση που θα εξαρτάται απ' τις παραμέτρους του συστήματος με τέτοιο τρόπο ώστε να μπορεί να γίνει αποδοτική μάθηση. Θα χρησιμοποιηθεί Αναδρομικό Νευρωνικό Δίκτυο με ένα ή δύο στρώματα υπολογισμού και συνάρτηση ενεργοποίησης  $\tanh$ . Θεωρώντας το παράθυρο διαφορών και τις παραμέτρους ως διανύσματα και ακολουθώντας τον συμβολισμό του κεφαλαίου 2, το νευρωνικό δίκτυο ενός επιπέδου παίρνει την μορφή

$$F_t = \text{sign}([r^T \quad F_{t-1} \quad 1] \hat{\theta}) \quad (6.3)$$

και με δύο επίπεδα :

$$F_t = \text{sign}([\tanh([r^T \quad F_{t-1} \quad 1] \cdot \hat{W}) \quad 1] \cdot \hat{\theta}) \quad (6.4)$$

Για την μάθηση, επειδή οι συναρτήσεις ενεργοποίησης χρειάζεται να είναι διαφορίσιμες, η  $\text{sign}$  αντικαθίσταται από την  $\tanh$ .

## Συνάρτηση Μάθησης

Υπάρχουν πολλές συναρτήσεις  $U(R_1, R_2, \dots, R_N)$  που μπορούν να αξιολογήσουν το σύστημα με βάση τις άμεσες ανταμοιβές. Πολλές από αυτές μπορούν να παραμετροποιηθούν ώστε να αντικατοπτρίζουν την διάθεση του συστήματος στο ρίσκο.

Μια απλή συνάρτηση είναι η  $\sum_{i=1}^N R_i$  η οποία αναπαριστά το τελικό κέρδος. Όμως έτσι αφήνουμε την διακύμανση των κερδών και όπως είδαμε στο κεφάλαιο 4, κατά το εμπόριο με περιθώριο πρέπει να κρατήσουμε την διακύμανση σε χαμηλά επίπεδα για να αποφύγουμε τυχόν *margin call*, το οποίο μας βγάζει εκτός παιχνιδιού πριν αρχίσει το σύστημα να πηγαίνει σε υψηλό τελικό κέρδος. Επομένως μια καλύτερη συνάρτηση μάθησης είναι ο λόγος Σαρπ

$$U = \frac{\text{mean}(R_1, R_2, \dots, R_N)}{\text{standardDeviation}(R_1, R_2, \dots, R_N)} \quad (6.5)$$

που επιβραβεύει μεγάλες ανταμοιβές ενώ ταυτόχρονα τιμωρεί μεγάλες διακυμάνσεις.

Πιο αναλυτικά, είναι

$$U = \frac{A_N}{K_N(B_N - A_N^2)^{1/2}} \quad (6.6)$$

όπου

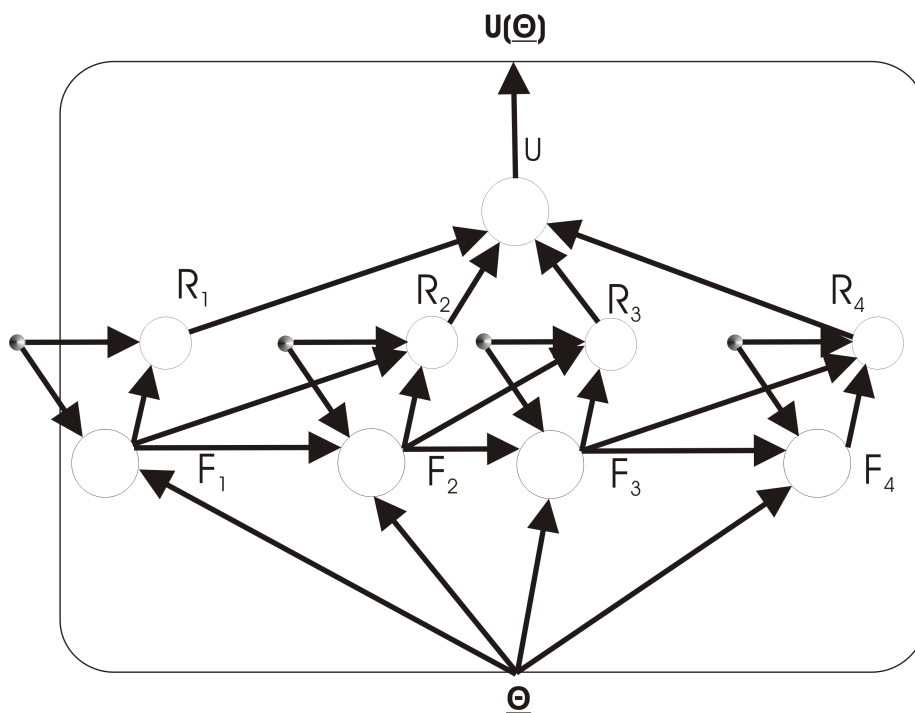
$$A_N = \frac{1}{N} \sum_{i=1}^N R_i \quad B_N = \frac{1}{N} \sum_{i=1}^N R_i^2 \quad K_N = \left(\frac{N}{N-1}\right)^{\frac{1}{2}} \quad (6.7)$$

## Μέθοδος Βελτιστοποίησης

Με τους παραπάνω ορισμούς, μπορεί να βρεθεί η αναλυτική κλίση της συνάρτησης μάθησης  $U$  ως προς κάθε ρυθμισίμη παράμετρο του συστήματος, με δεδομένη την ακολουθία από τις διαφορές των τιμών και τα κόστη συναλλαγών. Έτσι για την βελτιστοποίηση μπορεί να χρησιμοποιηθεί *gradientdescent* :

$$\underline{\theta}_{new} = \underline{\theta}_{old} + \rho \cdot \frac{dU}{d\theta} \quad (6.8)$$

Στην επόμενη παράγραφο θα δούμε πως υπολογίζεται η αναλυτική κλίση για τον *batch* τρόπο μάθησης, και στην μεθεπόμενη αφού οριστεί ο Διαφορικός Λόγος Σαρπ θα εφαρμόσουμε την ανάλογη στοχαστική *gradient descent* ώστε να συμβεί *online* μάθηση.



Σχήμα 6.1: Διάγραμμα *Batch* μάθησης

## 6.1 *Batch* μάθηση

Με τους συμβολισμούς της παραγράφου 3.3 η εξίσωση που περιγράφει την *batch* εκπαίδευση σε μια ακολουθία απο καταστάσεις είναι η

$$\underline{\theta}_{\kappa+1,1} = \underline{\theta}_{\kappa,1} + \rho \cdot G(\underline{\theta}_{\kappa,1}) \quad (6.9)$$

όπου  $G$  είναι η αναλυτική κλίση της  $U$ .

Στο σχήμα 4.1 φαίνεται αποδιπλωμένο το ANN που υλοποιεί την Πολιτική και γενικά όλες οι συσχετίσεις των μεταβλητών έτσι ώστε να φαίνεται η επιρροή από τις ρυθμίσιμες παραμέτρους στην συνάρτηση μάθησης  $U$ . Για να υπολογίσουμε την ζητούμενη κλίση θα πρέπει να κάνουμε οπισθοδιάδοση από το  $U$  έως το  $\underline{\theta}$  λαμβάνοντας υπόψιν κάθε διαφορετικό μονοπάτι που τα ενώνει.

Από το  $U$  έως τις εξόδους του ANN, η οπισθοδιάδοση δίνει

$$G(\underline{\theta}) = \frac{dU}{d\underline{\theta}} = \sum_{t=1}^N \frac{dU}{dR_t} \left\{ \frac{dR_t}{dF_t} \frac{dF_t}{d\underline{\theta}} + \frac{dR_t}{dF_{t-1}} \frac{dF_{t-1}}{d\underline{\theta}} \right\} \quad (6.10)$$

Από τον τύπο (6.1) που περιγράφει το κέρδος έχουμε

$$\frac{dR_t}{dF_t} = -\mu \cdot \delta_t \cdot \text{sign}(F_t - F_{t-1}), \quad \frac{dR_t}{dF_{t-1}} = -\mu \cdot (r_t + \delta \cdot \text{sign}(F_t - F_{t-1})) \quad (6.11)$$

Για να οπισθοχωρήσουμε απ' τις εξόδους του ANN έως τις ρυθμίσιμες παραμέτρους αρχικά χρησιμοποιούμε τον αναδρομικό τύπο

$$\frac{dF_t}{d\theta} = \frac{\partial F_t}{\partial \theta} + \frac{\partial F_t}{\partial F_{t-1}} \frac{dF_{t-1}}{d\theta} \quad (6.12)$$

Στην περίπτωση ANN με 2 επίπεδα, οι μερικοί παράγωγοι υπολογίζονται με τους τύπους της παραγράφου 2.2 ενώ με ένα επίπεδο βρίσκονται κατευθείαν

$$\frac{\partial F_t}{\partial \theta} = (1 - F_t^2) \cdot [r^T F_{t-1} \mathbf{1}]^T \quad (6.13)$$

Η τελευταία παράγωγος που χρειάζεται είναι η  $\frac{dU}{dR_t}$  η οποία βασιζόμενοι στο τύπο του λόγου Σαρπ είναι

$$\begin{aligned} \frac{dU}{dR_t} &= \frac{dU}{dA_N} \cdot \frac{dA_N}{dR_t} + \frac{dU}{dB_N} \cdot \frac{dB_N}{dR_t} = \\ &= \frac{1}{N} \cdot \frac{B_N - A_N \cdot R_t}{K_N \cdot (B_N - A_N^2)^{\frac{3}{2}}} \end{aligned} \quad (6.14)$$

Έτσι έχουμε όλες τις παραγώγους που χρειάζονται στον υπολογισμό της κλίσης της  $U$ .

Οι παράμετροι του αλγόριθμου είναι

- $M$  : οι είσοδοι διαφορών του ANN
- $N$  : ο αριθμός νευρώνων του κρυφού επιπέδου, αν χρησιμοποιηθεί ANN με 2 επίπεδα
- $\rho$  : ο ρυθμός μάθησης
- $n_e$  : ο αριθμός εποχών στις οποίες θα γίνει η μάθηση

Θέτοντας αυτές τις παραμέτρους και δίνοντας στο σύστημα μια ακολουθία από διαφορές τιμών και κόστη συναλλαγών, μήκους  $N$ , και ένα τυχαίο αρχικό σύνολο  $\theta_{init}$  από ρυθμίσιμες παραμέτρους, με την παραπάνω μέθοδο το σύστημα μπορεί να μάθει να κάνει αποδοτικές κινήσεις πάνω σε αυτά τα δεδομένα. Μερικά παραδείγματα απόδοσης του αλγόριθμου με τεχνητές τιμές παρουσιάζονται στην παράγραφο 4.3.



## 6.2 Online μάθηση

Με βάση την παραπάνω *batch* συνάρτηση  $U$  και ορίζοντας ως  $g$  την κλίση της τοπικής συνεισφοράς στην  $U$ , προκύπτει ο αντίστοιχος *online* αλγόριθμος

$$\underline{\theta}_{k,t+1} = \underline{\theta}_{k,t} + \rho \cdot g(ss_t, \underline{\theta}_{k,t}) \quad (6.15)$$

$$t = 1 \dots N, \quad \underline{\theta}_{k+1,1} = \underline{\theta}_{k,N+1}$$

Παίρνοντας μόνο έναν όρο από το άθροισμα  $G$ , το  $g$  μπορεί να οριστεί ως

$$g(ss_t, \underline{\theta}) = \frac{dU_t}{d\underline{\theta}_t} = \frac{dU_t}{dR_t} \left\{ \frac{dR_t}{dF_t} \frac{dF_t}{d\underline{\theta}_t} + \frac{dR_t}{dF_{t-1}} \frac{dF_{t-1}}{d\underline{\theta}_{t-1}} \right\} \quad (6.16)$$

και η αντίστοιχη της εξίσωσης (6.12) είναι η

$$\frac{dF_t}{d\underline{\theta}_t} = \frac{\partial F_t}{\partial \underline{\theta}_t} + \frac{\partial F_t}{\partial F_{t-1}} \frac{dF_{t-1}}{d\underline{\theta}_{t-1}} \quad (6.17)$$

Για να ολοκληρωθεί ο αλγόριθμος πρέπει να οριστεί η  $U_t$  ως τοπική συνεισφορά στο λόγο Σαρπ  $U$ .

### Διαφορικός λόγος Σαρπ

Ο ορισμός του λόγου Σαρπ που δόθηκε μπορεί να υπολογίζεται και με αναδρομή των  $A$  και  $B$

$$A_t = \left(\frac{1}{t}\right)R_t + \left(1 - \frac{1}{t}\right)A_{t-1} \quad B_t = \left(\frac{1}{t}\right)R_t^2 + \left(1 - \frac{1}{t}\right)B_{t-1} \quad (6.18)$$

όπου κάθε  $R_j, j = 1 \dots t$  συναθροίζεται με το ίδιο βάρος στα τελικά  $A_t, B_t$ . Μια γενικευμένη μορφή της παραπάνω αναδρομής παραμετροποιείται με τον αριθμό  $\eta \in [0, 1]$  έτσι ώστε να δίνονται διαφορετικά βάρη στα διάφορα  $R_j$ .

$$A_t = (\eta)R_t + (1 - \eta)A_{t-1} \quad B_t = (\eta)R_t^2 + (1 - \eta)B_{t-1} \quad (6.19)$$

Κάνοντας αναδρομικούς υπολογισμούς με βάση αυτούς τους τύπους που ξαναγράφονται ως

$$A_t = A_{t-1} + \eta \Delta A_t \quad B_t = B_{t-1} + \eta \Delta B_t \quad (6.20)$$

προκύπτει ο γενικευμένος λόγος Σαρπ

$$S_t = \frac{A_t}{K_\eta (B_t - A_t^2)^{1/2}} \quad (6.21)$$

με

$$K_\eta = \left(\frac{1 - \frac{\eta}{2}}{1 - \eta}\right)^{\frac{1}{2}}, \quad A_0 = 0, \quad B_0 = 0$$

Οι παραπάνω αναδρομές μπορούν να χρησιμοποιούνται για να γίνει πιο αποδοτική η *batch* μάθηση, ωστόσο για την *online* μάθηση χρειάζεται η τοπική συνεισφορά  $U_t$  η οποία μπορεί να βρεθεί προσεγγιστικά αναπτύσσοντας το  $S_t$  σε πρώτη τάξη ως προς το  $\eta$  στο σημείο  $\eta = 0$

$$S_t \approx S_{t-1} + \eta \cdot \left. \frac{dS_t}{d\eta} \right|_{\eta=0} + O(\eta^2) \quad (6.22)$$

και παίρνοντας τον πρώτης τάξης όρο. Ορίζοντας τον *διαφορικό λόγο*  $\Sigma$ ρπ ως

$$D_t = \frac{dS_t}{d\eta} = \frac{B_{t-1} \cdot \Delta A_t - \frac{1}{2} A_{t-1} \Delta B_t}{(B_{t-1} - A_{t-1}^2)^{\frac{3}{2}}} \quad (6.23)$$

η ζητούμενη παράγωγος στην εξίσωση (6.16) είναι

$$\frac{dU_t}{dR_t} \approx \eta \cdot \frac{dD_t}{dR_t} = \eta \cdot \frac{B_{t-1} - A_{t-1} \cdot R_t}{(B_{t-1} - A_{t-1}^2)^{\frac{3}{2}}} \quad (6.24)$$

### 6.3 Μάθηση με *bid/ask* τιμές

Κατά το εμπόριο μέσω ενός δημιουργού αγοράς (*market maker*) κάθε οικονομικό αντικείμενο συσχετίζεται με δύο τιμές, την τιμή προσφοράς και την τιμή ζήτησης. Όταν συμβαίνει εμπόριο υψηλής συχνότητας η διάκριση σε *bid/ask* τιμές είναι σημαντική επειδή η αλλαγή της τιμής ανάμεσα σε δύο αποφάσεις είναι συγκρίσιμη με το *spread*, την διαφορά δηλαδή μεταξύ *ask* και *bid* τιμών.

Με βάση τους τύπους ανταμοιβής της παραγράφου 4.3, για να εφαρμόσουμε την Αναδρομική Ενισχυτική Μάθηση με *bid/ask* τιμές μπορούμε να θέσουμε ως  $Pr_t$  του αντικειμένου την μέση τιμή των *bid* και *ask* τιμών

$$Pr_t = \frac{Pr_{t,ask} + Pr_{t,bid}}{2} \quad (6.25)$$

και ως κόστος συναλλαγής το μισό *spread*

$$\delta_t = \frac{Pr_{t,ask} - Pr_{t,bid}}{2} \quad (6.26)$$

Απ'τα πειραματικά αποτελέσματα φαίνεται ότι αυτή η προσέγγιση αποδίδει καλά αφού το κέρδος που μετρείται απ'τον τυπο (7.1) με τις παραπάνω αντικαταστάσεις έχει ασήμαντες διαφορές απ'τις μετρήσεις με την ακριβή εξίσωση (4.11).

## 6.4 Τεχνητές τιμές

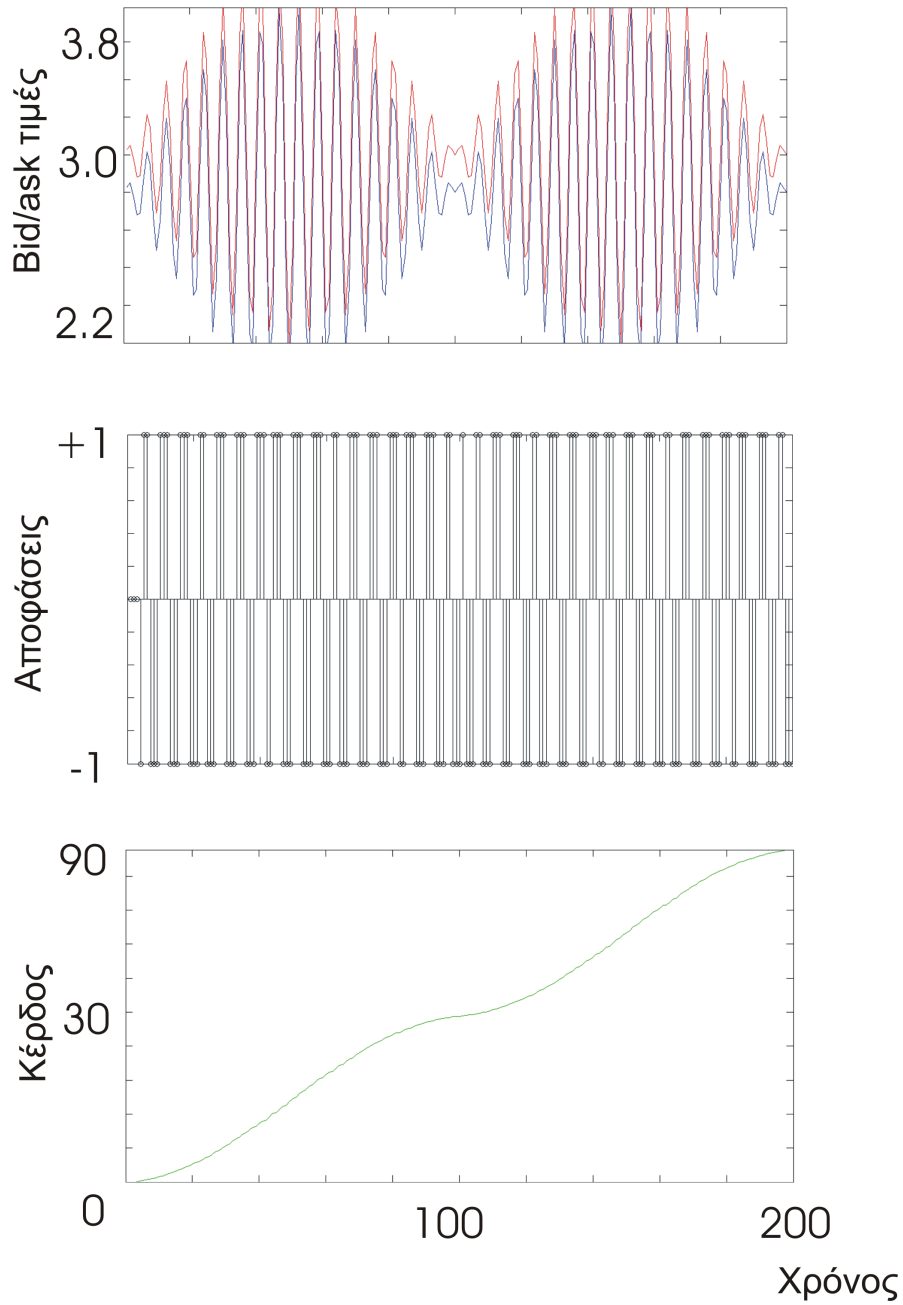
Ως μια πρώτη αξιολόγηση της Αναδρομικής Ενισχυτικής Μάθησης, σε αυτήν την παράγραφο θα δούμε παραδείγματα εφαρμογής της πάνω σε τεχνητές τιμές.

Στο σχήμα 6.2 βλέπουμε τιμές *bid/ask* διακροτήματος. Το σύστημα μαθαίνει με βάση αυτές τις τιμές και στην συνέχεια μπαίνει σε λειτουργία ελέγχου πάνω στις ίδιες τιμές. Οι αποφάσεις και το συσσωρευτικό κέρδος φαίνονται στο ίδιο σχήμα. Παρατηρούμε ότι όταν στην μέση της σειράς το *spread* γίνεται μεγάλο σχετικά με τις αλλαγές των τιμών, η συχνότητα εναλλαγής απόφασης ελαττώνεται. Στα σχήματα 6.3 και 6.4 αρχικά γίνεται μάθηση σε τιμές διακροτήματος και στην συνέχεια, στο σχήμα 6.3 συμβαίνει έλεγχος σε τιμές διακροτήματος με υπέρθεση θορύβου ενώ στο σχήμα 6.4 σε άσχετες θορυβώδεις τιμές. Το συμπέρασμα είναι ότι η μέθοδος δείχνει καλή ικανότητα γενίκευσης (*generalization*).

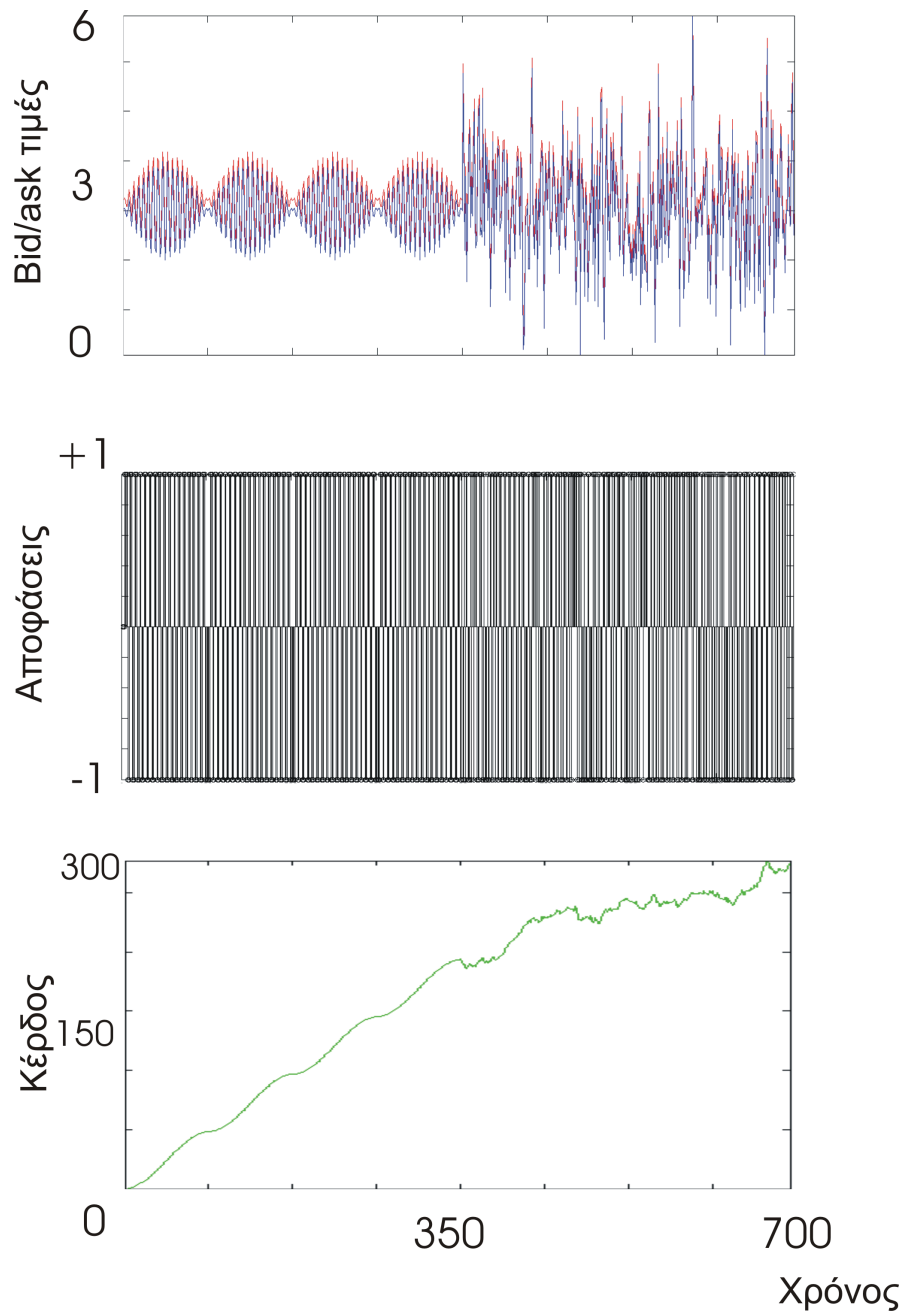
Για το τελευταίο παράδειγμα, η έξοδος  $p_t$  του παρακάτω στοχαστικού συστήματος τίθεται ως *ask* τιμές και ως *bid* τιμές είναι το ίδιο σήμα με αφαίρεση μικρού θετικού θορύβου. Οι αριθμοί  $\alpha$  και  $\beta$  μπορούν να πάρουν τις τιμές 0.1 και 0.5 αντίστοιχα όπως στο σχήμα 6.5.

$$\begin{pmatrix} p_{t+1} \\ h_{t+1} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} p_t \\ h_t \end{pmatrix} + \begin{pmatrix} noise1_t \\ noise2_t \end{pmatrix} \quad (6.27)$$

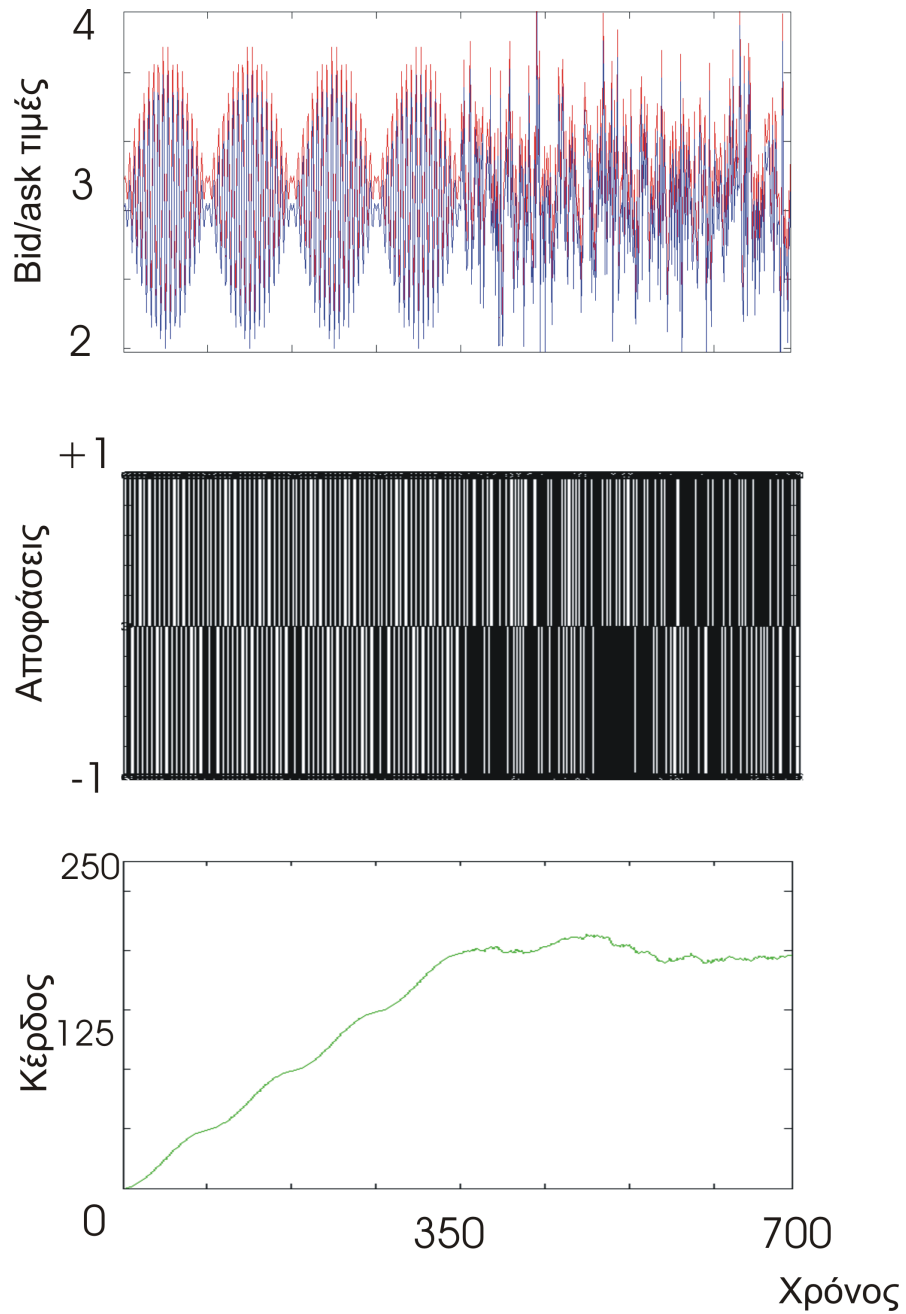
Σε αυτό το σχήμα, μάθηση γίνεται στα πρώτα 300 δείγματα και έλεγχος στα υπόλοιπα 100.



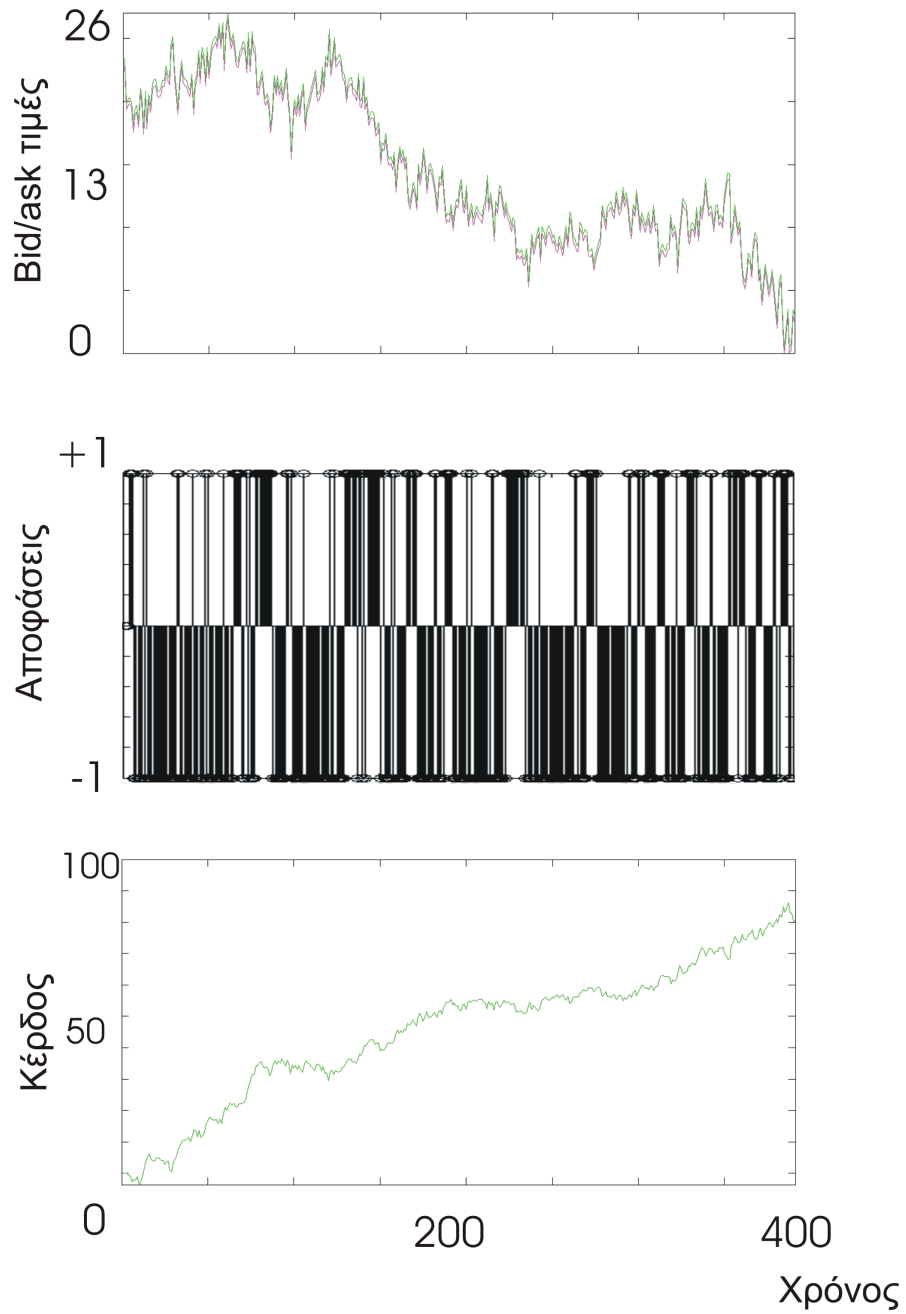
Σχήμα 6.2: Δείγμα αξιολόγησης της *RRL* μεθόδου πάνω σε *bid/ask* τιμές διακροτήματος. Το δεύτερο και το τρίτο διάγραμμα δείχνουν τις αποφάσεις και το συσσωρευμένο κέρδος που προκύπτουν. Έχει γίνει μάθηση σε όλη την ακολουθία



Σχήμα 6.3: Δείγμα αξιολόγησης της *RRL* μεθόδου πάνω σε *bid/ask* τιμές διακροτήματος με υπέρθεση θορύβου. Έχει γίνει μάθηση στο πρώτο μισό της ακολουθίας και έλεγχος στο υπόλοιπο μισό που είναι διακρότημα με υπέρθεση θορύβου



Σχήμα 6.4: Δείγμα αξιολόγησης της *RRL* μεθόδου πάνω σε *bid/ask* τιμές θορύβου. Έχει γίνει μάθηση στο πρώτο μισό της ακολουθίας και έλεγχος στο υπόλοιπο μισό που είναι θόρυβος



Σχήμα 6.5: Δείγμα αξιολόγησης της *RRL* μεθόδου πάνω σε *bid/ask* τιμές από έξοδο στοχαστικού συστήματος. Έχει γίνει μάθηση στο πρώτα 300 δείγματα της ακολουθίας και έλεγχος στο υπόλοιπα 100

## Κεφάλαιο 7

# Εμπόριο στην Αγορά Συναλλαγματος και Πειραματικά Αποτελέσματα

Σε αυτό το κεφάλαιο θα οριστεί και θα αξιολογηθεί μια εμπορική στρατηγική που περιλαμβάνει την *RRL* ως μονάδα μάθησης. Στην παράγραφο 7.1 εκτός από τον ορισμό της εμπορικής στρατηγικής θα παρουσιασθεί και το σύνολο δεδομένων *HFDF* – 96 της εταιρείας *Olsen*<sup>1</sup> πάνω στο οποίο θα γίνει η αξιολόγηση. Αυτό το σύνολο θα διαμερισθεί σε δεδομένα ρυθμίσης και δεδομένα ελέγχου. Στην παράγραφο 7.2 θα γίνει ρύθμιση της Εμπορικής Στρατηγικής με την έννοια ότι καθώς θα την έχουμε χαρακτηρίσει με ένα σύνολο  $\Sigma$  παραμέτρους θα ψάξουμε για το σημείο  $\Sigma^*$  που δίνει την καλύτερη απόδοση πάνω στα δεδομένα ρυθμίσης. Στην συνέχεια, στην παράγραφο 7.3 θα γίνει η αξιολόγηση της χαρακτηρισμένης με το σημείο  $\Sigma^*$  Εμπορικής Στρατηγικής μετρώντας την απόδοση της πάνω στα δεδομένα ελέγχου. Αυτή η απόδοση είναι η ενδιαφέρουσα μέτρηση επειδή τα δεδομένα ελέγχου είναι άγνωστα κατά το χτίσιμο της μεθόδου και έτσι αντιπροσωπεύουν τα μελλοντικά άγνωστα δεδομένα που καλείται να αντιμετωπίσει η στρατηγική στην πραγματικότητα. Τέλος, στην παράγραφο 7.4 θα παρατεθούν τα συμπεράσματα της αξιολόγησης και μερικές κατευθυντήριες γραμμές για σχετική μελλοντική έρευνα.

---

<sup>1</sup>[www.olsen.ch](http://www.olsen.ch)



## 7.1 Δεδομένα και Μέθοδος

Το πακέτο *HFDF* – 96 της *Olsen* που θα χρησιμοποιηθεί περιέχει *bid/ask* τιμές 25 σύμβολων για όλο το έτος 1996. Κάποια από αυτά τα δεδομένα φαίνονται στο σχήμα 7.1. Ο χρόνος που μεσολαβεί ανάμεσα σε δύο δείγματα σε αυτές τις χρονικές ακολουθίες είναι 30 λεπτά της ώρας.

Επειδή η λήψη μετρήσεων δεν ήταν πάντα δυνατή υπάρχουν μερικά μικρά χρονικά διαστήματα όπου οι τιμές παραμένουν σταθερές. Αυτά τα διαστήματα, καθώς και δείγματα *outliers* αφαιρούνται κατά το στάδιο προεπεξεργασίας των δεδομένων.

Ολοκληρώνοντας την λήψη δειγμάτων αξιολόγησης της *RRL* ως μεθόδου μάθησης, στο σχήμα 7.2 φαίνονται τα διαγράμματα τιμών, αποφάσεων και συσσωρευμένου κέρδους πάνω σε 500 δείγματα της αγοράς *DemFim*, όπου έχει γίνει μάθηση στα πρώτα 300 δείγματα και έλεγχος στα υπόλοιπα 200.

### Εμπορική Στρατηγική

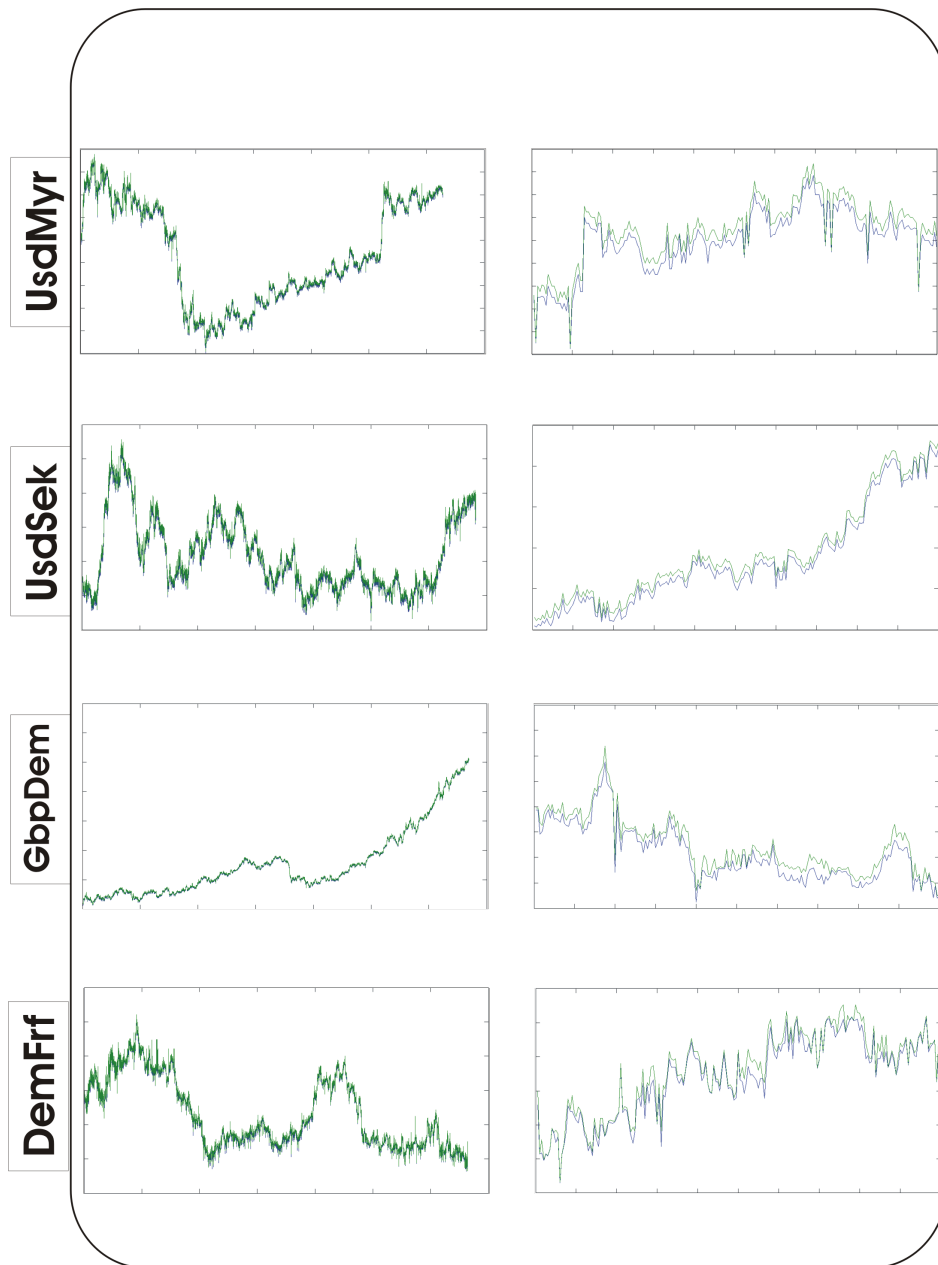
Γενικά μπορούν να οριστούν πολλές Εμπορικές Στρατηγικές βασισμένες στην *RRL*. Στη συνέχεια, στον αλγόριθμο 1, ορίζεται η Εμπορική Στρατηγική (ΕΣ) που θα χρησιμοποιήσουμε για πειραματικές μετρήσεις. Οι παράμετροι της είναι οι  $L_{train}, L_{trade}, M, N, \rho, n_e$  όπου οι τελευταίες τέσσερις είναι γνωστές και σχετίζονται με την διαδικασία *batch* μάθησης που θα χρησιμοποιηθεί και οι  $L_{train}, L_{trade}$  είναι το μήκος ακολουθίας μάθησης και το μήκος ακολουθίας εμπορίου αντίστοιχα.

Ο στόχος της επομένης παραγράφου είναι να βρεθούν οι τιμές των παραμέτρων που έχουν μέγιστη απόδοση πάνω στις αγορές ρύθμισης

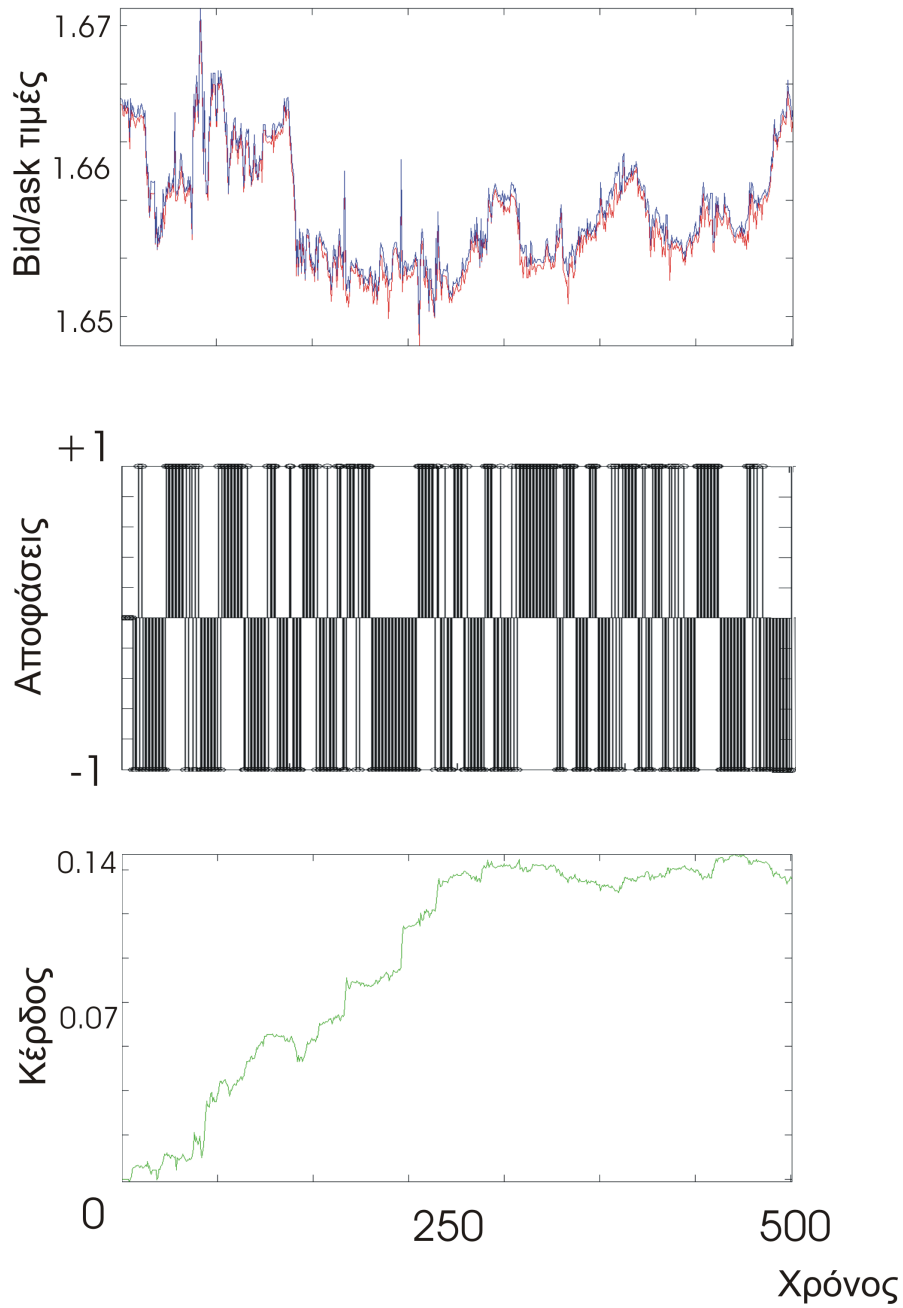
$$\Sigma^* = \operatorname{argmax}_{\Sigma \in \mathcal{S}} \{RRLTradingStrategy(\Sigma; tuningmarkets)\} \quad (7.1)$$

όπου  $\mathcal{S}$  είναι το σύνολο των πιθανών συνδυασμών για τις παραμέτρους.

Μερικές τελευταίες παρατηρήσεις πριν αρχίσει η ρύθμιση παραμέτρων είναι οι εξής: Οι διαφορές των τιμών που χρησιμοποιούνται από το σύνολο δεδομένων χρειάζεται να κανονικοποιηθούν έτσι ώστε αφενός να υπάρχει μια ομοιόμορφη εμβέλεια σε όλες τις διαφορετικές αγορές ώστε να είναι δυνατή μια ομοιόμορφη αντιμετώπιση τους και αφετέρου να έχουν συγκρίσιμες τιμές με την τιμή κατωφλίου του ANN η οποία έχει τεθεί ίση με 1. Επίσης, με βάση αρκετές αρχικές μετρήσεις εξοικείωσης με τα δεδομένα βρέθηκε ότι ο λόγος  $L_{trade}/L_{train}$  είναι αποδοτικό να πάρει την τιμή 1/4. Οπότε στην συνέχεια κρατάμε αυτήν την τιμή σταθερή για να διευκολύνουμε την διαδικασία αφαιρώντας την παράμετρο  $L_{trade}$ . Το  $L_{train}$  θα συμβολίζεται απλά ως  $L$ .



Σχήμα 7.1: Τιμές συναλλάγματος για τις αγορές *UsdMyr*, *UsdSek*, *GbpDem* και *DemFrfr*. Στα δεξιά το μήκος της ακολουθίας αντιστοιχεί σε όλο το έτος 1996 ενώ αριστερά φαίνεται μόνο ένα μικρό διάστημα 200 δειγμάτων δηλαδή 100 ωρών



Σχήμα 7.2: Δείγμα αξιολόγησης της *RRL* μεθόδου πάνω σε 500 δείγματα *bid/ask* τιμών της αγοράς *DemFim*. Έχει γίνει μάθηση στα πρώτα 300 και έλεγχος στα υπόλοιπα 200

---

**Algorithm 1** Trading Strategy Algorithm

---

$RRLTradingStrategy(L_{train}, L_{trade}, M, N, \rho, n_e; BidAskPrices)$

```
 $k = 0;$   
 $FinalAccumulativeProfit = 0;$   
repeat  
   $trainStart \leftarrow k + 1;$   
   $trainEnd \leftarrow k + L_{train};$   
   $trainingdata \leftarrow BidAskPrices(trainStart..trainEnd);$   
   $F_{pol} \leftarrow RRLearning(trainingdata, (M, N, \rho, n_e));$   
   $tradingdata \leftarrow BidAskPrices(trainEnd + 1..trainEnd + L_{trade});$   
   $faProfit \leftarrow Trading(tradingdata, F_{pol});$   
   $FinalAccumulativeProfit = FinalAccumulativeProfit + faProfit;$   
   $k \leftarrow k + L_{trade};$   
until  $k + L_{train} > length(BidAskPrices)$   
return  $FinalAccumulativeProfit$ 
```

---

## 7.2 Ρύθμιση Παραμέτρων

Οι 10 απίτις 25 αγορές του συνόλου  $HFDF - 96$  θα χρησιμοποιηθούν ως δεδομένα ρύθμισης. Η μονάδα μάθησης της ΕΣ χρησιμοποιεί ως μέτρο μάθησης τον λόγο Σάρπ ωστόσο αφού διαπιστώνεται ότι επιτυγχάνει να κρατάει την διακύμανση των κερδών σε ικανοποιητικά μικρά επίπεδα μπορούμε να θεωρήσουμε το τελικό συσσωρευτικό κέρδος ( $finalAccumulativeProfit - faP$ ) ως μέτρο απόδοσης. Στη συνέχεια του κεφαλαίου ως απόδοση της ΕΣ θα εννοείται αυτό το μέτρο. Επίσης, η σταθερή ποσότητα  $\mu$  των προηγούμενων κεφαλαίων που συμβολίζει την ποσότητα που εμπορεύεται, θα έχει πάντα την τιμή 1000 η οποία στην αγορά συναλλάγματος ονομάζεται 1 *microLot*. Πρακτικά, η ποσότητα εμπορίου μπορεί να τεθεί σε πολλές χιλιάδες κλιμακώνοντας την απόλυτη τιμή της απόδοσης η οποία ωστόσο πρέπει να ληφθεί υπόψιν ότι μπορεί να είναι αρνητική.

Το σύνολο παραμέτρων για ΕΣ με ANN ενός στρώματος είναι

$$\Sigma = \{L, M, n_e, \rho\}$$

Θέτουμε τις πιθανές τιμές για τις παραμέτρους αυτές ως

$$L \in \{1000, 1250, 1500, \dots, 2750, 3000\} \quad , \quad M \in \{2, 4, 6, 8, 10, 12, 14\}$$

$$\rho \in \{0.01, 0.1, 0.4, 0.7, 1\} \quad , \quad n_e \in \{25, 50, 75, 100\}$$

Με βάση τους παραπάνω μικρούς ορισμούς για τα πεδία των παραμέτρων ορίζεται το  $\mathcal{S}$  ώστε η εξίσωση (7.1) να είναι επιλύσιμη ωστόσο ακόμη και έτσι είναι σχεδόν αδύνατο να γίνει μια εξαντλητική αναζήτηση για την λύση της. Αντί αυτής της εξαντλητικής αναζήτησης θα χρησιμοποιηθεί η *greedy graduate student* τεχνική που παρουσιάζεται αμέσως παρακάτω.

### *Greedy graduate student*

Αρχικά επιλέγουμε μια τυχαία διάταξη των παραμέτρων και τυχαίες αρχικές τιμές για κάθε παράμετρο. Διατρέχουμε προς τα εμπρός την διάταξη κάνοντας *ceteris paribus* βελτιστοποίηση της τοπικής παραμέτρου, δηλαδή κρατώντας όλες τις υπόλοιπες παραμέτρους σταθερές. Όταν φτάσουμε στο τέλος της διάταξης, οπισθοχωρούμε με επιπλέον βελτιστοποιήσεις ως έλεγχοι ότι τα βέλτιστα σημεία δεν άλλαξαν παρά τις αλλαγές που ακολούθησαν.

Λίγο περισσότερο τυπικά, συμβολίζουμε το σύνολο παραμέτρων του βήματος  $i$  με  $\Sigma_i$ , την βελτιστοποίηση ως προς την παράμετρο  $x$  με  $CPOpt_x(\Sigma)$  και την ανανέωση του συνόλου  $\Sigma$  με καινούργια τιμή  $\chi$  της παραμέτρου  $x$  ως  $Upd_x(\chi, \Sigma)$ . Υποθέτοντας διάταξη  $(L, M, \rho, n_e)$  η *greedy graduate student* τεχνική φαίνεται παρακάτω ως αλγόριθμος 2.

Η όγδοη βελτιστοποίηση είναι ένα επιπλέον βήμα έτσι ώστε όλες οι παράμετροι να έχουν βελτιστοποιηθεί ίσο αριθμό φορών. Η παραπάνω μέθοδος εφαρμόζεται σε κάθε αγορά. Με βάση αυτά τα  $\Sigma^*$  που μαζεύονται, θα επιλεγεί κατά το τελικό στάδιο της ρύθμισης παραμέτρων το τελικό  $\Sigma^*$  που θα θεωρείται ο πιο αποδοτικός συνδυασμός παραμέτρων αν πρέπει να χρησιμοποιήσουμε μόνο έναν συνδυασμό πάνω σε όλες τις αγορές ρύθμισης. Τα αντίστοιχα συμβαίνουν και στην ΕΣ που χρησιμοποιεί ANN με 2 στρώματα όπου σε αυτήν την περίπτωση υπάρχουν πέντε παράμετροι  $L, M, N, \rho, n_e$ .

Χρησιμοποιώντας την παραπάνω τεχνική, στα σχήματα 7.3-7.6 φαίνεται η διαδικασία Ρύθμισης της ΕΣ με ANN 1 στρώματος στις αγορές *CadUsd*, *UsdDem*, *UsdBef* και *UsdZar*. Για κάθε συνδυασμό φαίνεται η μέση απόδοση και η διακύμανση της καθώς έχουν εφαρμοστεί 20 διαφορετικά διανύσματα αρχικών ρυθμίσιμων παραμέτρων του νευρωνικού δικτύου. Στο πρώτο από αυτά τα σχήματα, η διάταξη είναι  $(L, M, \rho, n_e)$  και οι αρχικές τιμές είναι  $(1000, 4, 1, 25)$ . Το  $\Sigma^*$  που προκύπτει για αυτήν την αγορά είναι το  $(1500, 4, 0.1, 75)$ .

Στον πίνακα 7.1 φαίνονται τα αποτελέσματα για τις 10 αγορές ρύθμισης με τους καλύτερους συνδυασμούς παραμέτρων που βρέθηκαν για κάθε αγορά. Κατά το τελικό στάδιο της ρύθμισης επιλέγουμε μόνο ένα από αυτά τα 6 διαφορετικά  $\Sigma^*$  που βρέθηκαν αθροίζοντας τις αποδόσεις τους σε δολάρια πάνω στις 10 αγορές ρύθμισης. Η μετατροπή σε δολάρια γίνεται σαν κανονικοποίηση των αποδόσεων ανάμεσα στις διαφορετικές αγορές. Σημειώνεται ότι τα δεδομέ-

---

**Algorithm 2** GreedyGraduateStudent technique for FineTuning, order  $(L, M, \rho, n_e)$  is supposed

---

GreedyGraduateStudent( $\Sigma_{init}$ )

```

 $\Sigma_1 \leftarrow \Sigma_{init}$ ;
 $L_{new} \leftarrow CPOpt_L(\Sigma_1)$ ;  $\Sigma_2 \leftarrow Upd_L(L_{new}, \Sigma_1)$ ;
 $M_{new} \leftarrow CPOpt_M(\Sigma_2)$ ;  $\Sigma_3 \leftarrow Upd_M(M_{new}, \Sigma_2)$ ;
 $\rho_{new} \leftarrow CPOpt_\rho(\Sigma_3)$ ;  $\Sigma_4 \leftarrow Upd_\rho(\rho_{new}, \Sigma_3)$ ;
 $n_{new} \leftarrow CPOpt_{n_e}(\Sigma_4)$ ;  $\Sigma_5 \leftarrow Upd_{n_e}(n_{new}, \Sigma_4)$ ;
if  $\Sigma_5 \neq \Sigma_3$  then
   $\rho_{new} \leftarrow CPOpt_\rho(\Sigma_5)$ ;  $\Sigma_6 \leftarrow Upd_\rho(\rho_{new}, \Sigma_5)$ 
else
   $\Sigma_6 \leftarrow \Sigma_5$ 
end if
if  $\Sigma_6 \neq \Sigma_2$  then
   $M_{new} \leftarrow CPOpt_M(\Sigma_6)$ ;  $\Sigma_7 \leftarrow Upd_M(M_{new}, \Sigma_6)$ ;
else
   $\Sigma_7 \leftarrow \Sigma_6$ 
end if
if  $\Sigma_7 \neq \Sigma_1$  then
   $L_{new} \leftarrow CPOpt_L(\Sigma_7)$ ;  $\Sigma_8 \leftarrow Upd_L(L_{new}, \Sigma_7)$ ;
else
   $\Sigma_8 \leftarrow \Sigma_7$ 
end if
if  $\Sigma_8 \neq \Sigma_4$  then
   $n_{new} = CPOpt_{n_e}(\Sigma_8)$ ;  $\Sigma^* = Upd_{n_e}(n_{new}, \Sigma_8)$ ;
else
   $\Sigma^* = \Sigma_8$ 
end if

```

---

Πίνακας 7.1: Ρυθμισμένες παράμετροι για ΕΣ με 1 στρώμα. Η τελευταία στήλη δίνει την απόδοση σε δολάρια ενώ η προ τελευταία σε μονάδες του δευτερεύοντος νομίσματος κάθε συμβόλου

σύμβολο	$M/S$	$\Sigma^*$	κέρδος ( $Yyy$ )	κέρδος (\$)
<i>AudUsd</i>	2.24	$(1500,6,0.1,50)=\Sigma_1^*$	192	192
<i>CadUsd</i>	1.45	$(1500,4,0.1,75)=\Sigma_2^*$	39	39
<i>DemFim</i>	0.97	$(2000,4,0.4,50)=\Sigma_3^*$	-316	-68
<i>DemJpy</i>	2.09	$(2000,8,0.1,50)=\Sigma_4^*$	17425	150
<i>GbpDem</i>	1.52	$(1500,6,0.1,50)$	-512.2	-325.1
<i>UsdBef</i>	3.13	$(2250,4,0.1,50)=\Sigma_5^*$	9860	317
<i>UsdChf</i>	2.57	$(1500,4,0.4,50)$	120	89
<i>UsdDem</i>	3.32	$(2000,4,0.4,50)$	341	217
<i>UsdDkk</i>	1.89	$(1500,4,0.1,75)$	1401	236
<i>UsdZar</i>	1.20	$(2000,4,0.4,75)=\Sigma_6^*$	-55	-12

να για να γίνει αυτή η μετατροπή εμπεριέχονται στο χρησιμοποιούμενο σύνολο δεδομένων αφού θεωρούμε ως *HOME* το αμερικάνικο δολλάριο και κάνουμε την μετατροπή σύμφωνα με τους τύπους του κεφαλαίου 3. Βλέποντας τα αθροίσματα στον πίνακα 7.2 ο συνδυασμός με το μεγαλύτερο άθροισμα είναι το  $\Sigma_3^*$  το οποίο είναι και το τελικό  $\Sigma^*$  και με αυτό χαρακτηρίζεται η καλύτερη δυνατή ΕΣ με 1 στρώμα. Στην παράγραφο 7.3 θα αξιολογηθεί αυτή η ΕΣ πάνω σε άγνωστες αγορές ελέγχου.

Στην συνέχεια παρουσιάζεται η αντίστοιχη διαδικασία ρύθμισης για εμπορική στρατηγική με νευρωνικό δίκτυο 2 στρωμάτων. Η επιπλέον παράμετρος  $N$  παίρνει τιμές από το σύνολο  $\{4, 8, 12, 16\}$ . Στα σχήματα 7.7-7.9 φαίνεται η ρύθμιση για 3 αγορές και στους πίνακες 7.3 και 7.4 οι ρυθμισμένες παράμετροι και το τελικό στάδιο της ρύθμισης αντίστοιχα.

### 7.3 Αξιολόγηση

Πριν αρχίσει η αξιολόγηση των ρυθμισμένων στρατηγικών που διαμορφώθηκαν στην προηγούμενη παράγραφο ας ορίσουμε μερικά άλλα συστήματα για σύγκριση :

- αφελές σύστημα 1 : χρησιμοποιεί ως Πολιτική ένα ANN όπως αυτά της μεθόδου *RRL*, με τυχαίες τιμές για τις ρυθμίσιμες παραμέτρους

Πίνακας 7.2: Τελικό στάδιο ρύθμισης παραμέτρων για ΕΣ με 1 στρώμα. Οι αποδόσεις έχουν κανονικοποιηθεί σε δολάρια και αθροίζονται

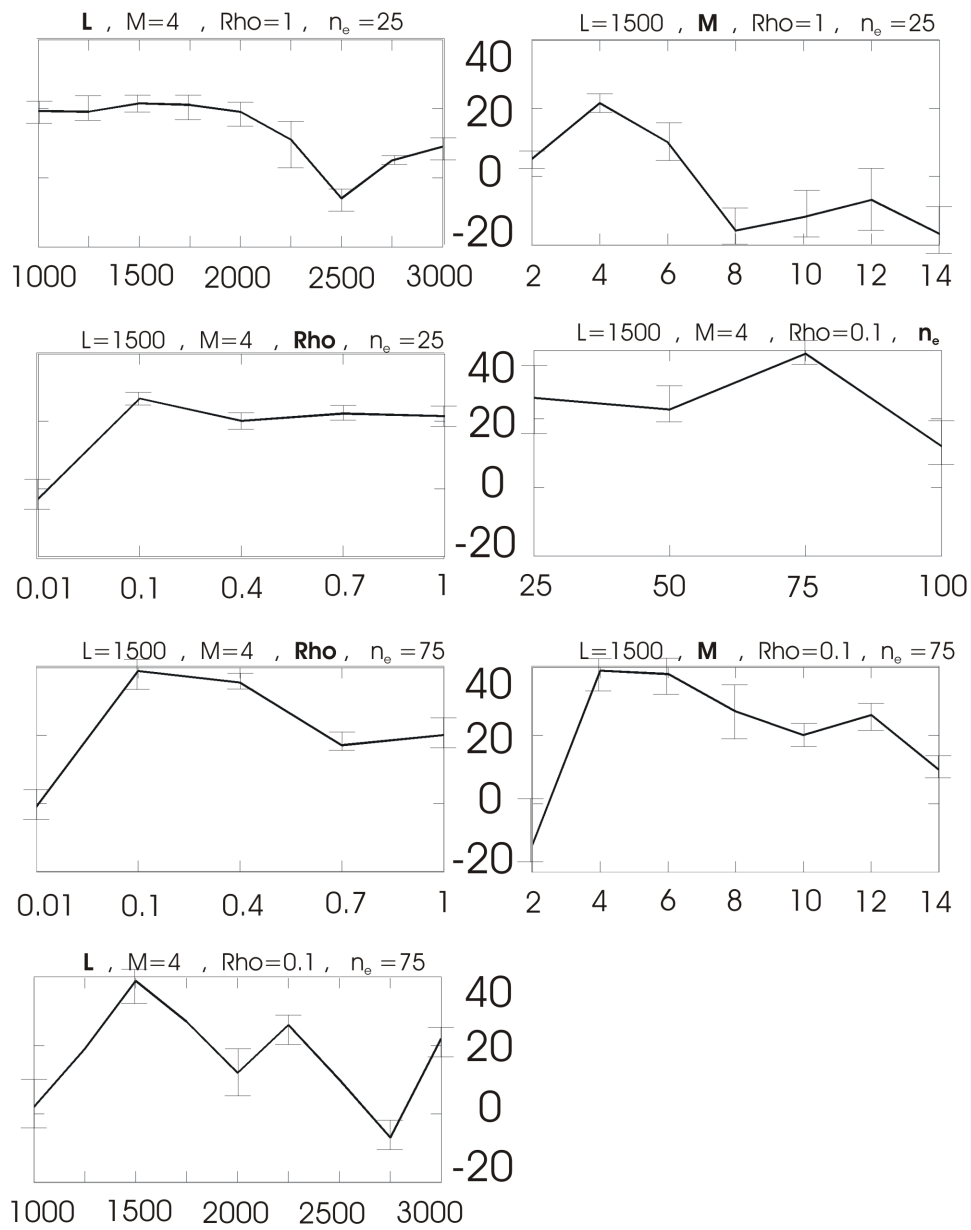
σύμβολο	$\Sigma_1^*$	$\Sigma_2^*$	$\Sigma_3^*$	$\Sigma_4^*$	$\Sigma_5^*$	$\Sigma_6^*$
<i>AudUsd</i>	192	188	125	-41	7	66
<i>CadUsd</i>	37	39	27	-15	18	16
<i>DemFim</i>	-85	-179	-68	-205	-124	-97
<i>DemJpy</i>	78	55	137	150	40	9
<i>GbpDem</i>	-170	-255	-203	-260	-235	-201
<i>UsdBef</i>	88	194	247	191	317	129
<i>UsdChf</i>	25	-2	89	27	-25	10
<i>UsdDem</i>	6	125	216	175	189	173
<i>UsdDkk</i>	152	236	205	107	-7	-12
<i>UsdZar</i>	-201	-21	-110	-253	-229	-12
άθροισμα	122	380	665	-124	-49	81

Πίνακας 7.3: Ρυθμισμένες παράμετροι για ΕΣ με 2 στρώματα. Η τελευταία στήλη δίνει την απόδοση σε δολάρια ενώ η προ τελευταία σε μονάδες του δευτερεύοντος νομίσματος κάθε συμβόλου

σύμβολο	$M/S$	$\Sigma^*$	κέρδος ( $Yyy$ )	κέρδος (\$)
<i>AudUsd</i>	2.24	$(2500,4,16,0.1,75)=\Sigma_1^*$	78	78
<i>CadUsd</i>	1.45	$(2000,6,12,0.4,75)=\Sigma_2^*$	39	39
<i>DemFim</i>	0.97	$(2500,4,16,0.1,75)$	-170.2	-102.1
<i>DemJpy</i>	2.09	$(2500,8,12,0.1,50)=\Sigma_3^*$	19891	171
<i>GbpDem</i>	1.52	$(1750,4,12,1,75)=\Sigma_4^*$	-72	-13
<i>UsdBef</i>	3.13	$(2000,6,12,0.4,50)=\Sigma_5^*$	12154	390
<i>UsdChf</i>	2.57	$(1000,8,16,0.4,50)=\Sigma_6^*$	150	111
<i>UsdDem</i>	3.32	$(1500,4,12,1,50)=\Sigma_7^*$	205	130
<i>UsdDkk</i>	1.89	$(1500,4,12,0.4,100)=\Sigma_8^*$	553	93
<i>UsdZar</i>	1.20	$(2000,6,12,0.4,75)$	-255.4	-56.8



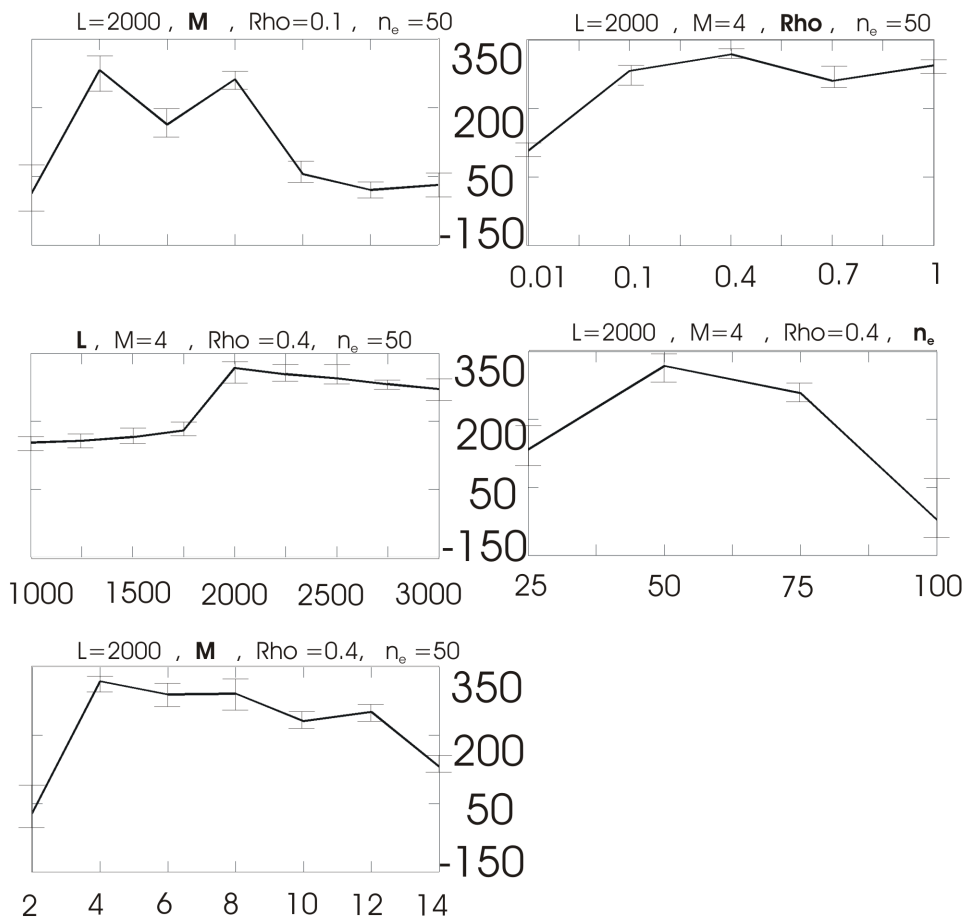
CadUsd



**L=1500, M=4, Rho=0.1, n\_e=75**

Σχήμα 7.3: Ρύθμιση παραμέτρων *CadUsd*, ANN με 1 στρώμα

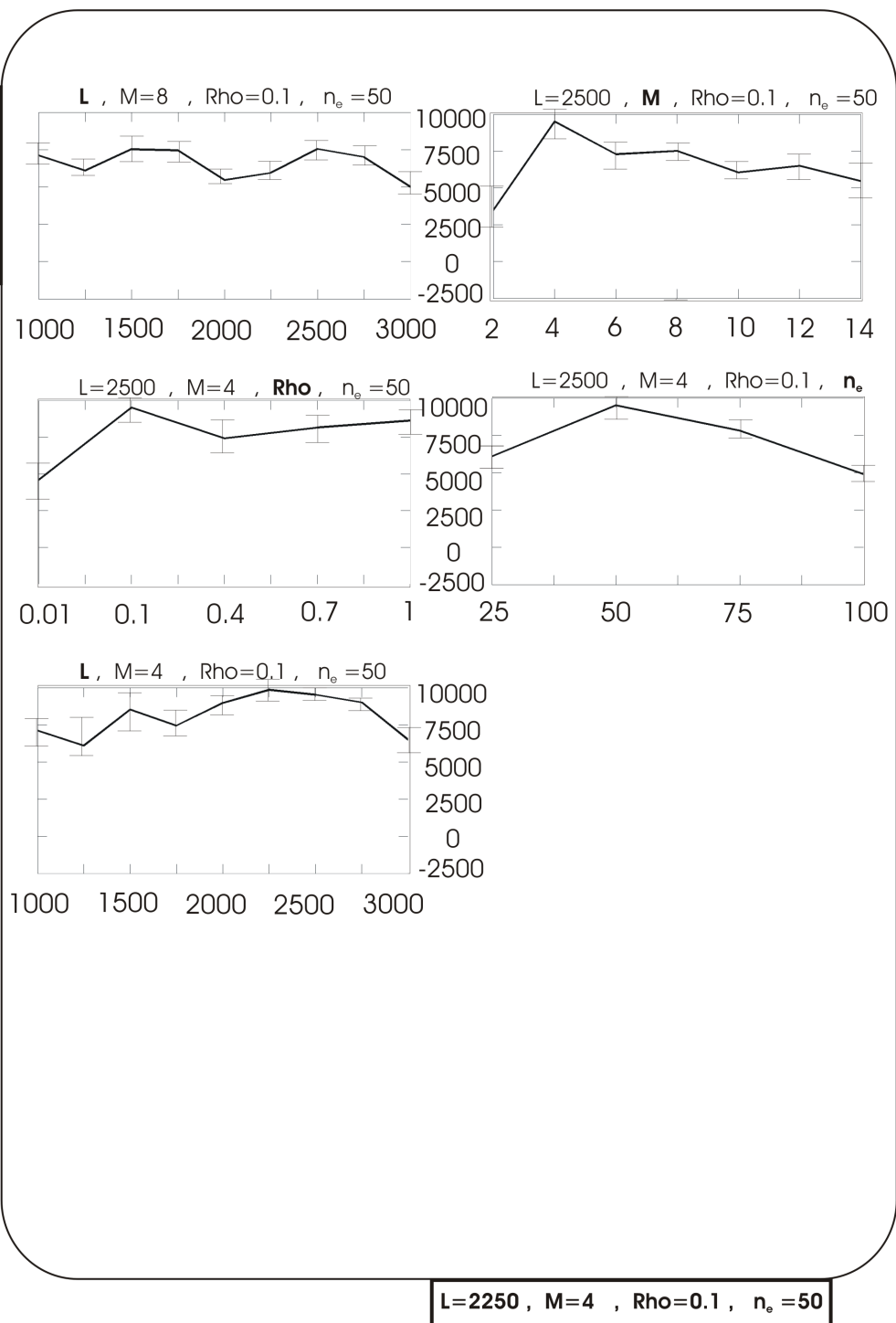
UsdDem



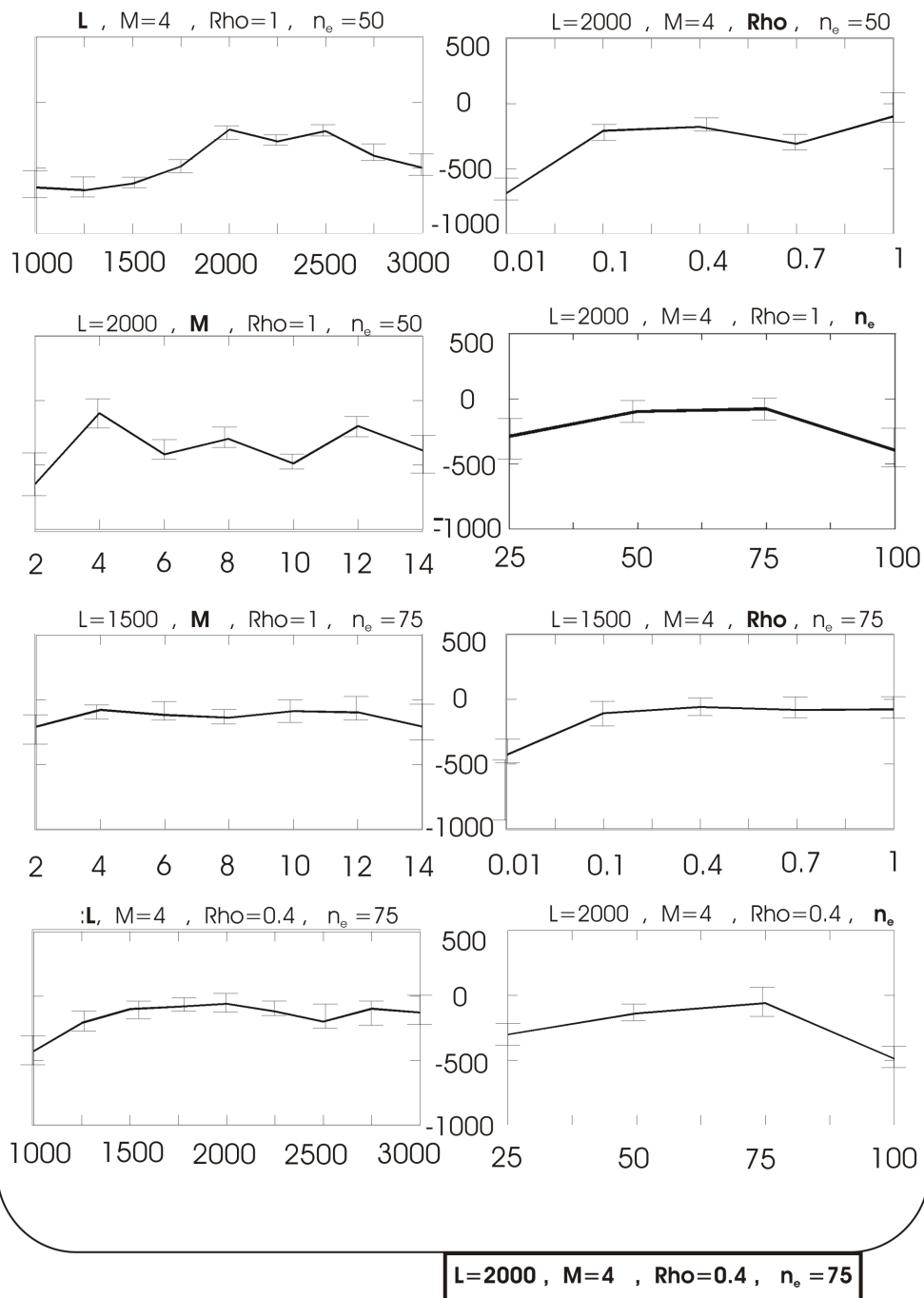
L=2000, M=4, Rho=0.4,  $n_e=50$

Σχήμα 7.4: Ρύθμιση παραμέτρων *UsdDem*, ANN με 1 στρώμα

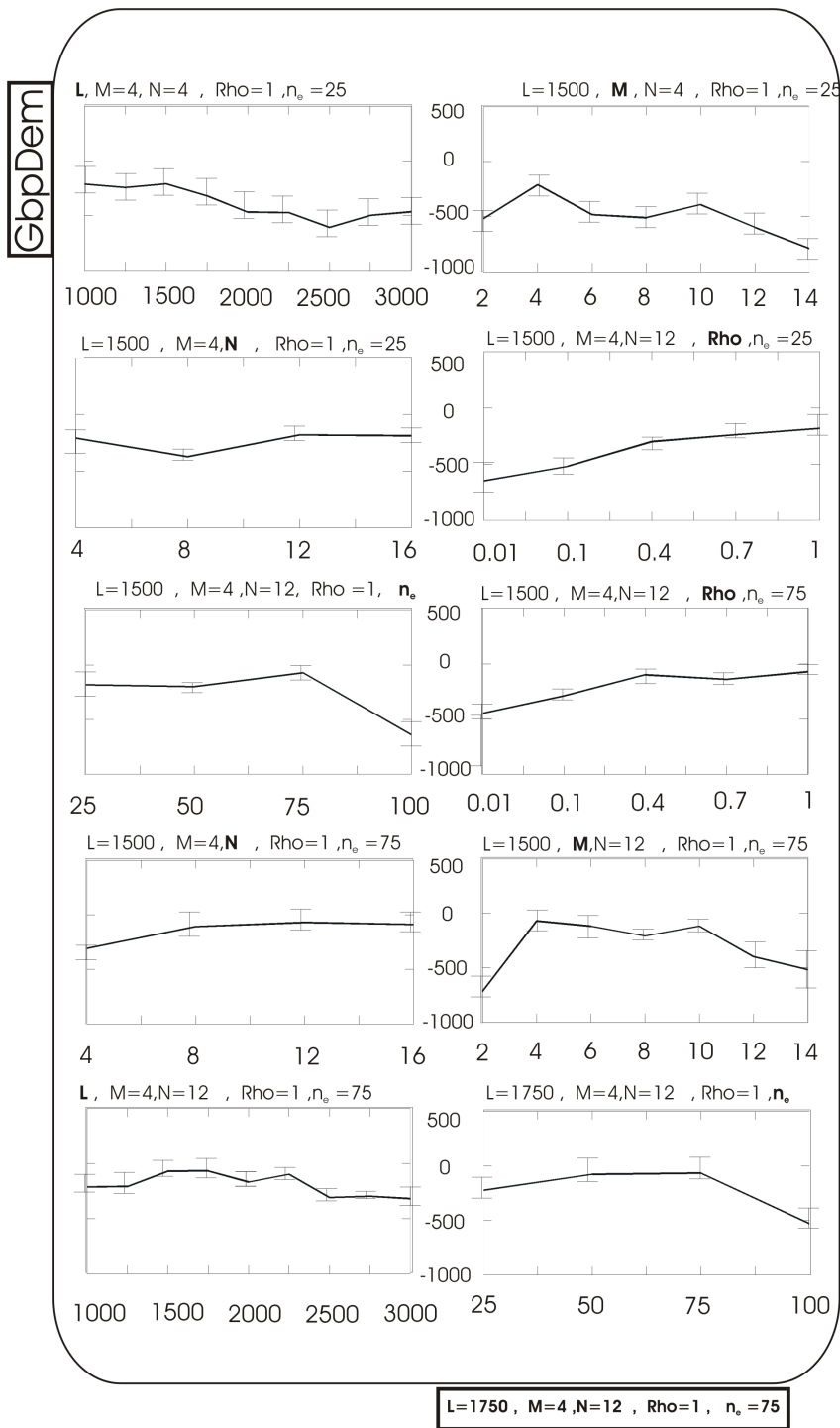
UsdBef



Σχήμα 7.5: Ρύθμιση παραμέτρων  $UsdBef$ , ANN με 1 στρώμα

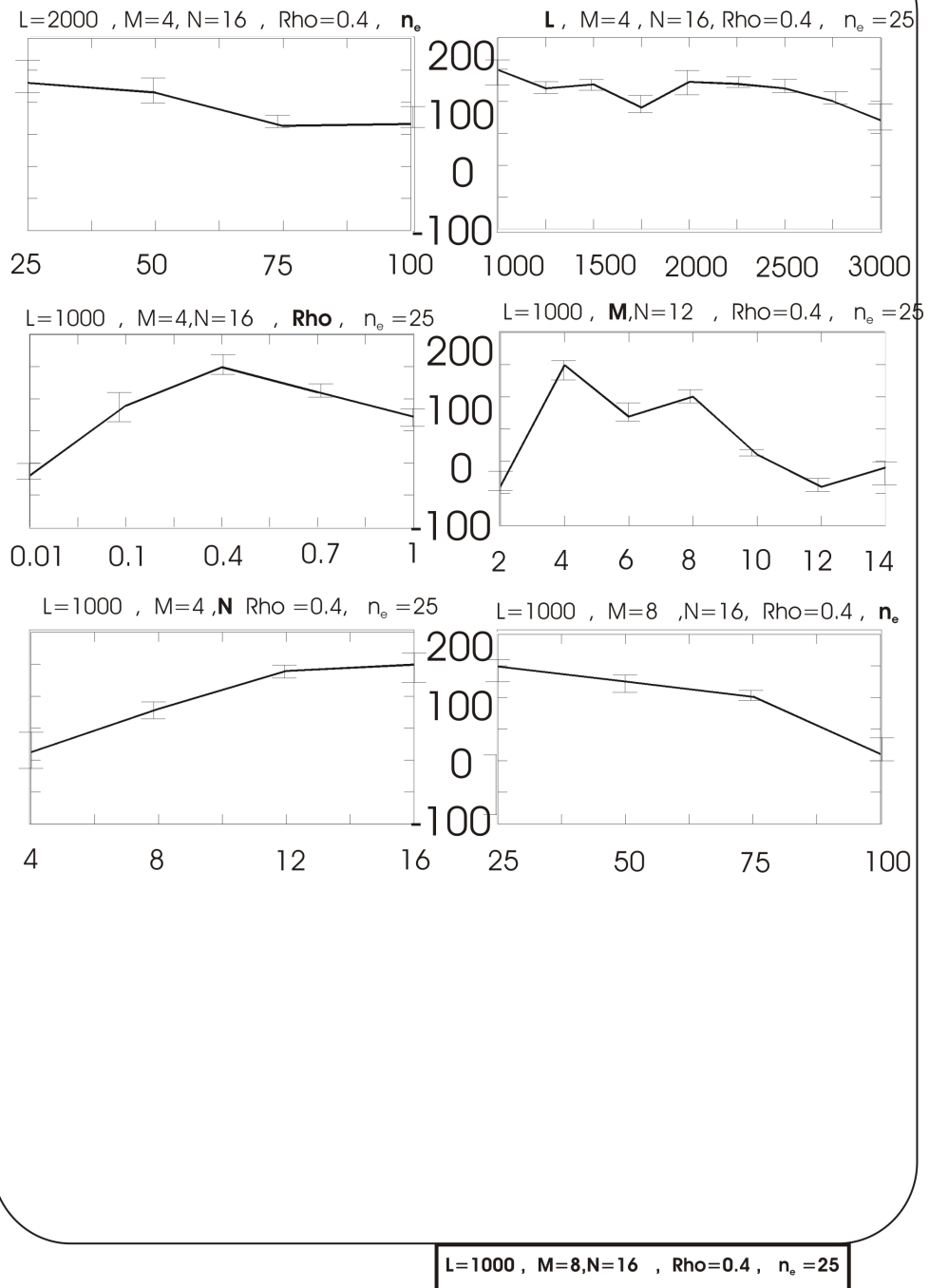


Σχήμα 7.6: Ρύθμιση παραμέτρων *UsdZar*, ANN με 1 στρώμα



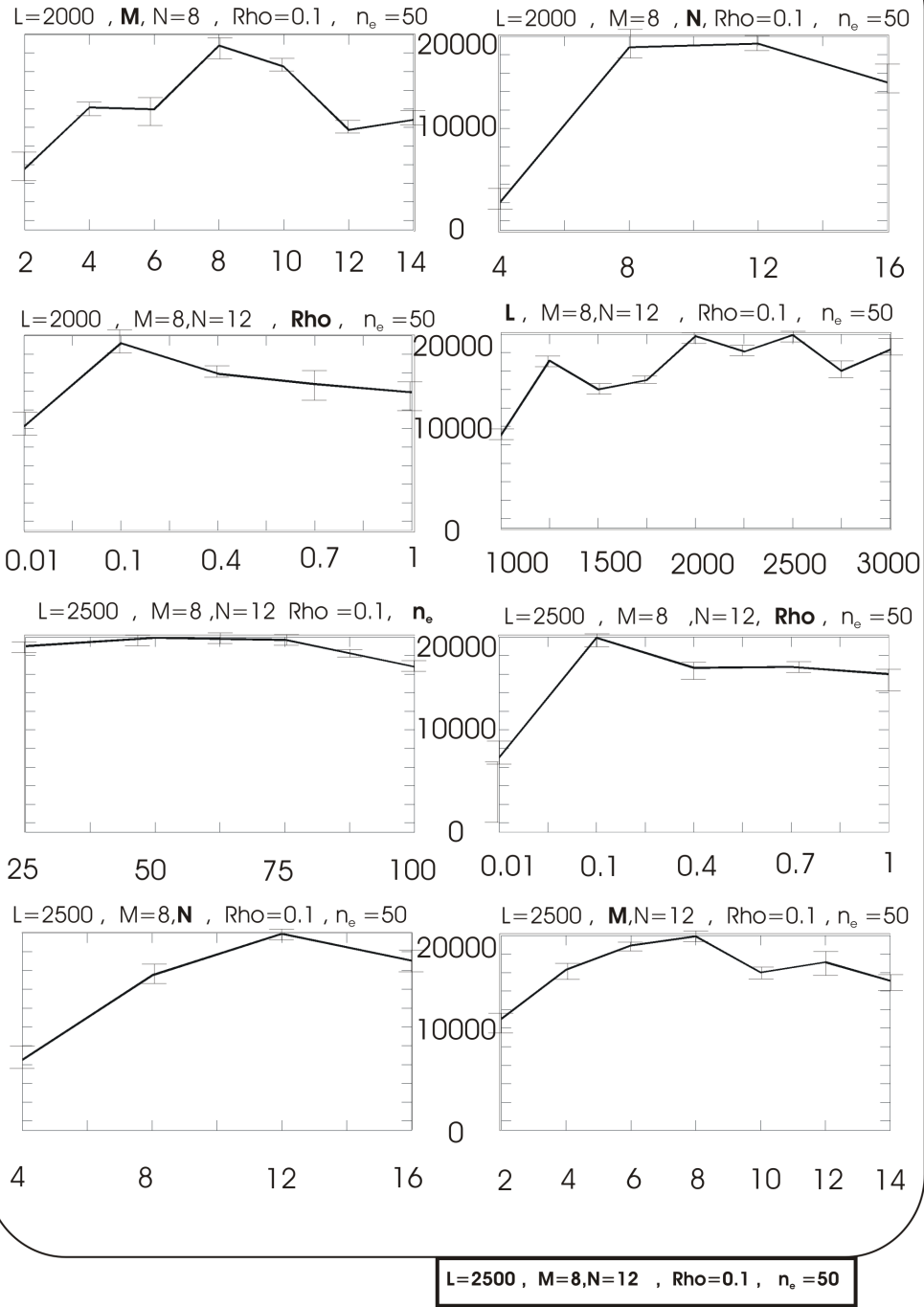
Σχήμα 7.7: Ρύθμιση παραμέτρων *GbpDem*, ANN με 2 στρώματα

UsdChf



Σχήμα 7.8: Ρύθμιση παραμέτρων  $UsdChf$ , ANN με 2 στρώματα

DemJpy



Σχήμα 7.9: Ρύθμιση παραμέτρων *DemJpy*, ANN με 2 στρώματα

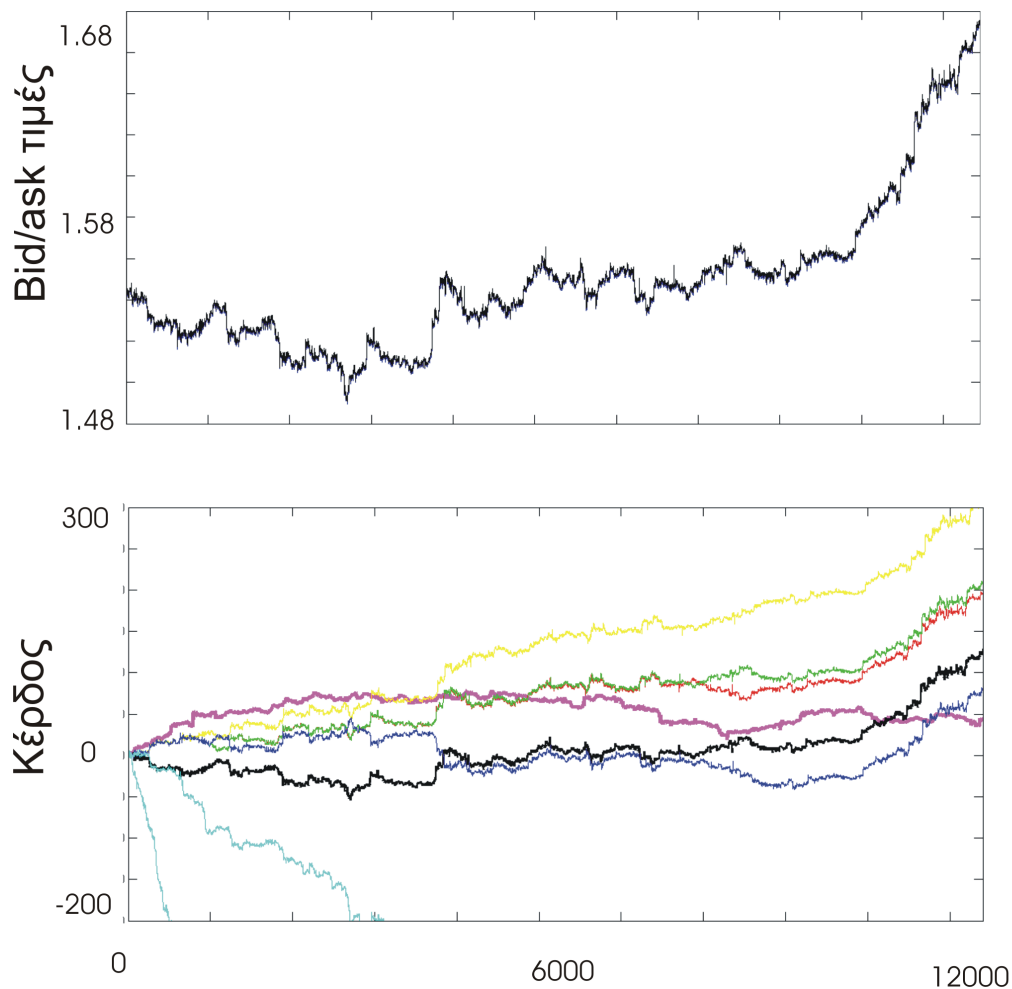
Πίνακας 7.4: Τελικό στάδιο ρύθμισης παραμέτρων για ΕΣ με 2 στρώματα. Οι αποδόσεις έχουν κανονικοποιηθεί σε δολάρια και αθροίζονται

σύμβολο	$\Sigma_1^*$	$\Sigma_2^*$	$\Sigma_3^*$	$\Sigma_4^*$	$\Sigma_5^*$	$\Sigma_6^*$	$\Sigma_7^*$	$\Sigma_8^*$
<i>AudUsd</i>	78	-45	47	40	65	7	27	-10
<i>CadUsd</i>	31	39	-4	26	25	20	19	-25
<i>DemFim</i>	-102	-185	-203	-121	-154	-279	-181	-206
<i>DemJpy</i>	106	72	171	107	125	81	49	77
<i>GbpDem</i>	-121	-145	-47	-13	-80	-44	-142	-305
<i>UsdBef</i>	26	56	322	121	390	107	20	86
<i>UsdChf</i>	17	97	80	102	20	111	-7	-15
<i>UsdDem</i>	7	61	108	27	97	-5	130	21
<i>UsdDkk</i>	49	71	56	50	67	79	25	93
<i>UsdZar</i>	-81	-56	-129	-105	-81	-121	-157	-69
άθροισμα	10	-35	401	234	474	-44	-217	-353

- N-αφελές σύστημα 2 : η συνολική ακολουθία τιμών διαμερίζεται σε διαστήματα μήκους  $N$  σε κάθε ένα από τα οποία η θέση μένει σταθερή και αποφασίζεται τυχαία
- αφελές σύστημα 3 : η συνολική ακολουθία τιμών διαμερίζεται σε διαστήματα μήκους  $L_{trade}$  σε κάθε ένα από τα οποία η θέση μένει σταθερή και είναι αγορά ή πώληση αντίστοιχα με το αν η τιμή αυξάνεται ή μειώνεται κατά μήκος του αμέσως προηγούμενου διαστήματος μήκους  $L_{train}$  .
- N-μαντείο : η συνολική ακολουθία τιμών διαμερίζεται σε διαστήματα μήκους  $\frac{datalength}{N}$  σε κάθε ένα από τα οποία η θέση μένει σταθερή και είναι αγορά ή πώληση αντίστοιχα με το αν η τιμή αυξάνεται ή μειώνεται κατά μήκος του ίδιου διαστήματος. Δηλαδή ένα N-μαντείο παίρνει σωστή απόφαση κάθε  $\frac{datalength}{N}$  δείγματα τιμών. Με *datalength* συμβολίζεται το μήκος όλης της ακολουθίας στην οποία καλείται το μαντείο να λειτουργήσει.

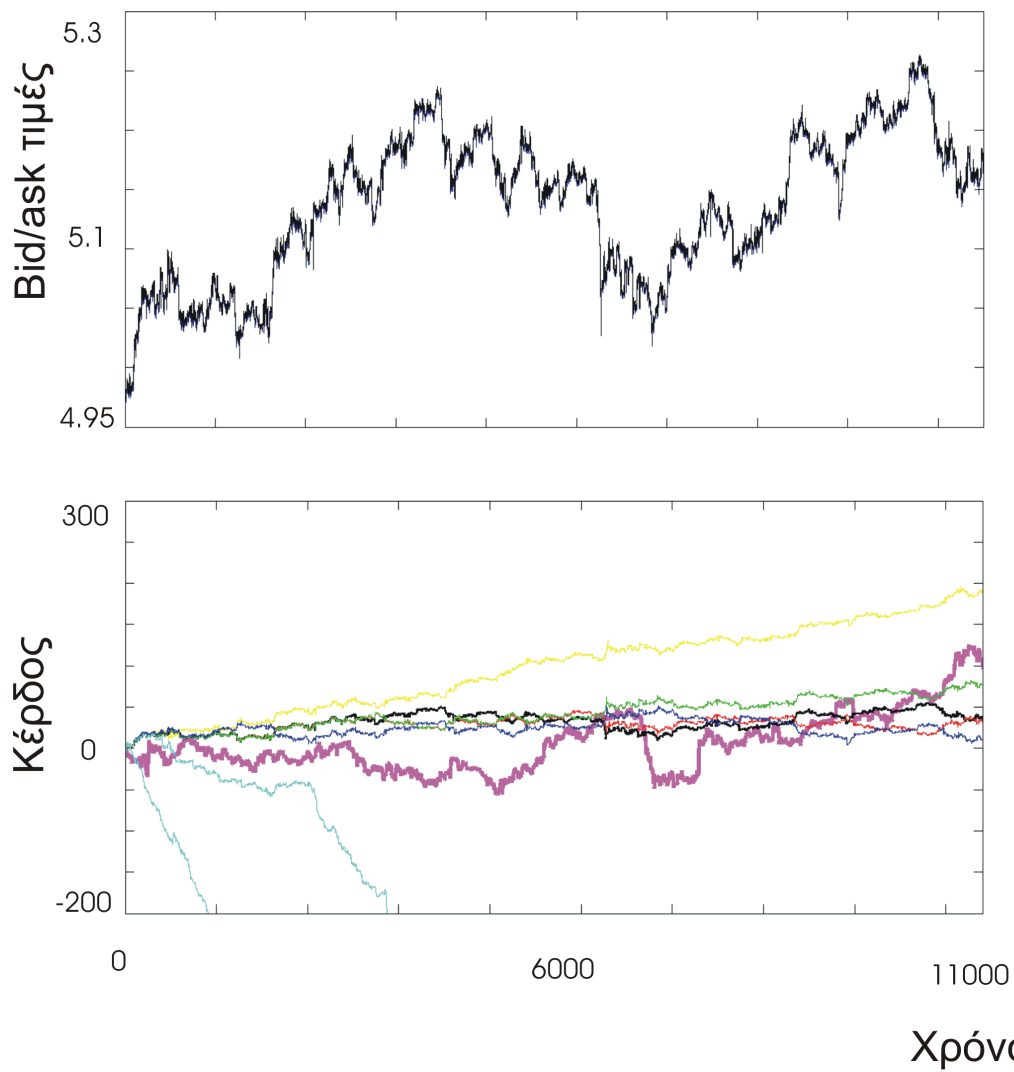
Στα σχήματα 7.10-7.13 φαίνονται εφαρμογές των ρυθμισμένων ΕΣ σε σύγκριση με αυτά τα παραπάνω συστήματα. Στα 7.10-7.12 είναι πάνω στις αγορές *GbpUsd* και *UsdFrf* για ΕΣ με 1 στρώμα, και στο 7.13 στην *UsdSgd* για ΕΣ με 2 στρώματα .



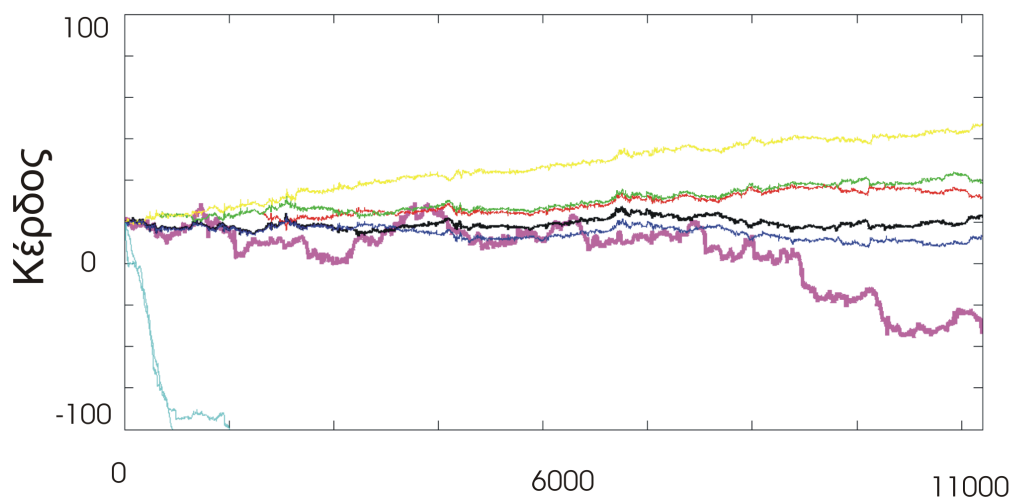
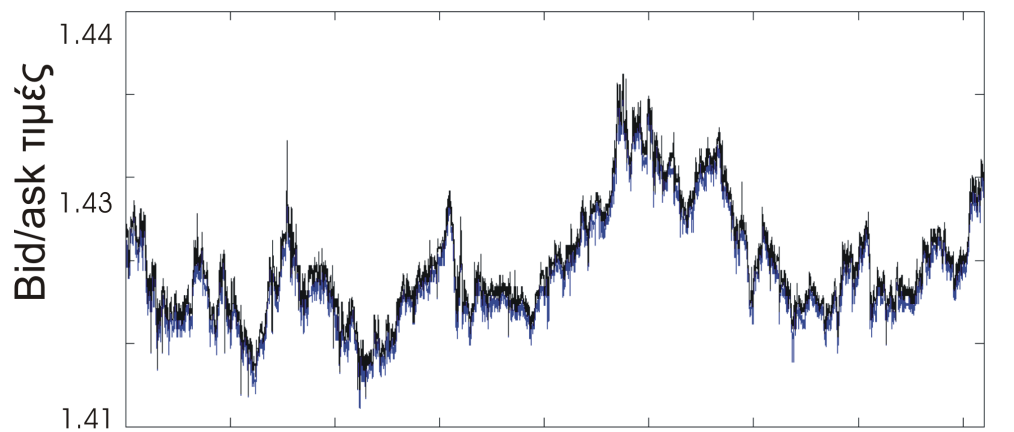


Χρόνος

Σχήμα 7.10: Εφαρμογή και σύγκριση της ρυθμισμένης ΕΣ με ANN 1 στρώματος στην αγορά *GbpUsd*. Η αντιστοιχία συστημάτων και χρωμάτων,σε αυτό και τα επόμενα δύο σχήματα,είναι η εξής : κίτρινο=(32)-μαντείο, πράσινο=(8)-μαντείο, κόκκινο=(4)-μαντείο, μαύρο = (1)-μαντείο, μπλέ = αφελές 3, ροζ=*RRL* , γαλάζιο = αφελές 1, γαλάζιο = (2)-αφελές 2



Σχήμα 7.11: Εφαρμογή και σύγκριση της ρυθμισμένης ΕΣ με ANN 1 στρώματος στην αγορά *UsdFrf*



Χρόνος

Σχήμα 7.12: Εφαρμογή και σύγκριση της ρυθμισμένης ΕΣ με ANN 2 στρωμάτων στην αγορά *UsdSgd*

Πίνακας 7.5: Συνολική αξιολόγηση και σύγκριση της *RRL* στρατηγικής. Η απόδοση είναι μετρημένη σε Δολλάρια

αγορά	M/S	RRL - 1	RRL - 2	αφελές 3	(1)-μαντείο	(4)-μαντείο	(8)-μαντείο	(32)-μαντείο
<i>DemEsp</i>	1.45	-10	-22	1	1	24	40	42
<i>DemFrf</i>	1.98	12	6	-1	11	27	31	35
<i>DemItl</i>	2.01	24	17	2	26	78	79	90
<i>DemSek</i>	1.89	54	45	18	37	39	59	92
<i>GbpUsd</i>	2.45	47	34	74	134	207	220	308
<i>UsdEsp</i>	1.02	-37	-7	19	39	73	87	120
<i>UsdFim</i>	0.97	-117	-30	31	22	167	170	221
<i>UsdFrf</i>	2.67	116	75	18	63	66	107	198
<i>UsdItl</i>	1.97	34	48	5	10	46	65	80
<i>UsdJpy</i>	2.55	149	35	-78	89	100	122	182
<i>UsdMyr</i>	1.45	-16	-4	20	4	42	47	51
<i>UsdNlg</i>	2.64	39	16	17	54	84	117	150
<i>UsdSek</i>	1.56	18	27	-26	17	31	86	160
<i>UsdSgd</i>	0.97	-45	-48	-18	8	14	20	41
<i>UsdXeu</i>	2.45	34	-3	-9	14	47	79	89
άθροισμα		302	189	73	529	1045	1329	1859

Στον πίνακα 6.5 φαίνεται η αξιολόγηση της *RRL* στρατηγικής με 1 και 2 στρώματα πάνω στις αγορές ελέγχου μετρημένη σε δολλάρια.

Τέλος, πέρα απ'την απόδοση, στον πίνακα 7.6 φαίνονται κάποιες μετρήσεις που χαρακτηρίζουν τον τρόπο που έγινε το εμπόριο στις αγορές ελέγχου.

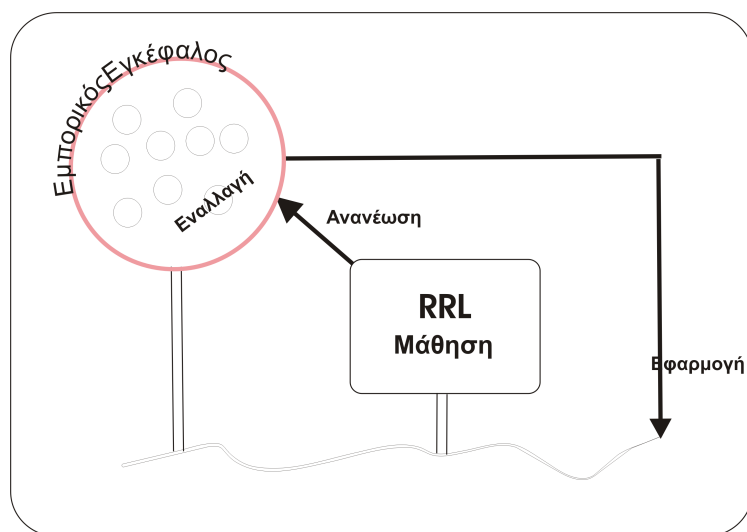
## 7.4 Συμπεράσματα και Μελλοντική Έρευνα

Απ'τα σχήματα της ρύθμισης παραμέτρων φαίνεται ότι η απόδοση είναι μέτρια ευαίσθητη στις αλλαγές των παραμέτρων. Αυτό σημαίνει ότι η ρύθμιση ήταν χρήσιμη αφού αν η απόδοση ήταν τελείως ευαίσθητη τότε παρα την αξιολόγηση θα ήταν αρνητικό χαρακτηριστικό για την ευστάθεια της ρυθμισμένης ΕΣ ανά τις διάφορες αγορές ελέγχου, ενώ αν ήταν τελείως αναίσθητη τότε δεν χρειαζόταν ρύθμιση.

Στον πίνακα 7.5 φαίνεται ότι η *RRL* ΕΣ τόσο για 1 και 2 στρώματα έχει θετικά αποτελέσματα. Το ANN με 1 στρώμα φαίνεται να αποδίδει καλύτερα, με την μεσή τιμή του κέρδους ανά τις 15 αγορές να είναι 20 δολλάρια και την τυπική απόκλιση στα 41, ενώ οι αντίστοιχες τιμές για το (1)-μαντείο είναι 37 και 40. Επιπλέον χαρακτηριστικές επιτυχίες της ΕΣ με 1 στρώμα είναι ότι στις αγορές *UsdFrf* και *UsdJpy* ξεπέρασε το (8)-μαντείο. Παρά αυτά τα καλά αποτελέσματα, η τελευταία γραμμή του πίνακα 7.6 δείχνει ότι το μέσο κέρδος για όλες τις αποφάσεις είναι αρκετά μικρό σε σχέση με τις απόλυτες τιμές των

Πίνακας 7.6: Μετρήσεις χαρακτηριστικών του εμπορίου στις αγορές ελέγχου. Συμβολίζοντας με  $h$  το σύνολο των διαστημάτων που ορίζονται με τα άκρα τους να είναι διαδοχικές αλλαγές θέσεων,  $\bar{H}$  = το μέσο μήκος των διαστημάτων του  $h$ ,  $\bar{H}^+$  = το μέσο μήκος των διαστημάτων του  $h$  που είχαν κέρδος,  $\bar{H}^-$  = το μέσο μήκος των διαστημάτων του  $h$  που είχαν απώλεια,  $P^+\%$  = ποσοστό των αποφάσεων κέρδους,  $\bar{P}^+$  = μέσο κέρδος για τις αποφάσεις κέρδους,  $\bar{P}^-$  = μέσο κέρδος για τις αποφάσεις απώλειας,  $\bar{P}$  = μέσο κέρδος για όλες τις αποφάσεις

αγορά	$\bar{H}$	$\bar{H}^+$	$\bar{H}^-$	$P^+\%$	$\bar{P}^+$	$\bar{P}^-$	$\bar{P}$
<i>DemEsp</i>	11.4	7.2	11.1	50.2	0.047	-0.049	-0.001
<i>DemFrf</i>	9.9	6.4	7.3	57.5	0.061	-0.080	0.001
<i>DemItl</i>	12.0	10.1	9.2	54.1	0.053	-0.058	0.002
<i>DemSek</i>	11.8	9.5	14.4	59.5	0.101	-0.136	0.005
<i>GbpUsd</i>	5.4	5.8	17.9	49.2	0.089	-0.076	0.005
<i>UsdEsp</i>	21.0	8.7	10.3	47.8	0.027	-0.032	-0.004
<i>UsdFim</i>	16.9	14.3	10.5	44.6	0.012	-0.031	-0.012
<i>UsdFrf</i>	12.2	12.3	12.1	67.1	0.058	-0.081	0.012
<i>UsdItl</i>	7.8	9.5	11.8	60.5	0.081	-0.116	0.003
<i>UsdJpy</i>	12.5	13.1	10.6	64.8	0.151	-0.233	0.015
<i>UsdMyr</i>	17.3	15.1	18.9	49.1	0.041	-0.043	-0.002
<i>UsdNlg</i>	22.6	19.5	16.6	56.2	0.103	-0.123	0.004
<i>UsdSek</i>	10.5	17.0	15.4	55.8	0.082	-0.099	0.002
<i>UsdSgd</i>	8.1	7.4	10.5	48.0	0.019	-0.027	-0.005
<i>UsdXeu</i>	11.6	19.3	15.4	60.2	0.094	-0.133	0.002
μέση τιμή	12.7	11.7	12.8	55.0	0.068	-0.088	0.002



Σχήμα 7.13: Εμπειρικός Εγκέφαλος

μέσων κερδών για αποφάσεις κέρδους και για αποφάσεις απώλειας θέτοντας ερωτήματα για την ευστάθεια της μεθόδου. Το γενικό συμπέρασμα είναι ότι τα αποτελέσματα αξιολογούν την μέθοδο *RRL* ως μια καλή μονάδα για μια εμπορική στρατηγική αφού ήταν πιο αποδοτική από κάθε αφελές σύστημα με το οποίο συγκρίθηκε, ωστόσο χρειάζεται έλεγχος σε δεδομένα σύγχρονων αγορών τα οποία διαφέροντας σε χαρακτηριστικά όπως το επίπεδο των *spreads* μπορεί να θέτουν την εξέλιξη της *RRL* στρατηγικής αναγκαία. Επιπλέον συμπέρασμα είναι ότι ο *Movement/Spread* λόγος είναι καλός δείκτης για το αν μια αγορά θα είναι κερδοφόρα αφού η ΕΣ με 1 στρώμα είχε θετικά αποτελέσματα σε κάθε αγορά ελέγχου με λόγο μεγαλύτερο του 2.

Όσον αφορά την σχετική μελλοντική έρευνα, μια βασική διάσταση είναι η αύξηση της κατάστασης του περιβάλλοντος με επιπλέον πληροφορίες που επηρεάζουν τις τιμές, όπως για παράδειγμα τα επιτόκια κάθε χώρας.

Για την άμεση συνέχεια των μετρήσεων που παρουσιάστηκαν χρειάζονται δεδομένα από σύγχρονες αγορές που μπορούν να βρεθούν στο διαδίκτυο<sup>2</sup> και επιπλέον έλεγχος παρόμοιων στρατηγικών με αυτήν που χρησιμοποιήθηκε. Μια ιδέα για επέκταση της στρατηγικής είναι η εξής : διατήρηση πολλών παράλληλων ANN με μάθηση *RRL* σε διάφορα διαστήματα τιμών και εναλλαγή μεταξύ αυτών ανά μια περίοδο όπως φαίνεται στο σχήμα 7.13. Το κριτήριο για την εναλλαγή μπορεί να τεθεί σε ευριστικό επίπεδο με μια επιλογή να είναι το πόσο καλή απόδοση έχει το κάθε ANN πρόσφατα, καθώς η λειτουργία του προσομοιώνεται χωρίς να κάνει εμπόριο πραγματικά.

<sup>2</sup>[www.histdata.com](http://www.histdata.com)

## Κεφάλαιο 8

# Υψηλής Συχνότητας Εμπόριο με το Πρωτόκολλο *FIX*

Έχοντας υλοποιήσει μια εμπορική στρατηγική το επόμενο βήμα είναι η εγκατάσταση ηλεκτρονικής επικοινωνίας με κάποιον μεσάζοντα της Αγοράς Συναλλάγματος μέσω ενός διαδικτυακού σταθμού. Το πρωτόκολλο *FIX* (Financial information eXchange) θέτει τους τυπικούς κανόνες για τέτοιου είδους επικοινωνίες οικονομικών συναλλαγών.

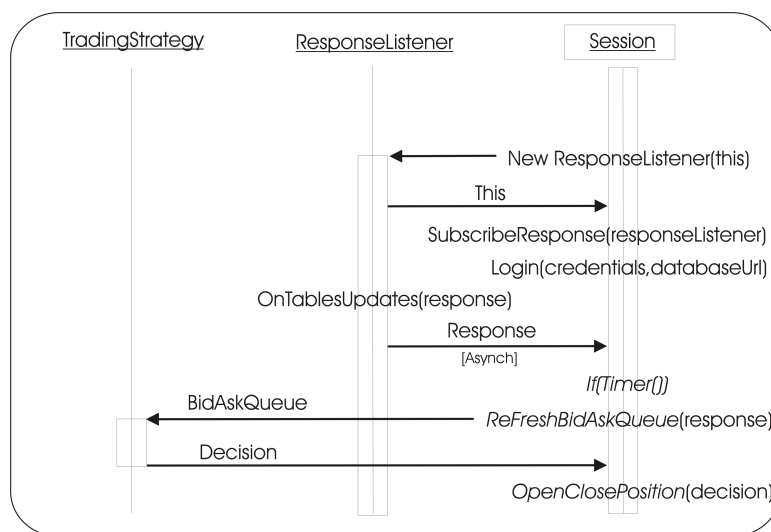
Το πρωτόκολλο *FIX* άρχισε να τυποποιείται από το 1992 για να γίνει δυνατό το ηλεκτρονικό εμπόριο μετοχών και ομολόγων μεταξύ οικονομικών οργανισμών και από τότε έχει επεκταθεί σε πολλές άλλες αγορές όπως η Αγορά Συναλλάγματος και η Αγορά Παραγώγων. Εν τω μεταξύ με το ποσοστό των συναλλαγών που συμβαίνει με αλγοριθμικό τρόπο να αυξάνεται έντονα στις μεγάλες αγορές και επιπλέον με την συχνότητα συναλλαγών να αυξάνεται επίσης η ανάγκη για μια τέτοια τυποποίηση έγινε όλο και πιο σημαντική. Υπάρχουν διαθέσιμες στο διαδίκτυο<sup>1</sup> διάφορες διεπαφές λογισμικού (APIs) που το υλοποιούν, όπως οι *ForexConnect*, *JavaTrading API* και *FIX API* καθώς και περιγραφές των περιεχομένων τους. Επιπλέον, υπάρχει ειδικός ιστότοπος που περιλαμβάνει τον τυπικό ορισμό του<sup>2</sup>.

Για την παρούσα εργασία χρησιμοποιήθηκε το *ForexConnect* σε γλώσσα *Java*. Το κεντρικό πρόγραμμα υλοποιήθηκε με *Java* και επικοινωνεί με τον μεσάζοντα με την παραπάνω διεπεφή, και με την εμπορική στρατηγική που υλοποιήθηκε με *Matlab* μέσω μιας *Java – Matlab* διεπαφής. Στην συνέχεια θα δούμε την γενική δομή που έχει μια τέτοια εφαρμογή με αναφορές σε κλάσεις του *ForexConnect* και στην αλληλεπίδραση μεταξύ τους ώστε να διατηρούμε την συμμετοχή μας στην αγορά για κάποιο σύμβολο *XxxYyy* και με μια δεδομένη

---

<sup>1</sup>[www.fxcodebase.com](http://www.fxcodebase.com)

<sup>2</sup>[www.fixprotocol.org](http://www.fixprotocol.org)



Σχήμα 8.1: Ακολουθιακό Διάγραμμα *UML* για την συνολική δομή της εφαρμογής

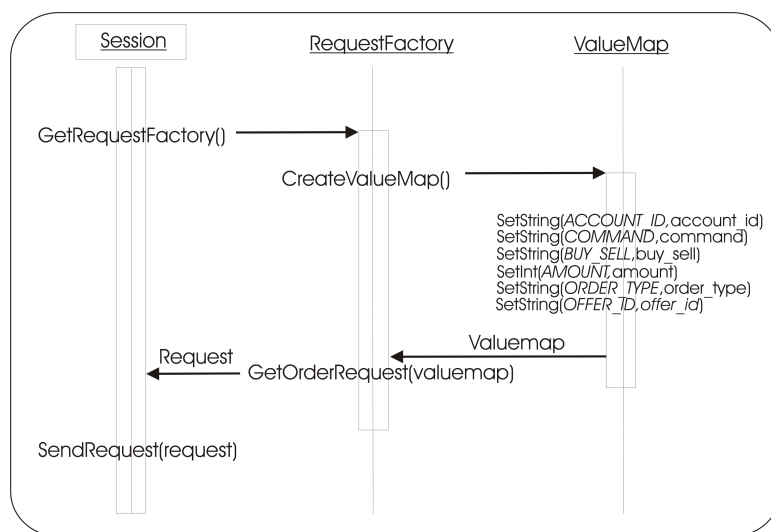
συχνότητα, ακριβώς δηλαδή όπως θεωρήσαμε την διαδικασία του εμπορίου στα προηγούμενα κεφάλαια.

Το σχήμα 8.2 δείχνει τη ακολουθία λειτουργίας των κύριων μεθόδων και την μεταφορά πληροφορίας έτσι ώστε να περιγράφεται η γενική δομή της εφαρμογής. Σημειώνουμε με πλάγια γραφή τις μεθόδους που δεν ανήκουν στο *ForexConnect*. Η *Timer()* απλά ελέγχει, με βάση την δεδομένη συχνότητα που πρέπει να συμβαίνει το εμπόριο, αν την τρέχουσα στιγμή πρέπει να ληφθεί απόφαση. Οι υπόλοιπες δύο μέθοδοι με πλάγια γράμματα θα αναλυθούν στα δύο επόμενα σχήματα.

Η βασική κλάση που περιγράφει την συνεδρία με τον μεσάζοντα είναι η *session* και δημιουργείται με την μέθοδο *createSession()* της κλάσης *O2GTransport*. Η συνεδρία εγγράφεται στην *ResponseListener* η οποία είναι μια διεπαφή υλοποιημένη έτσι ώστε να ενημερώνει την συνεδρία κάθε φορά που υπάρχει *response* με καινούργια δεδομένα από τον μεσάζοντα. Μετά την λήψη των δεδομένων, και αν η *Timer()* κρίνει ότι ο χρόνος είναι κατάλληλος δηλαδή έχει περάσει η δεδομένη περίοδος από την προηγούμενη απόφαση μας, ανανεώνουμε την ουρά που περιέχει τις *bid/ask* τιμές του συμβόλου *XxxYyy* με την μέθοδο *RefreshBidAskQueue*, στέλνουμε την ανανεωμένη ουρά στην εμπορική στρατηγική και με την απόφαση που επιστρέφεται διαχειριζόμαστε την θέση μας με την μέθοδο *OpenClosePosition*.

Ανάλογα με την τρέχουσα και την προηγούμενη απόφαση, η *OpenClosePosition* αποφασίζει αν πρέπει να αλλάξουμε θέση, κλείνοντας την προηγούμενη και ανοίγωντας την αντίθετη καινούργια ή αν πρέπει να διατηρήσουμε την θέση



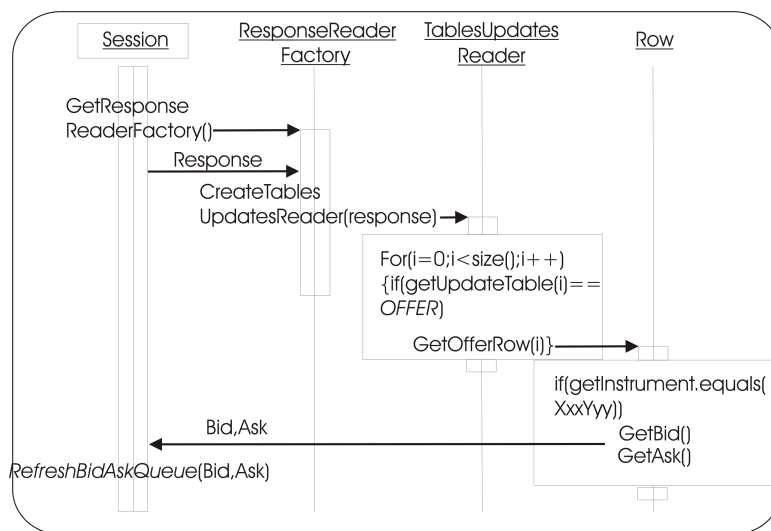


Σχήμα 8.2: Ακολουθιακό Διάγραμμα UML για αγορά και πώληση

μας δηλαδή να μην γίνει τίποτα.

Το σχήμα 8.3 δείχνει πως ανοίγει ή κλείνει μια θέση. Η συνεδρία δημιουργεί ένα RequestFactory το οποίο διαμορφώνει μια αντιστοιχία valuemap μεταξύ των χαρακτηριστικών μιας θέσης και των τιμών που θέλουμε να τους προσδώσουμε. Για παράδειγμα η σταθερά AMOUNT συμβολίζει την ποσότητα με την οποία συσχετίζεται η θέση, κάτι το οποίο μέχρι τώρα συμβολιζόταν με  $\mu$ , και η σταθερά OFFER ID αναφέρεται στον κωδικό του συμβόλου XxxYyy. Η υπόλοιπη ακολουθία δείχνει ότι το διαμορφωμένο valuemap χρησιμοποιείται από το RequestFactory για να δημιουργηθεί μια αίτηση η οποία στέλνεται τελικά από την συνεδρία στον μεσάζοντα.

Το σχήμα 8.4 περιγράφει την μέθοδο *RefreshBidAskQueue*, δηλαδή την εξαγωγή των τρέχοντων bid/ask τιμών από response που στέλνει ο μεσάζοντας. Αρχικά δημιουργείται ένα ResponseReaderFactory το οποίο με την σειρά του δημιουργεί έναν TablesUpdatesReader. Αυτός ο αναγνώστης είναι άμεσα συσχετισμένος με την πληροφορία που έρχεται μέσω του response η οποία είναι διαμορφωμένη σαν μια ακολουθία από γραμμές με την κάθε γραμμή να ανήκει σε έναν πίνακα. Οι βασικοί πίνακες που υπάρχουν είναι ο πίνακας OFFERS που έχει τα σύμβολα και τις τρέχουσες τιμές τους και ο πίνακας TRADES που έχει τις τρέχουσες ανοιχτές θέσεις. Τα κομμάτια ελέγχου που φαίνονται στο σχήμα 8.4 χρησιμεύουν ώστε να αναγνωριστεί η γραμμή που μας ενδιαφέρει, δηλαδή αυτή που ανήκει στον πίνακα OFFERS και αντιστοιχεί στο σύμβολο XxxYyy. Μετά την λήψη τους από αυτήν την γραμμή, οι ενδιαφέρουσες bid/ask τιμές ανανεώνουν την bid/ask ουρά ώστε να είναι έτοιμη να χρησιμοποιηθεί από την εμπορική στρατηγική.



Σχήμα 8.3: Ακολουθιακό Διάγραμμα UML για την ανανέωση των bid/ask δεδομένων

# Βιβλιογραφία

- [1] John Moody, Lizhong Wu, Yuansong Liao, Matthew Saffel, *Performance functions and Reinforcement Learning for trading systems and portfolios*, Journal of Forecasting, Volume 17 Issue 5 pp. 441-470, 1998
- [2] John Moody, Matthew Saffel, *Learning to trade via Direct Reinforcement*, Journal of IEEE Transactions on Neural Networks, Volume 12 Issue 4 pp. 875-889, 2001
- [3] John Moody, *Prediction risk and architecture selection for Neural Networks*, From Statistics to Neural Networks: Theory and Pattern Recognition Applications, 1994
- [4] M.Dempster, V.Leemans, *An automated FX trading system using Adaptive Reinforcement Learning*, Expert Systems with Applications, Volume 30 Issue 3 pp. 543-552, 2006
- [5] Gabriel Molina, *Stock Trading with Recurrent Reinforcement Learning*
- [6] Carl Gold, *FX Trading via Recurrent Reinforcement Learning*, IEEE International Conference on Computational Intelligence for Financial Engineering, pp. 363-370, 2003
- [7] Xin Du ,Koupin Lv, Jinjian Zhai, *Algorithm Trading using Q-Learning and Recurrent Reinforcement Learning*
- [8] C. Lee Giles, Steve Lawrence, Ah Chung Tsoi, *Noisy Time Series Prediction using a Recurrent Neural Network and Grammatical Inference*, Journal of Machine Learning, Volume 44 Issue 1-2 pp. 161-183, 2001
- [9] Denise Gorse, *Application of stochastic recurrent reinforcement learning to index trading*, European Symposium on Artificial Neural Networks, 2011

- [10] Francesco Bertoluzzo, Marco Corazza, *Making Financial Trading by Recurrent Reinforcement Learning*, Proceedings of the 11th International Conference on Neural Networks, 2007
- [11] David J.C. MacKay, *Information Theory, Inference and Learning Algorithms*, Cambridge University Press, 2002
- [12] Christopher M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, 2006
- [13] S.Russel, P.Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2010
- [14] R.Duda, R.Hart, D.Stork, *Pattern Classification*, Wiley-Interscience, 2000
- [15] Daphne Koller, Nir Friedman, *Probabilistic Graphical Models*, MIT Press, 2009
- [16] Pierre Baldi, Michal Rosen-Zvi, *On the Relationship Between Deterministic and Probabilistic Directed Graphical Models: from Bayesian Networks to Recursive Neural Networks*, Journal of Neural Networks, Volume 18 Issue 8 pp. 1080-1086, 2005
- [17] Barbara Hammer, *Learning with Recurrent Neural Networks*, Springer, 2008
- [18] Simone Santini, Alberto Del Bimbo, Ramesh Jain, *Block-structured Recurrent Neural Networks*, Journal of Neural Networks, Volume 8 Issue 1 pp. 135-147, 1995
- [19] Raul Rojas, *Neural Networks A Systematic Introduction*, Springer, 1996
- [20] Simon Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd edition, Prentice Hall, 1998
- [21] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: an Introduction*, MIT Press, 1998
- [22] Warren B.Powell, *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, 2nd edition, Wiley-Interscience, 2011

- [23] Dimitri Bertsekas, John Tsitsiklis *Neuro-Dynamic Programming*, Athena Scientific, 1996
- [24] Paul Krugman, Maurice Obstfeld, Marc Melitz, *International Economics: Theory and Policy*, 9th edition, Prentice Hall, 2011
- [25] Jürgen Schmidhuber, *Reinforcement learning with interacting continually running fully recurrent networks*, International Neural Network Conference, 1990
- [26] Katsunari Shibata, *Emergence of Intelligence Through Reinforcement Learning with a Neural Network*, Advances in Reinforcement Learning, 2011
- [27] Bank of International Settlements, *Report on global foreign exchange market activity in 2010*
- [28] Michael R King, Dagfinn Rime, *The 4 trillion question: what explains FX growth since the 2007 survey?*
- [29] Paul J. Werbos, *Backpropagation through time, what it does and how to do it*, Proceedings of the IEEE, Volume 78 Issue 10, 1990
- [30] Pieter Bram Bakker, *The State of Mind: Reinforcement Learning with Recurrent Neural Networks*, PhD thesis, Leiden University, 2004
- [31] James Nate Knight, Charles Anderson, *Stable reinforcement learning with recurrent neural networks*, Journal of Control Theory and Applications, Volume 9 Issue 3 pp. 410-420, 2011
- [32] Herrn Anton Maximilian Schafer, *Reinforcement Learning with Recurrent Neural Networks*, PhD thesis, Osnabrück University, 2008
- [33] Daniela Pucci de Farias, Benjamin van Roy, *The Linear Programming Approach to Approximate Dynamic Programming*, Operations Research, Volume 51 Issue 6 pp. 850-865, 2003
- [34] Eric Zivot, *Analysis of High Frequency Financial Time Series: Methods, Models and Software*, 2005
- [35] Nan Zhou, Wen Cheng, *Non-linear Supervised High Frequency Trading Strategies with Applications in US Equity Markets*, the 4th Annual Modeling High Frequency Data in Finance Conference, 2012

- [36] C.Ullrich, S.Chalup, D.Seese , *Investigating FX-Market Efficiency with Support Vector Machines*
- [37] Qinghua Wen, Zehong Yang, Yixu Song, Peifa Jia, *Automatic stock decision support system based on box theory and SVM algorithm*, Journal of Expert Systems with Applications, Volume 37 Issue 2 pp. 1015-1022, 2010
- [38] Hailin Li, Cihan H. Dagli, and David Enke, *Forecasting Series-based Stock Price Data using Direct Reinforcement Learning*, IEEE International Joint Conference on Neural Networks, 2004
- [39] Paolo Tenti, *Forecasting foreign exchange rates using recurrent neural networks*, Applied Artificial Intelligence: An International Journal, Volume 10 Issue 6 pp. 567-582, 1996
- [40] A.Nag, A.Mitra, *Forecasting Daily Foreign Exchange Rates Using Genetically Optimized Neural Networks*, Journal of Forecasting, Volume 21 Issue 7, pp.501-511, 2002
- [41] Gordon H. Dash, Nina Kajiji, *Prediction of FX Volatility via an RBF Neural Network with Closed- Form Regularization*, 8th International Conference on Forecasting Financial Markets, 2001
- [42] Yan Chen, Shingo Mabu and Kotaro Hirasawa, *Genetic Network Programming with Reinforcement Learning and Its Application to Creating Stock Trading Rules*, 2009
- [43] Jae Won Lee, Jonghun Park, Jangmin O, Jongwoo Lee, and Euyseok Hong, *A Multiagent Approach to Q-Learning for Daily Stock Trading*, Journal of IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, Volume 37 Issue 6, pp. 864-877, 2007
- [44] R.G. Bates, M.A.H. Dempster and Y.S. Romahi, *Evolutionary Reinforcement Learning in FX Order Book and Order Flow Analysis*, Proceedings of the IEEE International Conference on Computational Intelligence for Financial Engineering, 2003
- [45] T.W. Payne, M.A.H. Dempster and Y.S. Romahi, *Intraday FX trading: Reinforcement VS Evolutionary Learning*, Proceedings of the Third International Conference on Intelligent Data Engineering and Automated Learning, pp. 347-358, 2002

- [46] M.A.H. Dempster, T.W. Payne, Y.S. Romahi, and G.W.P. Thompson, *Computational Learning Techniques for Intraday FX Trading Using Popular Technical Indicators*, IEEE Transactions on Neural Networks, Volume 12 Issue 4 pp. 744-754, 2001
- [47] Alexander A. Sherstov and Peter Stone, *Three Automated Stock-Trading Agents: A Comparative Study*, Proceedings of the 6th AAMAS international conference on Agent-Mediated Electronic Commerce: theories for and Engineering of Distributed Mechanisms and Systems, pp. 173-187, 2005
- [48] Stuart Duerson, Farhan Saleem Khan, Victor Kovalev, Ali Hisham Malik, *Reinforcement Learning in Online Stock Trading Systems*
- [49] Naoki Abe, Edwin Pednault, Haixun Wang, Bianca Zadrozny, Wei Fan, Chid Apte, *Empirical Comparison of Various Reinforcement Learning Strategies for Sequential Targeted Marketing*, IEEE International Conference on Data Mining, 2002
- [50] Rui Pedro Barbosa and Orlando Belo, *Autonomous Forex Trading Agents*, Industrial Conference on Data Mining, pp.389-403, 2008
- [51] Patrik Idvall, Conny Jonsson, *Algorithmic Trading Hidden Markov Models on Foreign Exchange Data*, Master thesis, Linköpings University, 2008
- [52] István Szita, András Lőrincz, *PIRANHA: Policy Iteration for Recurrent Artificial Neural Networks with Hidden Activities*
- [53] Ralph Grothmann, *Multiagent Market Modelling based on Neural Networks*, PhD thesis, University of Bremen, Siemens AG, 2002