

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ  
ΘΕΣΣΑΛΟΝΙΚΗΣ  
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ -  
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ**

**Αυτοματοποιημένη δημιουργία  
προγραμμάτων εξεταστικής περιόδου.**

(Βελτιστοποίηση με εφαρμογή Παράλληλων Εξελικτικών Αλγορίθμων)

Πτυχιακή εργασία του : **Μιχαήλ Βρέττα**  
Εισηγητής Καθηγητής : **Δρ. Παναγιώτης Αδαμίδης**

ΘΕΣΣΑΛΟΝΙΚΗ 2004

# ΠΕΡΙΕΧΟΜΕΝΑ

<b>ΠΕΡΙΛΗΨΗ</b>	σελ.04
<b>ABSTRACT</b>	σελ.05
<b>ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ</b>	σελ.06
<b>1. Εισαγωγή.</b>	σελ.08
<b>2. Εξελικτικοί Αλγόριθμοι.</b>	σελ.10
2.1. Εισαγωγή	σελ.10
2.2. Βιολογικό υπόβαθρο	σελ.11
2.3. Ένας τυπικός ΕΑ	σελ.12
2.4. Βασικά μοντέλα εξελικτικών αλγορίθμων	σελ.14
2.4.1. Γενετικοί Αλγόριθμοι	σελ.14
2.4.2. Εξελικτικές Στρατηγικές	σελ.16
2.4.3. Εξελικτικός Προγραμματισμός	σελ.17
2.5. Παράλληλοι Εξελικτικοί Αλγόριθμοι	σελ.18
2.5.1. Τυπική παράλληλη προσέγγιση	σελ.18
2.5.2. Προσέγγιση του διαχωρισμού	σελ.19
2.6. Εφαρμογές Εξελικτικών Αλγορίθμων	σελ.21
<b>3. Πρόγραμμα εξετάσεων.</b>	σελ.23
3.1. Περιγραφή του προβλήματος	σελ.24
3.2. Επίλυση με διάφορες μεθόδους	σελ.29
<b>4. Εξελικτικοί αλγόριθμοι και πρόγραμμα εξετάσεων.</b>	σελ.32
4.1. Κωδικοποίηση του προβλήματος	σελ.33
4.2. Χρονοπρογραμματισμός Εξετάσεων & ΕΑ	σελ.38
<b>5. Υλοποίηση.</b>	σελ.45
5.1. Δημιουργία Χρωμοσώματος (Αναπαράσταση Λύσης)	σελ.45
5.2. Σχεδιασμός Συνάρτησης Αξιολόγησης	σελ.50
5.2.1. Περιορισμοί ικανοποίησης	σελ.51
5.2.2. Απόδοση Ποινών	σελ.55
5.3. Τελεστές	σελ.58
5.3.1. Τελεστές Ανασυνδυασμού	σελ.59
5.3.2. Τελεστές Μετάλλαξης	σελ.59
5.3.3. Τελεστές Επιλογής	σελ.60
5.3.4. Τελεστές Επιβίωσης	σελ.65
5.4. Ελιτισμός	σελ.66
5.5. Αναπροσαρμογή πιθανοτήτων	σελ.67
5.6. Σενάρια εκτέλεσης Εξελικτικού Αλγορίθμου	σελ.68
5.6.1. Ένας πληθυσμός – Μία παραμετροποίηση	σελ.69
5.6.2. 'N' πληθυσμοί – 'N' παραμετροποιήσεις	σελ.69
5.7. Προσομοίωση Πολέμου μεταξύ πληθυσμών	σελ.72
5.8. Προσομοίωση Ίωσης πληθυσμού	σελ.76
5.9. Κριτήρια τερματισμού εξέλιξης	σελ.77
<b>6. Αποτελέσματα.</b>	σελ.79
6.1. Χαρακτηριστικά πειραμάτων	σελ.79
6.2. Παρουσίαση αποτελεσμάτων	σελ.82

<b>7. Συμπεράσματα - Περαιτέρω Έρευνα.</b>	σελ.86
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b>	σελ.89
<b>ΠΑΡΑΡΤΗΜΑ Α:</b> Θεώρημα Σχημάτων (Schema Theorem)	σελ.92
<b>ΠΑΡΑΡΤΗΜΑ Β:</b> Exam Timetabling Glossary	σελ.96

# ΠΕΡΙΛΗΨΗ

Η δημιουργία ικανοποιητικών προγραμμάτων εξετάσεων, στην τριτοβάθμια εκπαίδευση, ήταν ανέκαθεν ένα από τα προβλήματα που αντιμετώπιζαν τα ανώτερα εκπαιδευτικά ιδρύματα. Πολλές μέθοδοι αναπτύχθηκαν, για την επίλυσή του, άλλες με επιτυχία και άλλες με όχι και τόσο καλά αποτελέσματα. Νέες προοπτικές όμως εμφανίστηκαν στο χώρο, παράλληλα με την ανάπτυξη των Εξελικτικών Αλγορίθμων (ΕΑ) οι οποίοι τις τελευταίες τρεις δεκαετίες κερδίζουν όλο και περισσότερο την προσοχή των επιστημόνων για την επίλυση δύσκολων και πολύπλοκων προβλημάτων.

Η αντιμετώπιση του προβλήματος του χρονοπρογραμματισμού των εξετάσεων από τους εξελικτικούς αλγορίθμους, ήρθε με την παρατήρηση ότι αυτό το πρόβλημα μπορεί να παρομοιασθεί και να μετατραπεί σε πρόβλημα χρωματισμού ενός γράφου (graph colouring problem), κάτι το οποίο οι ΕΑ αντιμετώπιζαν με επιτυχία. Έτσι λοιπόν η εκπαιδευτική επιστημονική κοινότητα έστρεψε το ενδιαφέρον της στους ΕΑ για να επιλύσει ένα πάγιο πρόβλημά της.

Στην παρούσα εργασία γίνεται μία προσπάθεια ανάπτυξης ενός ΕΑ για την επίλυση αυτού του προβλήματος στο τμήμα Πληροφορικής του Τεχνολογικού Εκπαιδευτικού Ιδρύματος της Θεσσαλονίκης. Για την επίτευξη καλύτερων ποιοτικά αποτελεσμάτων δίνεται η δυνατότητα εφαρμογής Παράλληλων ΕΑ, των οποίων τα πλεονεκτήματα και η χρησιμότητα αναφέρεται στα κείμενα που ακολουθούν.

Τέλος, ο ΕΑ που αναπτύχθηκε εφαρμόστηκε σε μία σειρά πειραμάτων, πάνω σε δοκιμαστικά δεδομένα, και τα αποτελέσματα που προέκυψαν ήταν κάτι παραπάνω από ενθαρρυντικά.

# ABSTRACT

The development of satisfying exam programs in tertiary education has always been one of the problems that the educational institutes are facing. Many methods were developed to solve this problem; some of them were successful while other did not have the expected results. New perspectives came up in the field, with the development of the Evolutionary Algorithms (EAs). The last three decades, EAs are gaining more and more attention from the scientists in solving difficult and complex problems.

The confrontation of the exam-timetabling problem using EAs started with the observation that this problem can be simulated by and transformed into a graph colouring problem, which the EAs were dealing with much success. Thus, the educational scientific community turn their interest to EAs in order to find a solution for this long-standing problem.

In this thesis, an effort is being made to develop a solution based on EAs, for the exam-timetabling problem at the department of Informatics, of the Technological Educational Institute of Thessaloniki. For the achievement of qualitatively better results the capability of applying parallel Evolutionary Algorithms is given. The advantages and usability of Parallel EAs is also described in this thesis.

Finally, the EA that has been developed was applied in a series of experiments, on test data, and the results were more than encouraging.

# ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα	Περιγραφή	Σελίδα
2.1	Ένας τυπικός εξελικτικός αλγόριθμος.	13
2.2	Εξελικτικός Κύκλος.	14
2.3	Διασταύρωση ενός σημείου, από την τέταρτη θέση.	15
2.4	Μετάλλαξη στη δεύτερη θέση του χρωμοσώματος.	16
2.5	Σχηματική αναπαράσταση της τυπικής παράλληλης προσέγγισης.	19
2.6.α	Πλήρης Διασύνδεση Νησίδων.	20
2.6.β	Διασύνδεση Δακτυλίου.	20
3.1	Κατηγοριοποίηση Χρονοπρογραμματισμού στην Ανώτατη Εκπαίδευση.	23
3.2	Ενδεικτική αναπαράσταση συνολικής χρονικής διάρκειας της εξεταστικής περιόδου. Σε κάθε μέρα αντιστοιχούμε το ίδιο σύνολο χρονοθυρίδων.	27
4.1	Παράδειγμα άμεσης αναπαράστασης χρωμοσώματος.	33
4.2	Παράδειγμα έμμεσης αναπαράστασης χρωμοσώματος.	34
4.3	Ανασυνδυασμός ενός σημείου.	36
4.4	Ανασυνδυασμός τεσσάρων σημείων.	37
4.5	Ομοιόμορφος Ανασυνδυασμός.	37
4.6	Παράδειγμα κατευθυνόμενης μετάλλαξης.	38
4.7	Λειτουργία ευριστικού τελεστή διασταύρωσης.	40
4.8	Άμεση αναπαράσταση στον "memetic" algorithm.	42
5.1	Α' αναπαράσταση χρωμοσώματος.	46
5.2	Πίνακας με τις διαθέσιμες χρονοθυρίδες (time slots).	46
5.3	Β' αναπαράσταση χρωμοσώματος.	48
5.4	Γ' αναπαράσταση χρωμοσώματος.	48
5.5	Τελική αναπαράσταση χρωμοσώματος.	49
5.6	Παράδειγμα συνδεδεμένων χρονοθυρίδων.	53
5.7	Αναπαράσταση τελεστών επιλογής σε 2Δ πλέγμα (9, 10, 11).	64
5.8	Α σενάριο εκτέλεσης EA.	69
5.9	Β σενάριο εκτέλεσης EA.	70
5.10	Γ σενάριο εκτέλεσης EA.	70
5.11	Δ σενάριο εκτέλεσης EA.	72
5.12	Παράδειγμα εφαρμογής βομβαρδισμού.	74
5.13	Μόλυνση χρωμοσώματος σε ποσοστό 40%.	76
6.1	Συνοπτικά αποτελέσματα πειράματος "Exp01".	82

6.2	Συνοπτικά αποτελέσματα πειράματος "Εxp02".	82
6.3	Συνοπτικά αποτελέσματα πειράματος "Εxp03".	82
6.4	Συνοπτικά αποτελέσματα πειράματος "Εxp04".	83
6.5	Συνοπτικά αποτελέσματα πειράματος "Εxp05".	83
6.6	Συνοπτικά αποτελέσματα πειράματος "Εxp06".	83

# ΚΕΦΑΛΑΙΟ 1

## Εισαγωγή

Το πρόβλημα της δημιουργίας ενός ικανοποιητικού προγράμματος εξεταστικής περιόδου υπάρχει σε όλα τα ανώτερα και ανώτατα εκπαιδευτικά ιδρύματα. Το ζήτημα που τίθεται όμως, σε κάθε ένα από αυτά, είναι τί αποτελεί ένα «ικανοποιητικό» πρόγραμμα. Η απάντηση δεν μια και ποικίλει από ίδρυμα σε ίδρυμα. Ο λόγος που συμβαίνει αυτό είναι ότι κάθε ίδρυμα χειρίζεται με διαφορετικό τρόπο τη δημιουργία και την υλοποίηση ενός τέτοιου προγράμματος. Αυτό οφείλεται κυρίως στο γεγονός ότι η δημιουργία ενός προγράμματος εξεταστικής περιόδου έχει κάποιες ιδιαιτερότητες που οφείλονται άμεσα στη φύση του προβλήματος και εξαρτώνται από τους διαθέσιμους πόρους και ανάγκες κάθε ιδρύματος.

Σε ορισμένα ιδρύματα η υλοποίηση ενός προγράμματος εξετάσεων που περιλαμβάνει όλα τα εξεταζόμενα μαθήματα είναι αρκετή. Σε άλλες περιπτώσεις δεν αρκεί μόνο αυτό, αλλά θα πρέπει το πρόγραμμα να ικανοποιεί και ένα σύνολο παραμέτρων, έτσι ώστε να μπορεί να χαρακτηριστεί ως ικανοποιητικό. Ακόμα όμως και σε αυτή την περίπτωση, το σύνολο των παραμέτρων που πρέπει να ικανοποιηθεί ποικίλει, από ίδρυμα σε ίδρυμα, τόσο ποιοτικά, όσο και ποσοτικά. Το γεγονός αυτό κάνει τη δημιουργία ενός προγράμματος εξετάσεων μία τόσο εξειδικευμένη διαδικασία που είναι πολύ δύσκολο, έως αδύνατο, να «τυποποιηθεί» και να αντιμετωπιστεί με κοινό τρόπο.

Στην παρούσα εργασία γίνεται μία προσπάθεια να αυτοματοποιηθεί η δημιουργία των προγραμμάτων εξεταστικής περιόδου στο τμήμα Πληροφορικής, της Σχολής Τεχνολογικών Εφαρμογών του Τεχνολογικού Εκπαιδευτικού Ιδρύματος Θεσσαλονίκης. Η εν λόγω διαδικασία μέχρι σήμερα πραγματοποιείται χειρωνακτικά από κάποιο μέλος του εκπαιδευτικού προσωπικού, του τμήματος, το οποίο αναλαμβάνει αυτήν τη διαδικασία χρησιμοποιώντας κυρίως την εμπειρία του και αξιοποιώντας παλαιότερα προγράμματα εξετάσεων. Για το σκοπό αυτό σχεδιάστηκε και υλοποιήθηκε ένα πλήρως αυτοματοποιημένο υπολογιστικό σύστημα το οποίο αναλαμβάνει τη δημιουργία των προγραμμάτων εξετάσεων κάνοντας χρήση εξελικτικών αλγορίθμων. Μάλιστα για τη βελτιστοποίηση των



αποτελεσμάτων δίνεται η δυνατότητα της εφαρμογής παράλληλων εξελικτικών αλγορίθμων. Η λειτουργία και η χρησιμότητα των εξελικτικών αλγορίθμων θα παρουσιαστεί, εν συντομία, σε σχετικό κεφάλαιο που ακολουθεί. Ο αλγόριθμος που υλοποιήθηκε, σχεδιάστηκε για να καλύπτει όχι μόνο τις παρούσες ανάγκες του τμήματος, αλλά και μελλοντικές που ενδεχομένως να προκύψουν. Με αυτόν τον τρόπο ανοίγει και ένα μικρό παράθυρο για την εφαρμογή του αλγορίθμου και σε άλλα ιδρύματα, καθώς δεν είναι σχεδιασμένος να λειτουργεί αποκλειστικά με τις ανάγκες ενός συγκεκριμένου τμήματος.

Στα κεφάλαια που ακολουθούν γίνεται, αρχικά μία περιγραφή των εξελικτικών αλγορίθμων (δεύτερο κεφάλαιο). Αναφέρεται τόσο το βιολογικό τους υπόβαθρο, όσο και τα βασικά μοντέλα που έχουν κυριαρχήσει σε αυτόν το χώρο. Τέλος γίνεται μία μικρή εισαγωγή στους παράλληλους εξελικτικούς αλγορίθμους, μίας και για τους σκοπούς της εργασίας υλοποιήθηκε ένας παράλληλος αλγόριθμος, ενώ αναφέρονται και μερικοί τομείς που μπορούν να βρουν εφαρμογή οι εξελικτικοί αλγόριθμοι.

Στο τρίτο κεφάλαιο, της παρούσης εργασίας, δίνεται ένας γενικός ορισμός του προβλήματος της δημιουργίας προγραμμάτων εξετάσεων. Επίσης περιγράφονται κύριες παράμετροι του προβλήματος, ενώ αναφέρονται και διάφοροι τρόποι που μπορούν να εφαρμοστούν για την επίλυσή του.

Στο τέταρτο κεφάλαιο αρχικά δίνεται ένα παράδειγμα κωδικοποίησης του προβλήματος σε μορφή κατανοητή από τους εξελικτικούς αλγορίθμους, ενώ ακολουθεί μία βιβλιογραφική ανασκόπηση των εξελικτικών αλγορίθμων που χρησιμοποιήθηκαν κατά καιρούς, από διάφορους επιστήμονες, σε διάφορα ιδρύματα για την επίλυση αυτού του προβλήματος, καθώς και τα αποτελέσματα που προέκυψαν από τα πειράματά τους.

Στο πέμπτο κεφάλαιο, περιγράφεται ο σχεδιασμός του παράλληλου εξελικτικού αλγορίθμου που υλοποιήθηκε. Αναφέρονται όλα τα στάδια της δημιουργίας του αλγορίθμου, καθώς και όλες οι παράμετροι που λήφθηκαν υπόψιν. Δίνεται ιδιαίτερη έμφαση στο σχεδιασμό του χρωμοσώματος καθώς και στα στάδια από τα οποία πέρασε για να πάρει την τελική του μορφή. Επίσης περιγράφονται όλοι οι περιορισμοί που ικανοποιεί ο αλγόριθμος, είτε άμεσα, είτε έμμεσα. Τέλος αναφέρονται με λεπτομέρειες όλοι οι τελεστές που χρησιμοποιήθηκαν καθώς και όλες οι υπόλοιπες παράμετροι που συμβάλουν στην ολοκλήρωση του αλγορίθμου.

Στο έκτο κεφάλαιο γίνεται μία παρουσίαση των πειραμάτων που πραγματοποιήθηκαν, καθώς και των αποτελεσμάτων που προέκυψαν από αυτά. Κατόπιν ακολουθεί ένα μικρό κεφάλαιο όπου περιλαμβάνει τα συμπεράσματα των πειραμάτων και γενικότερα ολόκληρης της εργασίας, ενώ τέλος προτείνει μερικά σημεία στα οποία θα πρέπει να γίνει περισσότερη έρευνα πάνω στην εφαρμογή των εξελικτικών αλγορίθμων για την επίλυση του προβλήματος της δημιουργίας προγραμμάτων εξεταστικής περιόδου.

Τέλος παρατίθενται, σε μορφή παραρτήματος, το θεώρημα των σχημάτων (schema theorem) που προτάθηκε το 1975 από τον John Holland, για την εξήγηση της λειτουργίας των Γενετικών Αλγορίθμων (ΓΑ), ενώ το τελευταίο παράρτημα είναι μία τυπική περιγραφή του προβλήματος χρονοπρογραμματισμού εξετάσεων, σε μορφή γλωσσάρου.

# ΚΕΦΑΛΑΙΟ 2

## Εξελικτικοί Αλγόριθμοι

### 2.1 Εισαγωγή

Οι Εξελικτικοί Αλγόριθμοι (ΕΑ) είναι αλγόριθμοι ανίχνευσης – αναζήτησης, βασισμένοι στη μηχανική της φυσικής επιλογής και της εξέλιξης των ειδών. Συνδυάζουν την επιβίωση του ικανότερου με μία οργανωμένη ανταλλαγή πληροφοριών, με στόχο τη διαμόρφωση ενός αλγόριθμου αναζήτησης που να διαθέτει, όσο μπορεί να είναι κάτι τέτοιο εφικτό, τη νεωτερίστικη διαίσθηση της ανθρώπινης αναζήτησης. Οι ΕΑ μιμούνται τις διαδικασίες βιολογικής εξέλιξης με την υλοποίηση των ιδεών της φυσικής επιλογής και της επικράτησης του ισχυρότερου, έτσι ώστε να παρέχουν αποτελεσματικές λύσεις σε προβλήματα αναζήτησης και βελτιστοποίησης.

Υπάρχουν διάφορα εξελικτικά υπολογιστικά μοντέλα τα οποία όμως βασίζονται στις ίδιες αρχές, δηλαδή στις αρχές προσομοίωσης της εξέλιξης ατομικών δομών μέσω των διαδικασιών της επιλογής και της αναπαραγωγής. Αυτές οι διαδικασίες βασίζονται στην ποιότητα/ικανότητα των ατομικών δομών όπως ορίζονται σε κάποιο περιβάλλον.

Το κεντρικό σημείο της έρευνας στους ΕΑ υπήρξε η **ευρωστία** (robustness), η ισορροπία δηλαδή ανάμεσα στην ικανότητα επίλυσης συγκεκριμένων προβλημάτων από τη μια μεριά και στην αποτελεσματικότητα που απαιτείται για την προσαρμογή σε πολλά διαφορετικά περιβάλλοντα από την άλλη. Όσο πιο εύρωστο είναι ένα τεχνητό σύστημα, τόσο καλύτερα και για μεγαλύτερο χρονικό διάστημα αποδίδει και επιτυγχάνει υψηλότερα επίπεδα προσαρμοστικότητας.

Πιο συγκεκριμένα, οι ΕΑ διατηρούν έναν **πληθυσμό ατόμων/δομών** (population of individuals) τον οποίο εξελίσσουν σύμφωνα με κάποιους κανόνες **επιλογής** (selection rules) και κάποιους **τελεστές** (operators), όπως **ανασυνδυασμός** (recombination) και **μετάλλαξη** (mutation). Κάθε **άτομο** (individual) του πληθυσμού αντιπροσωπεύει ένα σημείο του χώρου των πιθανών λύσεων ενός συγκεκριμένου προβλήματος. Επίσης περιέχει και κάποια γνώση

για τους κανόνες του περιβάλλοντος του προβλήματος. Σε κάθε άτομο του πληθυσμού αντιστοιχείται ένα μέτρο της **ποιότητας** (fitness) που διαθέτει στο συγκεκριμένο περιβάλλον του προβλήματος το οποίο αντιμετωπίζεται και το οποίο τυποποιείται/κωδικοποιείται μέσω κάποιας **συνάρτησης ποιότητας** (fitness function). Κατά την επιλογή η προσοχή εστιάζεται σε άτομα υψηλής ποιότητας αξιοποιώντας τη διαθέσιμη πληροφορία μέσω της ποιότητας των ατόμων. Ο ανασυνδυασμός και η μετάλλαξη διαταράσσουν τη δομή των ατόμων παρέχοντας δυνατότητες διερεύνησης του χώρου. Οι ΕΑ είναι αρκετά πολύπλοκοι έτσι ώστε να παρέχουν εύρωστους και αποτελεσματικούς μηχανισμούς αναζήτησης αν και φαίνονται πολύ απλοϊκοί από την πλευρά ενός βιολόγου.

Ο αρχικός πληθυσμός ενός ΕΑ συνήθως αρχικοποιείται σε τυχαίες τιμές και εξελίσσεται προς διαδοχικά καλύτερες περιοχές του χώρου αναζήτησης μέσω των προαναφερθέντων (λίγο ή πολύ) τυχαίων διαδικασιών της επιλογής, του ανασυνδυασμού και της μετάλλαξης. Το περιβάλλον αποδίδει πληροφορίες σχετικά με την **ποιότητα** (fitness value) των νέων σημείων αναζήτησης, και η διαδικασία επιλογής ευνοεί τα άτομα με καλύτερη ποιότητα να αναπαράγονται συχνότερα από τα άλλα άτομα του πληθυσμού. Ο μηχανισμός ανασυνδυασμού επιτρέπει τη μείξη της πληροφορίας που μεταφέρουν οι γονείς στους απογόνους και η μετάλλαξη εισάγει νέα στοιχεία, καινοτομίες στον πληθυσμό.

Οι κύριοι αντιπρόσωποι αυτού του υπολογιστικού μοντέλου περιλαμβάνουν τους **Γενετικούς Αλγόριθμους** (Genetic Algorithms) (Holland 1975, Goldberg 1989), τις **Εξελικτικές Στρατηγικές** (Evolutionary Strategies) (Schwefel 1981, Schwefel 1995) και τον **Εξελικτικό Προγραμματισμό** (Evolutionary Programming) (Fogel et al 1966, Fogel 1991) οι οποίες παρουσιάζονται παρακάτω σε επόμενη ενότητα (2.4).

## 2.2 Βιολογικό Υπόβαθρο

Οι εξελικτικοί αλγόριθμοι βασίζονται στο μοντέλο φυσικής, βιολογικής εξέλιξης το οποίο προτάθηκε από τον Κάρολο Δαρβίνο στο έργο του "The Origin of Species". Η θεωρία της εξέλιξης του Δαρβίνου εξηγεί την προσαρμοστική αλλαγή των ειδών μέσω της αρχής της φυσικής επιλογής, η οποία ευνοεί την επιβίωση και την περαιτέρω εξέλιξη εκείνων των ειδών που είναι καλύτερα προσαρμοσμένα στις περιβαλλοντικές τους συνθήκες (η λεγόμενη "επιβίωση του ικανότερου"). Εκτός από την επιλογή, ο άλλος σημαντικός παράγοντας που αναγνωρίζει ο Δαρβίνος για την εξέλιξη των ειδών, είναι η ύπαρξη μικρών, προφανώς τυχαίων και έμμεσων αποκλίσεων (μεταλλάξεων) ανάμεσα στους **φαινοτύπους** (phenotypes) των οργανισμών, δηλαδή τα φυσικά και πνευματικά χαρακτηριστικά όπως χρώμα ματιών, ύψος, μέγεθος εγκέφαλου, ευφυία κλπ, τα οποία καθορίζουν και τον τρόπο ανταπόκρισης και φυσικής ενσάρκωσης των γονέων και των παιδιών τους. Οι μεταλλάξεις αυτές υπερισχύουν μέσα από την επιλογή, εάν αποδείξουν την αξία τους στις συνθήκες του παρόντος περιβάλλοντος· διαφορετικά χάνονται. Η βασική κινητήρια δύναμη της επιλογής δίνεται από τη φυσική διαδικασία της αναπαραγωγής απογόνων. Υπό ευνοϊκές περιβαλλοντικές συνθήκες, το μέγεθος του πληθυσμού αυξάνεται εκθετικά, μία διαδικασία η οποία περιορίζεται από τους πεπερασμένους διαθέσιμους πόρους.

Όταν οι πόροι δεν επαρκούν για να στηρίξουν όλα τα άτομα ενός πληθυσμού, τότε ευνοούνται εκείνοι οι οργανισμοί οι οποίοι εκμεταλλεύονται πιο αποτελεσματικά τους διαθέσιμους πόρους.

Όλοι οι ζώντες οργανισμοί αποτελούνται από κύτταρα και κάθε κύτταρο περιέχει το ίδιο σύνολο από ένα ή περισσότερα **χρωμοσώματα** (chromosomes) τα οποία αποτελούν ακολουθίες DNA οι οποίες λειτουργούν ως προσχέδιο ανάπτυξης του οργανισμού. Ένα χρωμόσωμα διαιρείται σε **γονίδια** (genes) λειτουργικά τμήματα του DNA, κάθε ένα από τα οποία κωδικοποιεί μία συγκεκριμένη πρωτεΐνη. Τα γονίδια θεωρούνται οι μονάδες μεταβίβασης του συνόλου των κληρονομικών χαρακτηριστικών. Χονδρικά, μπορεί να θεωρηθεί ότι κάθε γονίδιο κωδικοποιεί ένα χαρακτηριστικό, όπως το χρώμα των ματιών. Οι διαφορετικές πιθανές τιμές ενός γονιδίου ονομάζονται **τιμές χαρακτηριστικών** - alleles (π.χ: καστανά, μπλέ, πράσινα, κτλ). Κάθε γονίδιο έχει μία συγκεκριμένη θέση μέσα στο χρωμόσωμα. Τα γονίδια μεταβάλλονται περιστασιακά μέσω μεταλλάξεων.

Πολλοί οργανισμοί έχουν πολλαπλά χρωμοσώματα σε κάθε κύτταρο. Το πλήρες σύνολο όλου του γενετικού υλικού (όλα τα χρωμοσώματα μαζί) αποτελούν το **γονιδίωμα** (genome) του οργανισμού. Ο όρος γονιδίωμα αναφέρεται στο σύνολο των γονιδίων τα οποία περιέχονται σε αυτό. Δύο άτομα τα οποία έχουν πανομοιότυπο γονιδίωμα λέγεται ότι έχουν τον ίδιο **γονότυπο** (genotype). Ο γονότυπος έχει σαν αποτέλεσμα, μέσω της εμβρυϊκής και της μετέπειτα ανάπτυξης, την εμφάνιση του συγκεκριμένου φαινοτύπου του οργανισμού.

Η επιλογή ενεργεί επί των ατόμων (τα άτομα είναι οι μονάδες επιλογής) τα οποία μέσω του φαινοτύπου τους εκφράζουν τις πολύπλοκες αλληλεπιδράσεις εντός του γονότυπού τους, δηλαδή τη συνολική γενετική πληροφορία του οργανισμού, καθώς επίσης και την αλληλεπίδραση του γονότυπου με το περιβάλλον κατά τον καθορισμό του φαινοτύπου.

Η **ποιότητα** (fitness) ενός οργανισμού τυπικά ορίζεται ως η πιθανότητα βιωσιμότητας (viability) του οργανισμού (η πιθανότητα να επιβιώσει και να αναπαραγάγει απογόνους) ή ως συνάρτηση του αριθμού των απογόνων του (**γονιμότητα** – fertility). Στο πλαίσιο της εξέλιξης, η ποιότητα ενός ατόμου μετρείται μόνο έμμεσα μέσω του ρυθμού αύξησής του σε σύγκριση με τα άλλα άτομα του πληθυσμού δηλ. με την τάση του να επιβιώσει και να αναπαραγάγει σε ένα συγκεκριμένο περιβάλλον. Επιπλέον η φυσική επιλογή δεν είναι η ενεργή κινητήρια δύναμη. Η επιλογή είναι το όνομα που χρησιμοποιείται για να περιγράψει την ικανότητα των ατόμων τα οποία έχουν καταφέρει να επιβιώσουν και να μεταφέρουν το γενετικό υλικό τους στην επόμενη γενιά.

## 2.3 Ένας τυπικός ΕΑ

Το παρακάτω σχήμα (2.1) περιγράφει έναν τυπικό ΕΑ. Ένας πληθυσμός δομών αρχικοποιείται και κατόπιν εξελίσσεται από γενιά σε γενιά με την εφαρμογή της επιλογής, του ανασυνδυασμού και της μετάλλαξης. Το μέγεθος του πληθυσμού είναι γενικά σταθερό στους ΕΑ, αν και δεν υπάρχει συγκεκριμένος λόγος (εκτός της προγραμματιστικής ευκολίας) να μένει σταθερό.

Τυπικά, ένας εξελικτικός αλγόριθμος αρχικοποιεί τον πληθυσμό του σε τυχαίες τιμές, αν και μπορεί να χρησιμοποιηθεί προηγούμενη γνώση του πεδίου εφαρμογής (εάν υπάρχει) για να επηρεάσει την αρχικοποίηση του πληθυσμού. Ακολουθεί η αξιολόγηση (evaluation) του πληθυσμού αποδίδοντας αντίστοιχες τιμές ποιότητας (fitness) σε κάθε άτομο του πληθυσμού στο συγκεκριμένο περιβάλλον. Η αξιολόγηση γίνεται μέσω της συνάρτησης ποιότητας (αντιπροσωπευτική του συγκεκριμένου περιβάλλοντος) η οποία μπορεί να είναι πολύ απλή, όπως ο υπολογισμός μιας απλής συνάρτησης, ή εξαιρετικά πολύπλοκη, όπως η εκτέλεση μιας πολύπλοκης προσομοίωσης. Η επιλογή (selection) συνήθως υλοποιείται σε δύο βήματα, στην επιλογή γονέων και στην επιβίωση απογόνων. Κατά την επιλογή των γονέων καθορίζεται ποιά άτομα θα γίνουν γονείς και πόσους απογόνους / παιδιά (offsprings/children) θα αποκτήσουν.

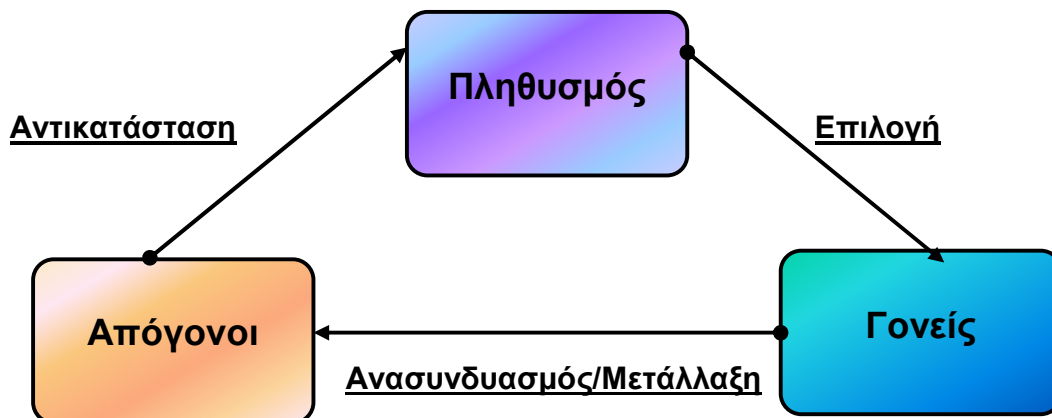
```

Procedure EA {
  t = 0;
  Initialize_Population_P (t);
  Evaluate_P (t);
  Until (done) {
    t = t + 1;
    Parent_Selection_P (t);
    Recombine_P (t);
    Mutate_P (t);
    Evaluate_P (t);
    Survive_P (t);
  }
}

```

**Σχήμα 2.1.** Ένας τυπικός εξελικτικός αλγόριθμος.

Οι απόγονοι δημιουργούνται μέσω ανασυνδυασμού των γονέων, δηλαδή με την ανταλλαγή πληροφορίας μεταξύ των γονέων και μέσω μετάλλαξης η οποία διαταράσσει περαιτέρω τους απογόνους. Ακολουθεί η χρήση της συνάρτησης ποιότητας για την αξιολόγηση των απογόνων και τελικά η επιλογή των ατόμων του πληθυσμού που θα επιβιώσουν στην επόμενη γενιά. Η διαδικασία αυτή ονομάζεται **εξελικτικός κύκλος** (evolutionary cycle). Το σχήμα 2.2 εμφανίζει τον συνήθη εξελικτικό κύκλο.



Σχήμα 2.2. Εξελικτικός Κύκλος.

## 2.4 Βασικά Μοντέλα Εξελικτικών Αλγορίθμων

Στο χώρο των ΕΑ έχουν κυριαρχήσει τρεις κυρίως μεθοδολογίες (κατηγορίες ΕΑ) όπως έχει ήδη αναφερθεί. Οι μεθοδολογίες αυτές περιλαμβάνουν τους **Γενετικούς Αλγόριθμους** (Genetic Algorithms) [8], τις **Εξελικτικές Στρατηγικές** (Evolutionary Strategies) και τον **Εξελικτικό Προγραμματισμό** (Evolutionary Programming). Αν και οι τρεις αυτές μεθοδολογίες στηρίζονται σε όμοιες αρχές, κάθε μία υλοποιείται με διαφορετικό τρόπο. Οι διαφορές αγγίζουν σχεδόν όλα τα θέματα υλοποίησης συμπεριλαμβάνοντας τις μεθόδους αναπαράστασης των ατόμων, τους μηχανισμούς επιλογής, τους τύπους των γενετικών τελεστών και την μέτρηση της απόδοσης.

Παρόλο που και οι τρεις μεθοδολογίες είναι γνωστές περισσότερο από 30 χρόνια, οι Γενετικοί Αλγόριθμοι έχουν κερδίσει περισσότερο ενδιαφέρον από τις Εξελικτικές Στρατηγικές και τον Εξελικτικό Προγραμματισμό.

Οι μεθοδολογίες αυτές έχουν δώσει την έμπνευση για την ανάπτυξη επιπρόσθετων μεθοδολογιών όπως τον "Γενετικό Προγραμματισμό - Genetic Programming" (Koza, 1991) τα "Συστήματα Κατάταξης - Classifier Systems" (Holland, 1986), τα "Συστήματα LS - LS systems" (Smith, 1983), κ.α.

### 2.4.1 Γενετικοί Αλγόριθμοι

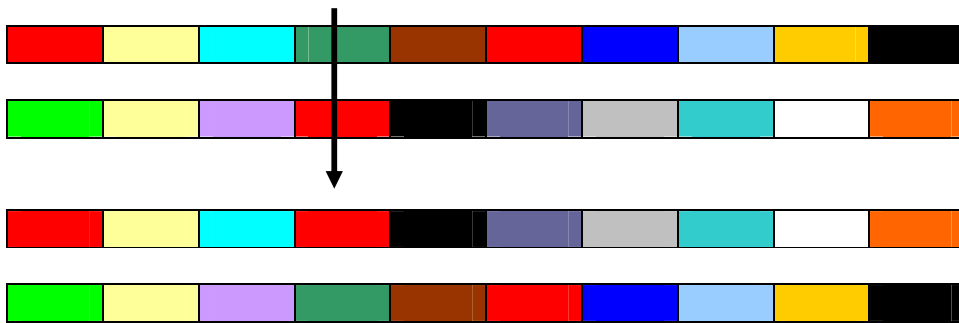
Η ανάπτυξή τους ξεκίνησε στη δεκαετία του 1960 από τον John Holland, τους συνεργάτες του και τους φοιτητές του, στο Πανεπιστήμιο του Michigan. Οι σκοποί της έρευνάς τους είχαν διπλή κατεύθυνση:

- να συνοψίσουν και να εξηγήσουν αυστηρά τις προσαρμοστικές και αναπαραγωγικές διαδικασίες των φυσικών συστημάτων, και
- να σχεδιάσουν λογισμικό τεχνητών συστημάτων που να διατηρεί τους πιο σημαντικούς από τους μηχανισμούς των φυσικών συστημάτων.

Παραδοσιακά, οι Γενετικοί Αλγόριθμοι (ΓΑ) χρησιμοποιούν μία αναπαράσταση η οποία είναι ανεξάρτητη από το πρόβλημα, δηλαδή συμβολοσειρές δυαδικών ψηφίων. Ωστόσο, αρκετές πρόσφατες εφαρμογές χρησιμοποιούν άλλους τρόπους αναπαράστασης όπως γράφοι, εκφράσεις LISP, διατεταγμένες λίστες και διανύσματα πραγματικών αριθμών.

Κατά την αρχικοποίηση του πληθυσμού δημιουργείται ένα σύνολο από δυαδικές συμβολοσειρές (χρωμοσώματα). Μετά την αρχικοποίηση επιλέγονται οι γονείς σύμφωνα με μία συνάρτηση πιθανότητας η οποία βασίζεται στην σχετική ποιότητα των ατόμων του πληθυσμού. Με άλλα λόγια, τα άτομα (χρωμοσώματα) με καλύτερη ποιότητα έχουν περισσότερες πιθανότητες να επιλεγούν ως γονείς. Όσο καλύτερη είναι η ποιότητα ενός ατόμου, τόσο αυξάνονται οι πιθανότητες να επιλεγεί περισσότερες φορές σαν γονέας για την αναπαραγωγή παιδιών. Γενικά, από  $N$  γονείς αναπαράγονται  $N$  παιδιά μέσω **διασταύρωσης** (crossover), όπως ονομάζεται ο ανασυνδυασμός στην περίπτωση των ΓΑ. Τυπικά, ακολουθεί η μετάλλαξη των  $N$  παιδιών σύμφωνα με κάποιο συντελεστή πιθανότητας μετάλλαξης και η επιβίωση των παιδιών αντικαθιστώντας τους  $N$  γονείς του πληθυσμού και δημιουργώντας μία νέα γενιά.

Ο τελεστής διασταύρωσης λειτουργεί ως εξής: Επιλέγεται τυχαία μία θέση του χρωμοσώματος και ανταλλάσσονται τα τμήματα πριν και μετά την θέση αυτή μεταξύ των δύο χρωμοσωμάτων όπως φαίνεται στο σχήμα 2.3. Για παράδειγμα, εάν οι συμβολοσειρές 10000100 και 11111111 διασταυρωθούν μετά την τρίτη θέση θα παραχθούν τα δύο παιδιά 10011111 και 11100100.



**Σχήμα 2.3.** Διασταύρωση ενός σημείου, από την τέταρτη θέση.

Ο τελεστής διασταύρωσης μιμείται χονδρικά τον βιολογικό ανασυνδυασμό μεταξύ δύο οργανισμών με μονό χρωμόσωμα (απλοειδείς). Μέσω της διασταύρωσης, οι ΓΑ εκμεταλλεύονται αποτελεσματικά ιστορικές πληροφορίες για να κάνουν υποθέσεις πάνω σε νέα σημεία έρευνας, με προσδοκώμενη βελτιωμένη απόδοση.

Ο τελεστής μετάλλαξης αλλάζει τυχαία κάποια από τα δυαδικά ψηφία ενός χρωμοσώματος, μετατρέποντάς τα από 0 σε 1 ή αντίστροφα όπως φαίνεται στο σχήμα 2.4. Για παράδειγμα, εάν στην συμβολοσειρά 00000100 γίνει μετάλλαξη στο δεύτερο δυαδικό ψηφίο της θα γίνει 01000100. Η μετάλλαξη μπορεί να συμβεί σε οποιαδήποτε θέση μιας συμβολοσειράς με κάποια πιθανότητα, συνήθως πολύ μικρή όσον αφορά τους ΓΑ (π.χ. 0.001)



Σχήμα 2.4. Μετάλλαξη στη δεύτερη θέση του χρωμοσώματος.

Είναι αρκετά σημαντικό να σημειώσουμε ότι στους ΓΑ η έμφαση δίνεται στον τελεστή ανασυνδυασμού και όχι στον τελεστή μετάλλαξης. Όπως ήδη ειπώθηκε η πιθανότητα μετάλλαξης (δηλ. αντιστροφής) των δυαδικών ψηφίων είναι πολύ μικρή και συχνά θεωρείται τελεστής που λειτουργεί στο παρασκήνιο. Ο ανασυνδυασμός, από την άλλη, θεωρείται ως ο κύριος τελεστής διερεύνησης.

### 2.4.2 Εξελικτικές Στρατηγικές

Οι Εξελικτικές Στρατηγικές (ΕΣ) αρχικά αναπτύχθηκαν το 1964 στο Technical University of Berlin (TUB) από τους Rechenberg και Schwefel ως μια πειραματική τεχνική βελτιστοποίησης. Οι πρώτες εφαρμογές είχαν σχέση με προβλήματα βελτιστοποίησης παραμέτρων όπως προβλήματα υδροδυναμικής (σχεδίαση της καμπής εύκαμπτων σωλήνων οι οποίοι διαρρέονται από κάποιο αέριο, έτσι ώστε να είναι ελάχιστη η απώλεια ενέργειας). Αυτή η πρώτη έκδοση δουλεύει χρησιμοποιώντας μόνο δύο άτομα (**δυμελής ΕΣ** - two membered ES), δηλαδή ένα γονέα και έναν απόγονο ανά γενιά. Ο απόγονος δημιουργείται εφαρμόζοντας διωνυμικές κατανομές (με αναμενόμενη τιμή μηδέν και διασπορά  $\sigma^2$ ) στον γονέα και είτε ο απόγονος γίνεται ο γονέας της επόμενης γενιάς (εάν είναι καλύτερος του γονέα), είτε ο γονέας "επιβιώνει". Πιο συγκεκριμένα, οι ΕΣ δουλεύουν με πίνακες πραγματικών διανυσμάτων. Στην περίπτωση των δυμελών ΕΣ ένα άτομο δημιουργείται από ένα μόνο γονέα μέσω της πρόσθεσης κανονικά κατανομημένων τυχαίων διανυσμάτων με αναμενόμενη τιμή μηδέν και τυπική απόκλιση 'σ' (το ίδιο 'σ' χρησιμοποιείται για όλα τα στοιχεία του διανύσματος). Το άτομο με την καλύτερη ποιότητα χρησιμοποιείται ως γονέας για την επόμενη γενιά. Αυτός ο αλγόριθμος ονομάζεται (1+1)-ES, δείχνοντας ότι επιλέγεται το καλύτερο άτομο προς επιβίωση, από ένα γονέα και ένα απόγονο. Η εξελικτική στρατηγική (1+1)-ES σύντομα χρησιμοποιήθηκε και για συνεχείς μεταβλητές (με κανονικές κατανομές), οι οποίες αποτελούν και το κυρίως μέρος των εφαρμογών σήμερα.

Αργότερα η (1+1)-ES αντικαταστάθηκε, κυρίως σε υπολογιστικές εφαρμογές, από παραλλαγές με περισσότερους από έναν γονείς ( $\mu > 1$ ) και περισσότερους από έναν απογόνους ( $\lambda > 1$ ) ανά γενιά, δηλαδή ( $\mu + \lambda$ )-ES.

Η σημειογραφία ( $\mu + \lambda$ )-ES σημαίνει ότι 'μ' γονείς δημιουργούν  $\lambda \geq 1$  απογόνους μέσω ανασυνδυασμού και μετάλλαξης. Τα 'μ' καλύτερα άτομα από τον συνολικό πληθυσμό των 'μ' γονέων και 'λ' απογόνων επιλέγονται για να σχηματίσουν την επόμενη γενιά.



### 2.4.3 Εξελικτικός Προγραμματισμός

Οι ΓΑ και οι ΕΣ αποτελούν τις δύο πιο συχνά χρησιμοποιούμενες και καλύτερα κατανοημένες μεθοδολογίες. Ο Εξελικτικός Προγραμματισμός (ΕΠ) είναι η πιο σπάνια χρησιμοποιούμενη μεθοδολογία μεταξύ των τριών βασικών μεθοδολογιών Εξελικτικών Αλγόριθμων.

Ο ΕΠ αναπτύχθηκε από τους Fogel, Owens και Walsh (1966). Παραδοσιακά χρησιμοποιεί αναπαραστάσεις προσαρμοσμένες στο πρόβλημα. Για παράδειγμα, σε προβλήματα βελτιστοποίησης πραγματικών αριθμών, τα άτομα είναι διανύσματα πραγματικών αριθμών, ενώ για το πρόβλημα του περιοδεύοντος πωλητή (traveling salesperson problem) χρησιμοποιούνται διατεταγμένες λίστες και για εφαρμογές μηχανών πεπερασμένων καταστάσεων χρησιμοποιούνται γράφοι.

Ο ΕΠ συχνά χρησιμοποιείται ως μέσο βελτιστοποίησης, αν και από την αρχή δημιουργήθηκε με στόχο την επίτευξη ευφυούς συμπεριφοράς μέσω της προσομοίωσης της φυσικής εξέλιξης. Ο D. Fogel ορίζει την νοημοσύνη ως την «ικανότητα ενός συστήματος να προσαρμόσει την συμπεριφορά του για να πετύχει τους στόχους του σε διάφορα περιβάλλοντα», διευκρινίζοντας πως μπορεί να επιτευχθεί αυτό χρησιμοποιώντας σαν βάση την προσομοίωση της φυσικής εξέλιξης. Ενώ η αρχική μορφή του ΕΠ προτάθηκε για να λειτουργεί σε μηχανές πεπερασμένων καταστάσεων και τις αντίστοιχες διακριτές αναπαραστάσεις, οι περισσότερες εφαρμογές του ΕΠ αναφέρονται σε προβλήματα βελτιστοποίησης συνεχών μεταβλητών.

Το μοντέλο ΕΠ, όπως υλοποιήθηκε από τον Fogel, δούλευε με πληθυσμό  $\mu > 1$  ατόμων τα οποία δημιουργούν 'μ' απογόνους μέσω μετάλλαξης του κάθε γονέα. Η μετάλλαξη υλοποιείται ως μία τυχαία αλλαγή της περιγραφής της μηχανής πεπερασμένων καταστάσεων σύμφωνα με πέντε διαφορετικές τροποποιήσεις:

- αλλαγή ενός συμβόλου εξόδου
- αλλαγή μίας κατάστασης μετάβασης
- πρόσθεση μίας κατάστασης
- διαγραφή μίας κατάστασης
- αλλαγή της αρχικής κατάστασης

Τυπικά οι μεταλλάξεις γίνονται με ομοιόμορφη κατανομή και ο αριθμός των μεταλλάξεων για ένα απόγονο είτε είναι σταθερός, είτε επιλέγεται σύμφωνα με κάποια κατανομή πιθανότητας. Μετά την αξιολόγηση των απογόνων, επιλέγονται τα 'μ' καλύτερα άτομα από το σύνολο των γονέων και των απογόνων.

Η γενική αρχή ενός αλγορίθμου μετάλλαξης – επιλογής, όπως ο ΕΠ, ο οποίος δε χρησιμοποιεί ανασυνδυασμό, δέχτηκε ισχυρή κριτική από ερευνητές οι οποίοι δουλεύουν στο χώρο των ΓΑ [33] και καταλήγουν ότι δεν είναι μία αρκετά ισχυρή μέθοδος. Όμως είναι φανερό, από διάφορα εμπειρικά και θεωρητικά αποτελέσματα ότι ο ρόλος της μετάλλαξης έχει υποτιμηθεί στο χώρο των ΓΑ για περισσότερο από 30 χρόνια, ενώ έχει υπερτιμηθεί ο ρόλος του ανασυνδυασμού.

## 2.5 Παράλληλοι Εξελικτικοί Αλγόριθμοι

Ενώ η φυσική εξέλιξη είναι μία υψηλά παράλληλη διεργασία, οι παραδοσιακοί εξελικτικοί αλγόριθμοι υλοποιούνται με σειριακό τρόπο. Όπως αναφέρει και ο D.E. Goldberg σε μία εργασία του, “Genetic Algorithms in Search, Optimization, and Machine Learning”, :

*“In a world where serial algorithms are usually made parallel through countless tricks and contortions, it is no small irony that genetic algorithms (highly parallel algorithms) are made serial through equally unnatural tricks and turns.”*

[«Σε έναν κόσμο όπου οι σειριακοί αλγόριθμοι μετατρέπονται σε παράλληλοι διαμέσου αμέτρητων τεχνασμάτων και παραμορφώσεων, δεν είναι μικρή ειρωνία ότι οι γενετικοί αλγόριθμοι (που είναι άκρως παράλληλοι αλγόριθμοι) υλοποιούνται σειριακά διαμέσου εξίσου αφύσικων τεχνασμάτων και μεταβολών»].

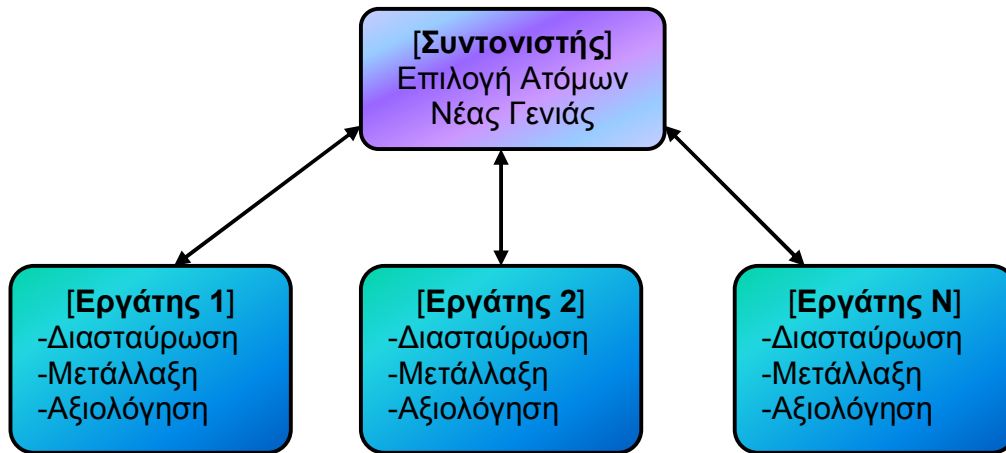
Το ενδιαφέρον για την παράλληλη φύση των εξελικτικών αλγορίθμων ξεκινάει από την εργασία του Holland. Τα τελευταία χρόνια έχουν γίνει πολλές προσπάθειες για τη δημιουργία παράλληλων μοντέλων ΕΑ, τα οποία να αντικατοπτρίζουν μερική από την πολυπλοκότητα των φυσικών συστημάτων.

Οι κύριες κατηγορίες παράλληλων ΕΑ είναι δύο και περιγράφονται με συντομία παρακάτω [1, 3, 8, 32].

### 2.5.1 Τυπική Παράλληλη Προσέγγιση

Σε αυτήν την προσέγγιση (global or standard parallelization) χρησιμοποιείται ένας παράλληλος ηλεκτρονικός υπολογιστής (H/Y) για την υλοποίηση, συνήθως, ενός Σειριακού Γενετικού Αλγόριθμου (ΣΓΑ). Παρόλο που η αξιολόγηση της ποιότητας των χρωμοσωμάτων και η εφαρμογή των γενετικών τελεστών είναι μία πλήρως παραλληλίσιμη διεργασία ο πληθυσμός παραμένει ενιαίος. Με αυτόν τον τρόπο κάθε μονάδα του πληθυσμού (χρωμόσωμα) έχει τη δυνατότητα να «ζευγαρώσει» με οποιοδήποτε άλλο μέλος του πληθυσμού. Σε αυτό λοιπόν το σημείο (του ζευγαρώματος) η τυπική παράλληλη προσέγγιση δε διαφέρει σε τίποτα από τον ΣΓΑ.

Για την υλοποίηση αυτής της προσέγγισης μπορούμε να χρησιμοποιήσουμε έναν επεξεργαστή, ο οποίος παίζει το ρόλο του συντονιστή και αναλαμβάνει την επιλογή των ατόμων που θα αποτελέσουν την επόμενη γενιά, και ένα πλήθος από άλλους επεξεργαστές οι οποίοι παίζουν το ρόλο του εργάτη και αναλαμβάνουν να εφαρμόσουν στα άτομα (που τους δίνει ο συντονιστής) όλους τους γενετικούς τελεστές και αφού δημιουργήσουν τους απογόνους και τους αξιολογήσουν επιστρέφουν πίσω τα άτομα στο συντονιστή H/Y.



Σχήμα 2.5. Σχηματική αναπαράσταση της τυπικής παράλληλης προσέγγισης.

## 2.5.2 Προσέγγιση του Διαχωρισμού

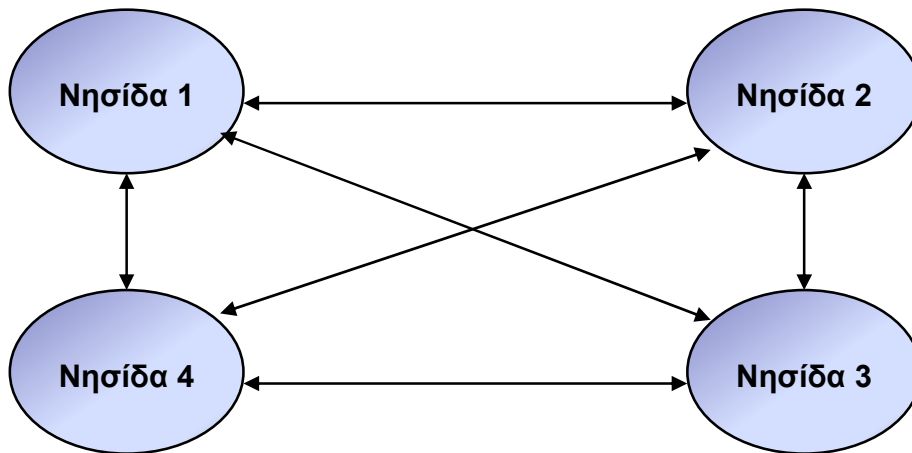
Η προσέγγιση του Διαχωρισμού (decomposition approach) αποτελεί ένα ξεχωριστό μοντέλο Εξελικτικού Αλγορίθμου και έχει σαν βασικό χαρακτηριστικό το διαχωρισμό (τμηματοποίηση) του συνολικού πληθυσμού σε μικρότερους υποπληθυσμούς. Το μέγεθος του κάθε υποπληθυσμού αποτελεί σημαντικό κριτήριο της εν λόγω προσέγγισης και τη χωρίζει σε δύο υποκατηγορίες:

- Χαμηλής Ανάλυσης (Coarse Grained)
- Υψηλής Ανάλυσης (Fine Grained)

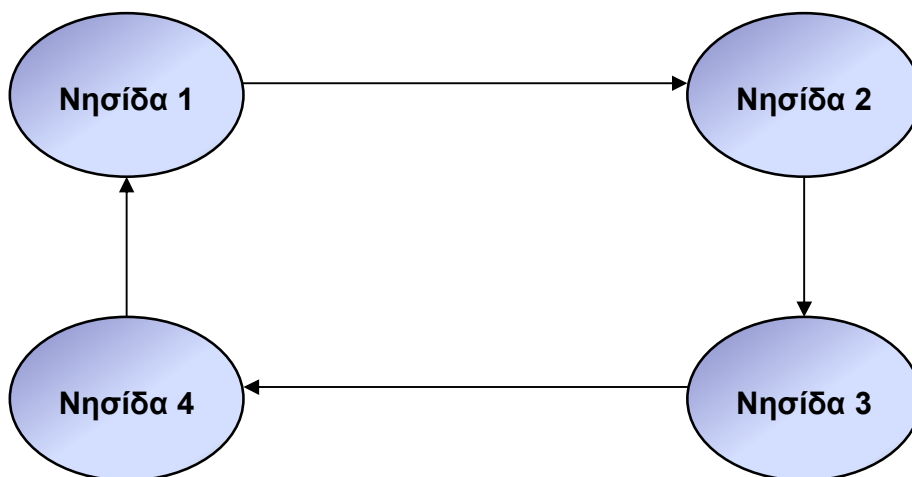
Στους παράλληλους ΓΑ χαμηλής ανάλυσης (Coarse – Grained Parallel GA) το πλήθος των υποπληθυσμών είναι, συνήθως, μικρό σε σύγκριση με αυτούς της υψηλής ανάλυσης. Κάθε υποπληθυσμός ανατίθεται σε έναν επεξεργαστή και τα άτομα που περιλαμβάνει εξελίσσονται ξεχωριστά από τα άτομα των άλλων υποπληθυσμών. Περιοδικά, σε συγκεκριμένα χρονικά διαστήματα, δίνεται η δυνατότητα στους διάφορους υποπληθυσμούς να ανταλλάξουν μεταξύ τους άτομα, συνήθως υψηλής ποιότητας. Με αυτόν τον τρόπο έχουμε ανταλλαγή γενετικού υλικού και την μετακίνηση ατόμων υψηλής ποιότητας σε όλους τους πληθυσμούς.

Οι τρόποι με τους οποίους μπορούν να μεταναστεύσουν τα άτομα μεταξύ των πληθυσμών είναι αρκετοί και εξαρτώνται από τη διασύνδεση του δικτύου των νησίδων (islands), όπου κάθε νησίδα αντιστοιχείται σε έναν υποπληθυσμό.

Έτσι έχουμε τους παρακάτω τρόπους διασύνδεσης:



Σχήμα 2.6.α. Πλήρης Διασύνδεση Νησίδων.



Σχήμα 2.6.β. Διασύνδεση Δακτυλίου.

Στην πλήρη διασύνδεση των νησίδων (βλέπε σχήμα 2.6.α) κάθε πληθυσμός είναι ελεύθερος να ανταλλάξει άτομα με όποιον πληθυσμό επιθυμεί, ακόμα και με όλους μαζί. Ενώ στη διασύνδεση δακτυλίου κάθε πληθυσμός έχει τη δυνατότητα να στείλει και να δεχθεί άτομα από συγκεκριμένους πληθυσμούς (βλέπε σχήμα 2.6.β).

Εκτός από τους παραπάνω τρόπους διασύνδεσης, που είναι οι πιο απλοί, υπάρχουν και άλλοι, πιο σύνθετοι, οι οποίοι τοποθετούν τις νησίδες σε ένα δυσδιάστατο πλέγμα (2Δ) και δημιουργούν γειτονιές. Σε αυτήν την περίπτωση τα άτομα μπορούν να μεταναστεύουν μόνο σε γειτονικούς υποπληθυσμούς.

Εκτός από τους ΠΓΑ χαμηλής ανάλυσης, υπάρχουν και αυτοί της υψηλής ανάλυσης (Fine – Grained Parallel GA). Σε αυτούς τους αλγορίθμους το πλήθος των υποπληθυσμών είναι πολύ μεγάλο με ιδανική περίπτωση αυτήν όπου κάθε χρωμόσωμα ανατίθεται σε ξεχωριστό επεξεργαστή. Εδώ τα χρωμοσώματα δεν

μπορούν να μεταναστεύουν ανεξέλεκτα σε όποιο επεξεργαστή θέλουν, αλλά μόνο σε αυτούς που ορίζεται μέσα στη γειτονιά τους. Επειδή όμως οι γειτονιές είναι επικαλυπτόμενες διασφαλίζεται το γεγονός της διάχυσης των χρωμοσωμάτων υψηλής ποιότητας σε όλο τον πληθυσμό.

Η τοπολογία του δικτύου των νησίδων, η συχνότητα μετανάστευσης των ατόμων, το πλήθος της μετακίνησης καθώς και ο τρόπος επιλογής των ατόμων που θα μεταναστεύσουν, αποτελούν πολύ σημαντικές παράμετροι στους Παράλληλους ΓΑ και καθορίζουν σημαντικά την απόδοση του αλγορίθμου και την ποιότητα των τελικών αποτελεσμάτων.

Τα πλεονεκτήματα των ΠΓΑ σε αντίθεση με τους ΣΓΑ είναι: **α)** η επιτάχυνση του υπολογιστικού χρόνου, **β)** η εύρεση καλύτερης, ποιοτικά, λύσης καθώς οι ΠΓΑ έχουν τη δυνατότητα να εξερευνούν μεγαλύτερο χώρο αναζήτησης και **γ)** μικρότερη πιθανότητα να κολλήσει η αναζήτηση σε κάποιο τοπικό βέλτιστο (φαινόμενο πρόωρης σύγκλισης). Ειδικότερα τα δύο τελευταία πλεονεκτήματα ισχύουν ακόμα και στην περίπτωση όπου ο ΠΓΑ εφαρμόζεται σε σειριακό Η/Υ και όχι σε παράλληλο.

## 2.6 Εφαρμογές Εξελικτικών Αλγορίθμων

Οι εξελικτικοί αλγόριθμοι (ΕΑ) λόγο του ότι αποτελούν πολύ ισχυρές μεθόδους αναζήτησης βρίσκουν εφαρμογή σε ένα ευρύ πεδίο επιστημών, που περιλαμβάνει επιστήμες όπως τα μαθηματικά, τη μηχανική, τη βιολογία, την ιατρική και άλλες. Μερικές σημαντικές κατηγορίες εφαρμογών των ΕΑ είναι οι εξής:

- ❖ Εύρεση μέγιστης τιμής σε αριθμητικές συναρτήσεις (πχ: συναρτήσεις πολλών μεταβλητών με ασυνέχεια και θόρυβο)
- ❖ Επεξεργασία εικόνων (πχ: αναγνώριση προτύπων σε ψηφιοποιημένες εικόνες)
- ❖ Συνδυαστική Βελτιστοποίηση [πχ: Travelling Salesman Problem (TSP), αποθήκευση κιβωτίων σε περιορισμένο χώρο (bin packing), καταμερισμός εργασιών (job-shop scheduling), κατασκευή προγραμμάτων διαλέξεων / εξετάσεων (Timetabling problems)]
- ❖ Σχεδίαση (πχ: VLSI κυκλωμάτων, γεφυρών, μηχανημάτων, σπιτιών, κτλ)
- ❖ Μηχανική μάθηση [πχ: συστήματα κατηγοριοποίησης (classifier systems), κτλ]
- ❖ κ.α.

Γίνεται λοιπόν φανερό ότι η εφαρμογές των ΕΑ καλύπτουν ένα μεγάλο φάσμα των ανθρώπινων δραστηριοτήτων. Ένα μεγάλο πλεονέκτημα των ΕΑ είναι ότι δεν απαιτούν οποιαδήποτε βοηθητική γνώση πάνω στο πρόβλημα που επιλύουν. Αυτό τους κάνει ανεξάρτητους και εφαρμόσιμους σε πολλά διαφορετικά προβλήματα. Από την άλλη μεριά όμως αυτό μπορεί να αποτελεί αδυναμία και μειονέκτημα σε προβλήματα που υπάρχει βοηθητική γνώση και μπορεί να χρησιμοποιηθεί από κάποια άλλη μέθοδο αναζήτησης / βελτιστοποίησης. Σε αυτές τις περιπτώσεις είναι, σαφώς, προτιμότερη η επιλογή κάποιας άλλης μεθόδου που θα επιλύει το πρόβλημα ευκολότερα.

Θα πρέπει, εν κατακλείδι, να επισημάνουμε ότι οι ΕΑ δεν είναι «εξαντλητικοί» (exhaustive), δηλαδή δεν εξετάζουν ολόκληρο το χώρο αναζήτησης, αλλά κατεθύνονται, με την εφαρμογή του τελεστή επιλογής, σε περιοχές του χώρου με υψηλότερη ποιότητα (exploitation). Ακόμα οι ΕΑ είναι πλήρεις (complete), γιατί πάντα βρίσκουν μία λύση, αλλά δεν είναι «αποδεκτοί» (admissible) γιατί δεν εγγυόνται ότι η λύση που θα βρουν θα είναι και η βέλτιστη του χώρου αναζήτησης.

# ΚΕΦΑΛΑΙΟ 3

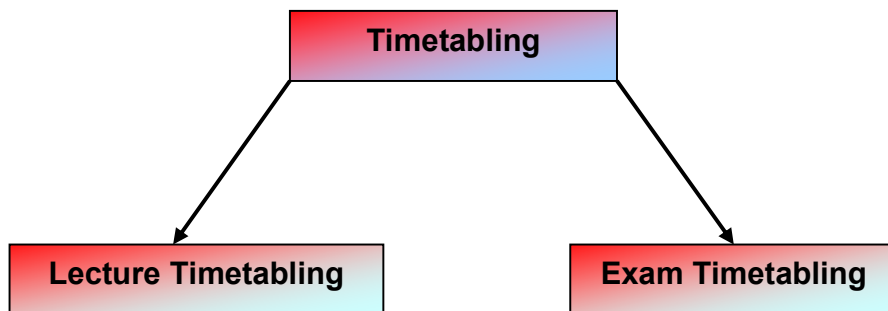
## Πρόγραμμα Εξετάσεων

Το πρόγραμμα εξετάσεων, στα ανώτερα και ανώτατα εκπαιδευτικά ιδρύματα, ανήκει σε μία γενικότερη κατηγορία προβλημάτων που ασχολείται με τον χρονοπρογραμματισμό (timetabling). Ο χρονοπρογραμματισμός ορίζεται από τον A. Wren (Wren, 1996) ως εξής:

*“Timetabling is the allocation, subject to constraints, of given resources to objects being placed in space time, in such a way as to satisfy as nearly as possible a set of desirable objectives.”*

[«Χρονοπρογραμματισμός είναι ο καταμερισμός, που υπόκειται σε περιορισμούς, πόρων σε αντικείμενα που τοποθετούνται σε ένα χρονικό διάστημα με τέτοιο τρόπο που να ικανοποιεί όσο το δυνατόν περισσότερο ένα σύνολο επιθυμητών αντικειμενικών στόχων.»]

Ο χρονοπρογραμματισμός, στο πεδίο της ανώτατης εκπαίδευσης, χωρίζεται σε δύο υποκατηγορίες : **α)** στον προγραμματισμό διαλέξεων (lecture timetabling) και **β)** στον προγραμματισμό εξετάσεων (exam timetabling).



Σχήμα 3.1. Κατηγοριοποίηση Χρονοπρογραμματισμού στην Ανώτατη Εκπαίδευση.

Και στις δύο υποκατηγορίες εμφανίζονται κοινοί περιορισμοί, όπως για παράδειγμα καμία οντότητα (πχ: φοιτητής, καθηγητής, ...) δεν μπορεί να βρίσκεται σε δύο διαφορετικά μέρη ταυτόχρονα. Από την άλλη μεριά όμως υπάρχουν σημαντικές διαφορές που φέρνουν τις δύο υποκατηγορίες σε εκ διαμέτρου αντίθετες κατευθύνσεις. Έτσι, ενώ στο χρονοπρογραμματισμό διαλέξεων, συνήθως, σε κάθε αίθουσα διεξάγεται μόνο ένα μάθημα, στο χρονοπρογραμματισμό εξετάσεων δύναται να συνυπάρχουν δύο ή και περισσότερα εξεταζόμενα μαθήματα στην ίδια αίθουσα. Επίσης ενώ στην πρώτη υποκατηγορία επιδιώκουμε να μην υπάρχουν μεγάλα κενά, χρονικά διαστήματα, μεταξύ των διδασκόμενων μαθημάτων, κάθε φοιτητή, στη δεύτερη υποκατηγορία επιδιώκουμε το ακριβώς αντίθετο, δηλαδή τα εξεταζόμενα μαθήματα κάθε φοιτητή να έχουν όσο το δυνατό μεγαλύτερο κενό, χρονικό διάστημα, δίνοντάς του έτσι περισσότερο χρόνο για μελέτη.

Η παρούσα εργασία ασχολείται με τη δεύτερη υποκατηγορία, χρονοπρογραμματισμός εξετάσεων, και περιγράφει το πρόβλημα στις ενότητες που ακολουθούν. Επίσης στο ίδιο κεφάλαιο μαζί με την περιγραφή του προβλήματος θα αναφερθούν και θα περιγραφούν οι περισσότερες μέθοδοι που έχουν εφαρμοστεί στην πράξη για την επίλυση αυτού του προβλήματος.

### **3.1 Περιγραφή του προβλήματος**

Η δημιουργία των προγραμμάτων εξεταστικής περιόδου αποτελεί μία διοικητική διαδικασία μείζονος σημασίας για όλα τα ιδρύματα της ανώτερης και ανώτατης εκπαίδευσης.

Κάθε ίδρυμα καλείται δύο ή και περισσότερες, ίσως, φορές το χρόνο να εφαρμόσει ένα πρόγραμμα σύμφωνα με το οποίο οι φοιτητές του θα εξεταστούν στα μαθήματα που έχουν δηλώσει να παρακολουθήσουν για το τρέχον ακαδημαϊκό έτος. Κάθε ίδρυμα, ανάλογα με την πολιτική που ακολουθεί πάνω στην εκπαιδευτική διαδικασία δίνει τη δυνατότητα στους φοιτητές / σπουδαστές του να επιλέγουν μαθήματα με έναν αρκετά δυναμικό τρόπο δίνοντας τους έτσι τη δυνατότητα να επιλέγουν διαφορετικό σύνολο μαθημάτων τα οποία μπορεί να προέρχονται από διαφορετικά εξάμηνα σπουδών. Φυσικά πάντα υπάρχουν περιορισμοί, οι οποίοι όμως είναι διαφορετικοί για κάθε εκπαιδευτικό ίδρυμα.

Θεωρώντας λοιπόν όλα τα προηγούμενα και συνυπολογίζοντας το γεγονός ότι κάθε ίδρυμα έχει διαφορετικούς διαθέσιμους πόρους, για την διεξαγωγή των εξετάσεων, μπορούμε να αντιληφθούμε ότι η δημιουργία ενός προγράμματος εξετάσεων είναι μία διαδικασία που είναι σε μεγάλο βαθμό εξειδικευμένη και ανταποκρίνεται στις απαιτήσεις του εκάστοτε ιδρύματος.

Η δημιουργία ενός προγράμματος εξετάσεων, με λίγα λόγια, περιλαμβάνει την ανάθεση ενός πεπερασμένου συνόλου συμβάντων (πχ: μαθήματα που πρόκειται να εξεταστούν), σε ένα, επίσης, πεπερασμένο σύνολο χρονικών περιόδων.

Η παραπάνω περιγραφή αναφέρεται σε ένα πρόγραμμα που είναι γενικά εφικτό (feasible), και δεν κάνει καμία αναφορά στην ικανοποίηση ενός ή περισσότερων περιορισμών που θα έκαναν το πρόγραμμα των εξετάσεων «καλύτερο». Σε αυτό το σημείο θα πρέπει να επισημάνουμε ότι με τον όρο «καλύτερο» εννοούμε ένα πρόγραμμα το οποίο όχι μόνο θα είναι εφικτό, αλλά θα ικανοποιεί



στο μέγιστο δυνατό βαθμό ένα σύνολο περιορισμών οι οποίοι συνήθως αποβλέπουν τόσο στην διευκόλυνση των φοιτητών κατά τη διεξαγωγή των εξετάσεων, δίνοντάς τους περισσότερο ελεύθερο χρόνο για μελέτη, όσο και στην αποτελεσματικότερη εκμετάλλευση των διαθέσιμων πόρων του ιδρύματος, όπως οι αίθουσες διεξαγωγής των εξετάσεων, οι επιβλέποντες, κτλ.

Όπως αντιλαμβανόμαστε λοιπόν η έννοια του «καλού» προγράμματος εξετάσεων μπορεί να διαφέρει πολύ από ίδρυμα σε ίδρυμα. Φυσικά υπάρχουν κάποιοι περιορισμοί οι οποίοι είναι καθιερωμένοι και σε μικρό ή μεγάλο βαθμό συμβάλλουν στη δημιουργία των προγραμμάτων κάθε ιδρύματος.

Οι περιορισμοί που συναντούμε μπορούν να ταξινομηθούν στις εξής κατηγορίες:

- **Μοναδιαίοι Περιορισμοί** (Unary Constraints). Οι μοναδιαίοι περιορισμοί περιλαμβάνουν μόνο ένα συμβάν. Έτσι, για παράδειγμα, μπορούμε να έχουμε τον εξής περιορισμό: «*Το μάθημα 'Προγραμματισμός Η/Υ 1' πρέπει να διεξαχθεί την πρώτη ημέρα των εξετάσεων*». Αυτή η κατηγορία χωρίζεται σε δύο υποκατηγορίες:
  - **Αποκλεισμού** (Exclusion). Σε αυτήν την υποκατηγορία ένα συμβάν **δεν πρέπει** να τοποθετείται σε μία συγκεκριμένη αίθουσα, ή να διεξάγεται σε μία συγκεκριμένη χρονοθυρίδα, κτλ.
  - **Προσδιορισμού** (Specification). Σε αυτήν την υποκατηγορία ένα συμβάν **πρέπει** να τοποθετείται σε μία συγκεκριμένη αίθουσα, ή να διεξάγεται σε μία συγκεκριμένη χρονοθυρίδα, κτλ.
- **Δυαδικοί Περιορισμοί** (Binary Constraints). Οι δυαδικοί περιορισμοί περιλαμβάνουν απαγορεύσεις μεταξύ δύο συμβάντων. Και αυτοί χωρίζονται σε δύο υποκατηγορίες:
  - **Ακμής** (Edge). Είναι οι πιο συχνά εμφανιζόμενοι περιορισμοί και αφορούν την ύπαρξη ενός ατόμου σε δύο μέρη την ίδια χρονική στιγμή. Ένα γενικό παράδειγμα αυτής της κατηγορίας είναι: «*Κανένας φοιτητής δεν θα πρέπει να εξετάζεται σε δύο διαφορετικά μαθήματα, ή σε δύο διαφορετικές αίθουσες την ίδια χρονική στιγμή*». Ο όρος **ακμή** (edge), προκύπτει από την παρομοίωση των προβλημάτων χρονοπρογραμματισμού με αυτά του χρωματισμού γραφημάτων (graph colouring problems).
  - **Αντιπαράθεση** (Juxtaposition). Σε αυτήν την υποκατηγορία περιλαμβάνονται περιορισμοί οι οποίοι σχετίζονται με την σειρά τοποθέτησης των συμβάντων ή με το χρονικό διάστημα που μεσολαβεί μεταξύ δύο συμβάντων. Έτσι, για παράδειγμα, έχουμε τον παρακάτω περιορισμό: «*Το μάθημα 'Α' και το μάθημα 'Β' πρέπει να εξεταστούν την ίδια χρονική στιγμή*» ή «*Το μάθημα 'Α' πρέπει να εξεταστεί δύο χρονικές περιόδους νωρίτερα από το μάθημα 'Β'*».
- **Περιορισμοί Χωρητικότητας** (Capacity Constraints). Σε αυτούς τους περιορισμούς περιγράφεται η αναγκαιότητα να μη υπερβαίνεται ένα άνω όριο στη διαθεσιμότητα κάποιων πεπερασμένων πόρων, κατά την διεξαγωγή ενός ή περισσότερων συμβάντων ταυτόχρονα. Έτσι, για

παράδειγμα, μπορούμε να πούμε ότι αν η χωρητικότητα μίας αίθουσας είναι πχ: 120 καθίσματα, να μην προγραμματίζονται μαθήματα με μεγαλύτερο πλήθος δηλώσεων.

- **Περιορισμοί διασποράς συμβάντων (Event – Spread Constraints).** Αυτοί οι περιορισμοί περιγράφουν την επιθυμία που υπάρχει στον χρονοπρογραμματισμό για την διασπορά των συμβάντων που προγραμματίζονται με συγκεκριμένο τρόπο. Έτσι, για παράδειγμα, στον χρονοπρογραμματισμό εξετάσεων είναι επιθυμητό τα μαθήματα στα οποία θα εξεταστούν οι φοιτητές να είναι όσο το δυνατόν πιο «απλωμένα» μέσα στη συνολική διάρκεια της εξεταστικής περιόδου. Δηλαδή αν κάποιος φοιτητής εξετάζεται σε πέντε μαθήματα και η συνολική χρονική διάρκεια είναι δέκα ημέρες, τότε είναι επιθυμητό ο φοιτητής να εξετάζεται κάθε δύο μέρες.

Σε προβλήματα όπου πρέπει να ικανοποιηθούν πολλοί περιορισμοί, συνήθως, τους κατηγοριοποιούμε σε «σκληρούς» (hard constraints) και σε «χαλαρούς» (soft constraints). Ο διαχωρισμός αυτός γίνεται γιατί, κατά κανόνα, είναι πολύ δύσκολο εως ανέφικτο να τους ικανοποιήσουμε όλους πλήρως.

Για το λόγο αυτό επικεντρώνουμε κυρίως την προσπάθειά μας στους σκληρούς περιορισμούς τους οποίους πρέπει να ικανοποιούμε πλήρως και στο μέγιστο βαθμό. Μάλιστα πολύ συχνά θεωρούμε ότι παραβίαση ή η μη ικανοποίηση ενός τέτοιου περιορισμού χαρακτηρίζει ένα πρόγραμμα ως ανέφικτο (infeasible)! Άλλες φορές πάλι, όταν είμαστε σίγουροι πώς με τους διαθέσιμους πόρους δεν μπορούμε να ικανοποιήσουμε πλήρως όλους τους σκληρούς περιορισμούς, γινόμαστε πιο ελαστικοί και αποδεχόμαστε την ύπαρξη ενός τέτοιου προγράμματος.

Από την άλλη μεριά υπάρχουν οι χαλαροί περιορισμοί των οποίων η ικανοποίηση είναι επιθυμητή, αλλά δεν είναι αποφασιστική. Το σύνολο αυτών των περιορισμών μπορεί να είναι πολύ μεγάλο και δίνει μία ξεχωριστή ευελιξία σε κάθε πρόγραμμα.

Παρακάτω δίνουμε μερικούς από τους περιορισμούς και των δύο κατηγοριών που συναντάμε συνήθως στη δημιουργία των προγραμμάτων εξετάσεων:

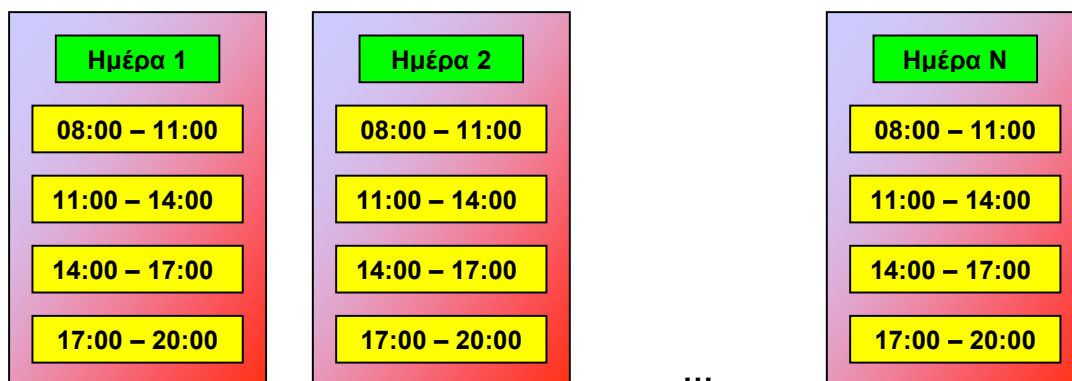
- **Σκληροί Περιορισμοί (Hard Constraints):**
  - Κανένας φοιτητής δεν θα πρέπει να βρίσκεται σε δύο διαφορετικά μέρη την ίδια χρονική στιγμή. Δηλαδή να μη δίνει ταυτόχρονα δύο ή και περισσότερα μαθήματα.
  - Κάθε μάθημα που έχει δηλωθεί για εξέταση θα πρέπει να βρίσκεται μέσα στο πρόγραμμα των εξετάσεων.
  - Δε θα πρέπει να προγραμματίζονται, σε μία αίθουσα, περισσότερα μαθήματα απ'όσα μπορεί να χωρέσει για κάθε χρονική περίοδο.
  - Το σύνολο των εξεταζομένων φοιτητών δε θα πρέπει να υπερβαίνει το σύνολο των διαθέσιμων καθισμάτων όλων των αιθουσών για κάθε χρονική περίοδο.

- Χαλαροί Περιορισμοί (Soft Constraints):
  - Να δίνεται η δυνατότητα σε ορισμένα μαθήματα να εξετάζονται σε συγκεκριμένες χρονικές περιόδους.
  - Να δίνεται η δυνατότητα σε ορισμένα μαθήματα να εξετάζονται σε συγκεκριμένες αίθουσες, που προφανώς θα υπάρχει απαραίτητη υλικοτεχνική υποδομή.
  - Να δίνεται η δυνατότητα σε κάποια μαθήματα να προγραμματίζονται νωρίτερα ή αργότερα από κάποια άλλα.
  - Οι φοιτητές να μην εξετάζονται σε συνεχόμενες χρονικές περιόδους ή να εξετάζονται μετά από συγκεκριμένο χρονικό διάστημα (πχ: κάθε δύο μέρες).
  - Μαθήματα με μεγάλη ζήτηση να εξετάζονται στην αρχή της εξεταστικής περιόδου ή μετά από κάποιο πιθανό κενό (πχ: μετά από αργία ή μετά από Σαβ/κο).

Όπως βλέπουμε λοιπόν, το σύνολο τόσο των σκληρών περιορισμών, όσο και των χαλαρών μπορεί να διαφέρει από ίδρυμα σε ίδρυμα. Ακόμη όμως και στην περίπτωση που έχουμε το ίδιο σύνολο περιορισμών η σημασία που θα δώσουμε στην ικανοποίηση κάθε περιορισμού, δηλαδή η βαρύτητα που θα έχει στη δημιουργία του «καλού» προγράμματος εξετάσεων μπορεί να ποικίλει.

Εκτός όμως από τους περιορισμούς, που επιθυμούμε να περιλάβουμε στη δημιουργία του προγράμματος, θα πρέπει να γνωρίζουμε και κάποιες άλλες παραμέτρους που έμμεσα εμφανίζονται στους περιορισμούς.

Καταρχήν θα πρέπει να γνωρίζουμε, εκ των προτέρων, τη συνολική διάρκεια της εξεταστικής περιόδου. Αυτή αποτελείται από το γινόμενο των χρονικών περιόδων (time slots) κάθε ημέρας επί το σύνολο των ημερών που θα διαρκέσουν οι εξετάσεις, αν υποθέσουμε ότι κάθε ημέρα εξετάσεων περιλαμβάνει το ίδιο πλήθος περιόδων. Σε αντίθετη περίπτωση η συνολική διάρκεια ορίζεται απλώς από το άθροισμα όλων των χρονικών περιόδων (time slots) όλων των εξεταζομένων ημερών.



**Σχήμα 3.2.** Ενδεικτική αναπαράσταση συνολικής χρονικής διάρκειας της εξεταστικής περιόδου. Σε κάθε μέρα αντιστοιχούμε το ίδιο σύνολο χρονοθυρίδων (time slots).

Για παράδειγμα στην παραπάνω σχηματική αναπαράσταση σε κάθε ημέρα της εξεταστικής περιόδου έχουμε τοποθετήσει τέσσερις χρονοθυρίδες, οι οποίες έχουν όλες την ίδια διάρκεια (τρεις ώρες). Αυτές είναι οι εξής:

1. 08:00 – 11:00
2. 11:00 – 14:00
3. 14:00 – 17:00
4. 17:00 – 20:00

Οπότε αν η εξεταστική περίοδος διαρκεί, πχ, 10 ημέρες ( $N = 10$ ), τότε το σύνολο των χρονοθυρίδων θα είναι 40 ( $10 \cdot 4$ ). Δηλαδή θα πρέπει να τοποθετήσουμε όλα τα συμβάντα (μαθήματα που πρόκειται να εξεταστούν) σε σαράντα χρονοθυρίδες.

Φυσικά το παραπάνω παράδειγμα είναι τυχαίο και τόσο το πλήθος των χρονοθυρίδων ανά ημέρα, όσο και η διάρκεια των χρονοθυρίδων μπορούν να διαφέρουν.

Μία άλλη παράμετρος που πρέπει να λαμβάνεται σοβαρά υπόψιν είναι η διαθέσιμη χωρητικότητα των αιθουσών, στις οποίες θα διεξαχθούν οι εξετάσεις. Σε κάθε περίπτωση θα πρέπει το σύνολο των φοιτητών που πρόκειται να εξεταστούν σε ένα ή και περισσότερα μαθήματα να μην υπερβαίνει τη συνολική διαθεσιμότητα των καθισμάτων, των αιθουσών. Δηλαδή αν μία χρονοθυρίδα έχει συνολική χωρητικότητα, πχ, 300 καθισμάτων, τότε οι φοιτητές των μαθημάτων που πρόκειται να εξεταστούν, σε αυτήν τη χρονοθυρίδα, δε θα πρέπει να υπερβαίνουν αυτό το άνω όριο.

Τέλος μία ακόμη παράμετρος που λαμβάνεται συχνά υπόψιν από τα εκπαιδευτικά ιδρύματα, κατά τη δημιουργία των προγραμμάτων εξετάσεων, είναι η τοποθέτηση των επιτηρητών καθηγητών (invigilator) στις αίθουσες διεξαγωγής των εξετάσεων. Η ανάθεση των καθηγητών στις αίθουσες είναι συνήθως δευτερευούσης σημασίας και πολλές φορές δεν παίζει κανέναν απολύτως ρόλο τόσο στη δημιουργία, όσο και στην «ποιότητα» ενός καλού προγράμματος εξετάσεων.

Όπως μπορούμε εύκολα να συμπεράνουμε από τα ανωτέρω, η δημιουργία ενός εφικτού (feasible) προγράμματος είναι μία σχετικά εύκολη διαδικασία. Η δημιουργία όμως ενός «καλού» προγράμματος είναι μία διαδικασία που είναι πολύ δύσκολη. Αυτό προκύπτει από το γεγονός ότι το σύνολο των παραμέτρων που πρέπει να συνυπολογιστούν, καθώς και το σύνολο των περιορισμών που πρέπει να ικανοποιηθούν είναι τόσο μεγάλο που μετατρέπει το σύνολο των πιθανών λύσεων (δηλ. του χώρου αναζήτησης του προβλήματος) σε εξαιρετικά μεγάλο.

Για την επίλυση αυτού του τόσο απαιτητικού προβλήματος υπάρχουν αρκετοί τρόποι και έχουν αναπτυχθεί πολλές μεθοδολογίες. Η εξειδίκευση όμως του προβλήματος, από ίδρυμα σε ίδρυμα, καθιστά την εύρεση μιάς ενιαίας τυποποιημένης μεθοδολογίας μία εξαιρετικά δύσκολη έως αδύνατη διαδικασία.

## 3.2 Επίλυση με διάφορες μεθόδους

Για την αντιμετώπιση του προβλήματος της δημιουργίας προγραμμάτων εξετάσεων, έχουν προταθεί και εφαρμοστεί πολλές μεθοδολογίες. Όσον αφορά την χρήση ή όχι ηλεκτρονικού υπολογιστή (Η/Υ) οι τρόποι επίλυσης μπορούν να κατηγοριοποιηθούν ως εξής:

- a. Χειρωνακτικοί (manually)
- b. Μερικώς αυτοματοποιημένοι (δηλαδή χειρωνακτικοί με μικρή βοήθεια κάποιου αυτοματοποιημένου εργαλείου σε Η/Υ)
- c. Πλήρως αυτοματοποιημένοι (κάνοντας χρήση κάποιου ολοκληρωμένου πληροφοριακού συστήματος)

Ο πρώτος τρόπος επίλυσης (χειρωνακτικός) είναι ο λιγότερο δαπανηρός σε χρήματα, μιάς και τη δημιουργία του προγράμματος αναλαμβάνει συνήθως κάποιο μέλος του εκπαιδευτικού προσωπικού του ιδρύματος, αλλά και ο λιγότερο αποτελεσματικός γιατί η τελική ποιότητα των προγραμμάτων απέχει, συνήθως, πολύ από την έννοια του καλού προγράμματος. Ο λόγος δεν είναι φυσικά η ανικανότητα αυτού που θα αναλάβει να φτιάξει το πρόγραμμα, αλλά κυρίως η έλλειψη χρόνου, γιατί όπως αναφέραμε και νωρίτερα ο χώρος λύσεων του προβλήματος είναι τόσο μεγάλος που θα ήταν αδύνατο για έναν άνθρωπο να βρεί μία καλή λύση σε λογικά χρονικά πλαίσια. Για την εύρεση λοιπόν μίας «καλής» λύσης με χειρωνακτικό τρόπο γίνονται κυρίως κάποιες υποχωρήσεις στην ικανοποίηση κάποιων περιορισμών και εφαρμόζεται παλαιότερη γνώση του προβλήματος από παλιότερα προγράμματα εξετάσεων.

Ο δεύτερος τρόπος, που είναι και ο πιο συχνά χρησιμοποιούμενος, κάνει χρήση κάποιου αυτοματοποιημένου εργαλείου σε Η/Υ. Η βοήθεια αυτού του εργαλείου περιλαμβάνει συνήθως τη συλλογή, την ομαδοποίηση και την επεξεργασία κάποιων δεδομένων, για να μπορέσει ο χρήστης να βγάλει κάποια χρήσιμα συμπεράσματα. Η τελική δημιουργία του προγράμματος όμως είναι πάλι ευθύνη του χρήστη.

Στην τελευταία περίπτωση κάνουμε χρήση κάποιου ολοκληρωμένου πληροφοριακού συστήματος το οποίο όχι μόνο επεξεργάζεται τα δεδομένα που του εισάγουμε, αλλά προτείνει και κάποιες λύσεις. Οι λύσεις αυτές είναι τα προγράμματα εξετάσεων που παρήγαγε κατά την επεξεργασία των δεδομένων. Με αυτόν τον τρόπο έχουμε καλύτερης ποιότητας λύσεις, καθώς με την εφαρμογή του πληροφοριακού συστήματος (ΠΣ) μπορούμε να εξερευνήσουμε καλύτερα το χώρο αναζήτησης λύσεων του προβλήματος και να ικανοποιήσουμε περισσότερους περιορισμούς, σε σύντομο χρονικό διάστημα.

Όσον αφορά τις μεθοδολογίες που έχουν χρησιμοποιηθεί, κατά καιρούς, και χρησιμοποιούνται ακόμα, αυτές μπορούν να χωριστούν στις εξής τέσσερις κατηγορίες (σύμφωνα με τους Carter and Laporte 1996 και 1998) : **1)** Ακολουθιακές Μέθοδοι (Sequential Methods), **2)** Μέθοδοι Ομαδοποίησης (Cluster Methods), **3)** Μέθοδοι Βασισμένοι στους Περιορισμούς (Constraint based Methods), **4)** Μέτα – Ευριστικοί Αλγόριθμοι (Meta – Heuristic Algorithms).

1. **Ακολουθιακές Μέθοδοι (Sequential Methods).** Αυτές οι μέθοδοι ταξινομούν τα συμβάντα (events) με τη βοήθεια ευριστικών μεθόδων που

σχετίζονται άμεσα με το συγκεκριμένο πρόβλημα που επιλύουν και κατόπιν αναθέτουν τα συμβάντα με ακολουθιακό τρόπο σε έγκυρες χρονοθυρίδες (valid time slots), έτσι ώστε κανένα συμβάν μέσα στην ίδια χρονοθυρίδα να μην έρχεται σε αντίθεση με κάποιο άλλο. Στις ακολουθιακές μεθόδους το πρόβλημα του χρονοπρογραμματισμού των εξετάσεων (exam timetabling problem) παριστάνεται ως γράφημα (graph), όπου τα συμβάντα, δηλαδή τα εξεταζόμενα μαθήματα, αναπαριστώνται σαν κορυφές (vertices) και οι συγκρούσεις (conflicts) μεταξύ των συμβάντων σαν τις ακμές (edges) του γραφήματος. Έτσι, για παράδειγμα, αν κάποιοι φοιτητές πρέπει να παρακολουθήσουν δύο συμβάντα, τότε υπάρχει μία ακμή ανάμεσα σε αυτούς τους δύο κόμβους που δηλώνει τη σύγκρουση. Ως επί τούτου η δημιουργία ενός προγράμματος εξετάσεων το οποίο δεν θα περιέχει τέτοιες συγκρούσεις μπορεί να «μοντελοποιηθεί» σαν ένα πρόβλημα χρωματισμού γραφήματος (graph coloring problem). Κάθε χρονοθυρίδα του προγράμματος αντιστοιχεί σε ένα ξεχωριστό χρώμα και οι κορυφές / κόμβοι του γραφήματος χρωματίζονται με τέτοιο τρόπο ώστε κανένας γειτονικός κόμβος να μην έχει το ίδιο χρώμα. Υπάρχει μία μεγάλη ποικιλία ευριστικών μεθόδων που χρησιμοποιούνται για την ταξινόμηση των συμβάντων που λειτουργούν, κυρίως, με μία εκτίμηση του πόσο δύσκολο είναι να τοποθετηθεί ένα συμβάν σε καλή χρονοθυρίδα [7, 11, 17, 18]. Οι ευριστικές τεχνικές που χρησιμοποιούνται συχνά είναι οι εξής:

- a. **Largest Degree First.** Συμβάντα που έχουν το μεγαλύτερο πλήθος συγκρούσεων τοποθετούνται νωρίτερα στο πρόγραμμα. Η λογική πίσω από αυτήν τη μέθοδο είναι ότι τα συμβάντα με το μεγαλύτερο πλήθος συγκρούσεων είναι δυσκολότερο να τοποθετηθούν στο πρόγραμμα γ'αυτό τα τοποθετούμε όσο το δυνατό νωρίτερα.
  - b. **Largest Weighted Degree.** Αυτή είναι μία παραλλαγή της προηγούμενης μεθόδου η οποία συνυπολογίζει σε κάθε σύγκρουση και το πλήθος των φοιτητών που σχετίζονται με αυτή.
  - c. **Saturation Degree.** Σε κάθε βήμα της δημιουργίας του προγράμματος επιλέγεται εκείνο το συμβάν με το μικρότερο, διαθέσιμο, αριθμό έγκυρων χρονοθυρίδων.
  - d. **Colour Degree.** Σε αυτή τη μέθοδο δίνεται μεγαλύτερη προτεραιότητα στα συμβάντα που έχουν το μεγαλύτερο πλήθος συγκρούσεων με συμβάντα που έχουν ήδη τοποθετηθεί μέσα στο πρόγραμμα.
2. **Μέθοδοι Ομαδοποίησης (Cluster Methods).** Σε αυτές τις μεθόδους το σύνολο των συμβάντων χωρίζεται σε μικρότερα υποσύνολα τα οποία ικανοποιούν τους σκληρούς περιορισμούς (hard constraints). Κατόπιν τα υποσύνολα ανατίθενται σε χρονοθυρίδες με τέτοιο τρόπο ώστε να ικανοποιούνται και οι χαλαροί περιορισμοί (soft constraints). Πολλές και διαφορετικές τεχνικές βελτιστοποίησης έχουν εφαρμοστεί για την ανάθεση των υποσυνόλων, των συμβάντων, στις χρονοθυρίδες. Το κύριο μειονέκτημα αυτών των μεθόδων είναι ότι οι ομαδοποιήσεις των συμβάντων γίνονται στην αρχή του αλγορίθμου χωρίς να έχουν τη δυνατότητα να

αλλάξουν και αυτό μπορεί να συντελέσει στην δημιουργία προγραμμάτων χαμηλής ποιότητας.

3. **Μέθοδοι Βασισμένοι στους Περιορισμούς (Constraint based Methods).** Σε αυτές τις μεθόδους το πρόβλημα του χρονοπρογραμματισμού (timetable problem) μοντελοποιείται ως ένα σύνολο από μεταβλητές (πχ: συμβάντα) των οποίων οι τιμές (πχ: αίθουσες, χρονοθυρίδες) πρέπει να ανατεθούν ικανοποιώντας ένα σύνολο περιορισμών. Συνήθως η ανάθεση των τιμών στις μεταβλητές γίνεται με κάποιους κανόνες. Όταν κανένας κανόνας δεν είναι εφαρμόσιμος στην τρέχουσα λύση εφαρμόζεται ένα είδος οπισθοδρόμησης μέχρις ότου βρεθεί μία λύση που να ικανοποιεί όλους τους περιορισμούς [2, 7, 31].
4. **Μέτα – Ευριστικοί Αλγόριθμοι (Meta – Heuristic Algorithms).** Τις τελευταίες δύο δεκαετίες μία ποικιλία από μέτα – ευριστικές προσεγγίσεις όπως simulated annealing, tabu search, evolutionary algorithms, hybrid algorithms και άλλες, έχουν ερευνηθεί και εφαρμοστεί πάνω σε προβλήματα χρονοπρογραμματισμού. Αυτές οι μέθοδοι ξεκινούν με μία ή περισσότερες αρχικές λύσεις (συνήθως τυχαία αρχικοποιημένες) και εφαρμόζουν διάφορες στρατηγικές αναζήτησης για την αποφυγή τοπικών βέλτιστων (local optima). Όλοι αυτοί οι αλγόριθμοι μπορούν να παράγουν λύσεις υψηλής ποιότητας αλλά συχνά εμφανίζουν υψηλό υπολογιστικό κόστος [2, 10, 14, 16, 19, 20, 26, 28, 34, 35].

Από την παραπάνω, σύντομη, περιγραφή των μεθοδολογιών που έχουν αναπτυχθεί για την επίλυση του προβλήματος της δημιουργίας προγραμμάτων εξετάσεων, βλέπουμε ότι το πλήθος τους δεν είναι μικρό. Αυτό φανερώνει το ενδιαφέρον της επιστημονικής κοινότητας να βρει λύση σε ένα πρόβλημα που απασχολεί όλα τα εκπαιδευτικά ιδρύματα της τριτοβάθμιας εκπαίδευσης.

# ΚΕΦΑΛΑΙΟ 4

## Εξελικτικοί Αλγόριθμοι και Πρόγραμμα Εξετάσεων

Όπως αναφέραμε και σε προηγούμενη ενότητα (Κεφάλαιο 2<sup>ο</sup>), οι Εξελικτικοί Αλγόριθμοι (ΕΑ) είναι αλγόριθμοι αναζήτησης – βελτιστοποίησης, γενικού σκοπού, οι οποίοι προσομοιώνουν τις λειτουργίες της φυσικής εξέλιξης και φυσικής επιλογής για την εύρεση της βέλτιστης λύσης σε έναν τεράστιο χώρο αναζήτησης με πολλαπλά τοπικά βέλτιστα. Η αναζήτηση της καλύτερης λύσης παρόλο που πραγματοποιείται με πιθανολογικούς (probabilistic) και όχι με αιτιοκρατικούς (deterministic) κανόνες δεν είναι και τόσο «τυχαία», όπως φαίνεται με μία πρώτη ματιά, αλλά καθοδηγείται σε ποιοτικά καλύτερες περιοχές του χώρου αναζήτησης με βάση τον τελεστή επιλογής που εφαρμόζεται (exploitation). Ένα από τα πλεονεκτήματα των ΕΑ είναι ότι για την επίλυση ενός προβλήματος δεν απαιτούν βοηθητικές γνώσεις για το πρόβλημα. Αυτό τους επιτρέπει να εφαρμόζονται σε ένα ευρύ φάσμα προβλημάτων όπου δεν υπάρχει βοηθητική γνώση για την επιτάχυνση της εύρεσης λύσης, ή την εύρεση καλύτερης ποιοτικά λύσης.

Η επιστημονική κοινότητα αντιλήφθηκε, γρήγορα, τα οφέλη που θα είχε η εφαρμογή των ΕΑ σε πολλά προβλήματα (κυρίως βελτιστοποίησης) και στράφηκε σε αυτούς εφαρμόζοντάς τους τόσο σε πειραματικά προβλήματα, όσο και σε πραγματικά προβλήματα (real world problems). Ένας από τους τομείς των προβλημάτων είναι και αυτός του χρονοπρογραμματισμού και πιο συγκεκριμένα (στα πλαίσια της παρούσης εργασίας) του χρονοπρογραμματισμού των εξετάσεων (exam timetabling). Αμέτρητες είναι οι εφαρμογές τόσο «καθάρων» ΕΑ, όσο και υβριδικών ΕΑ, που έγιναν στην πράξη από πολλούς επιστήμονες και από πολλά εκπαιδευτικά ιδρύματα. Μία μικρή ανασκόπηση αυτών των προσπαθειών γίνεται σε παρακάτω ενότητα (4.2).

Για την αντιμετώπιση οποιουδήποτε προβλήματος, από τους ΕΑ, είναι απαραίτητη η κωδικοποίηση των παραμέτρων του με τέτοιο τρόπο ώστε να



μπορεί να πραγματοποιηθεί η επεξεργασία τους. Τα βασικότερα στοιχεία που πρέπει να περιέχει ένας ΕΑ είναι: **1)** η αναπαράσταση μίας πιθανής λύσης του προβλήματος, **2)** ο αρχικός πληθυσμός, **3)** η συνάρτηση ποιότητας / καταλληλότητας και **4)** οι γενετικοί τελεστές που θα εφαρμοστούν. Στην ενότητα που ακολουθεί γίνεται μία προσπάθεια κωδικοποίησης του προβλήματος, το οποίο πραγματεύεται η εργασία, για την αντιμετώπισή του από τους ΕΑ.

## 4.1 Κωδικοποίηση του προβλήματος

Όπως αναφέραμε και παραπάνω για την επίλυση οποιουδήποτε προβλήματος οι ΕΑ απαιτούν μία συγκεκριμένη κωδικοποίηση των παραμέτρων του. Η κωδικοποίηση αυτή δεν είναι αυστηρή σε επίπεδο υλοποίησης, αλλά πρέπει να υπακούει σε κάποιες γενικές αρχές.

- **Αναπαράσταση μίας πιθανής λύσης.** Όπως περιγράφηκε και σε προηγούμενη ενότητα (2.1) οι ΕΑ διατηρούν έναν πληθυσμό από άτομα (δομές δεδομένων), τα οποία εξελίσσουν σύμφωνα με κάποιους κανόνες επιλογής και την εφαρμογή κάποιων γενετικών τελεστών. Τα άτομα αυτά αποτελούν τα σημεία του χώρου αναζήτησης και το κάθε ένα από αυτά αναπαριστά μία πιθανή λύση του εκάστοτε προβλήματος. Στην περίπτωση του προβλήματος χρονοπρογραμματισμού εξετάσεων (exam timetabling problem), τα άτομα αυτά αποτελούν τα προγράμματα των εξετάσεων. Η αναπαράσταση αυτών των ατόμων μπορεί να είναι είτε άμεση (direct representation), όπου κάθε άτομο αποτελεί ένα ολοκληρωμένο πρόγραμμα και περιέχει όλη την απαραίτητη πληροφορία, είτε έμμεση (indirect representation), όπου το άτομο δεν περιέχει όλη την απαραίτητη πληροφορία, αλλά υπάρχει ένα άλλο τμήμα του αλγορίθμου το οποίο αναλαμβάνει να διερμηνεύσει το άτομο και να σχηματίσει το πρόγραμμα.
  - Στην περίπτωση της άμεσης αναπαράστασης (direct representation) μπορούμε να φανταστούμε το χρωμόσωμα σαν έναν πίνακα όπου κάθε κελί του πίνακα αναπαριστά πχ: μία χρονική περίοδο (time slot). Και η τιμή που περιέχει αυτό το κελί να είναι ένα συμβάν (η εξέταση ενός μαθήματος) που έχει προγραμματιστεί σε αυτήν τη χρονική περίοδο.

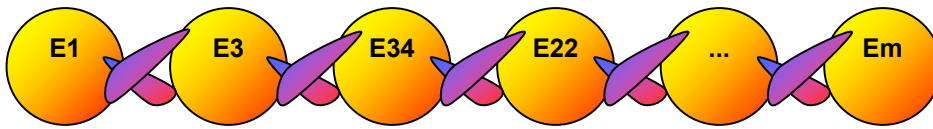
Ts: 0	Ts: 1	Ts: 2	...	Ts: n
Event: e2	Event: e43	Event: e3	...	Event: em

Σχήμα 4.1. Παράδειγμα άμεσης αναπαράστασης χρωμοσώματος.

Η ερμηνεία που θα μπορούσαμε να δώσουμε στην παραπάνω σχηματική αναπαράσταση, είναι η εξής:  
Έστω ότι έχουμε τα εξής σύνολα:

- $T_s = \{0, 1, 2, 3, \dots, n\}$ , όπου  $T_s$  είναι το σύνολο που περιέχει όλες τις δυνατές χρονοθυρίδες (time slots).
- $Event = \{e_1, e_2, e_3, \dots, e_m\}$ , όπου  $Event$  είναι το σύνολο που περιέχει όλα τα συμβάντα που πρέπει να προγραμματιστούν. Σε κάθε χρονοθυρίδα τοποθετείται ένα και μόνο ένα συμβάν. Οπότε μπορούμε να πούμε ότι: «Στη χρονοθυρίδα 0, έχει προγραμματιστεί το συμβάν  $e_2$ », κοκ.

- Στην περίπτωση της έμμεσης αναπαράστασης (indirect representation) μπορούμε να φανταστούμε το χρωμόσωμα σαν μία συνδεδεμένη λίστα, όπου κάθε κόμβος της να αναπαριστά και ένα συμβάν.



Σχήμα 4.2. Παράδειγμα έμμεσης αναπαράστασης χρωμοσώματος.

Η ερμηνεία που θα μπορούσαμε να δώσουμε στην παραπάνω σχηματική αναπαράσταση, είναι η εξής:

Έστω ότι έχουμε το σύνολο:

- $Event = \{e_1, e_2, e_3, \dots, e_m\}$ , όπου  $Event$  είναι το σύνολο που περιέχει όλα τα συμβάντα που πρέπει να προγραμματιστούν.

Οπότε μπορούμε να πούμε: «Χρησιμοποίησε την ευριστική συνάρτηση  $H$  για να προγραμματίσεις πρώτα το συμβάν  $e_1$ , χρησιμοποίησε την ευριστική συνάρτηση  $H$  για να προγραμματίσεις δεύτερο το συμβάν  $e_3$ , χρησιμοποίησε την ευριστική συνάρτηση  $H$  για να προγραμματίσεις τρίτο το συμβάν  $e_{34}$ , κοκ».

Το τί θα αποτελεί την συνάρτηση  $H$ , ποικίλει ανάλογα με το τί θέλουμε να πετύχουμε. Η συνάρτηση αυτή θα μπορούσε πχ: να τοποθετεί τα συμβάντα μέσα στις χρονοθυρίδες με τέτοιο τρόπο ώστε να ικανοποιούνται όλοι οι σκληροί περιορισμοί (hard constraints).

Η επιλογή μίας καλής αναπαράστασης επηρεάζει σε πολύ μεγάλο βαθμό την απόδοση του EA και καθορίζει τον τρόπο που θα επενεργούν πάνω στο χρωμόσωμα (άτομο) οι γενετικοί τελεστές. Επίσης ένα άλλο σημαντικό πρόβλημα το οποίο εμφανίζεται έντονα σε προβλήματα χρονοπρογραμματισμού και συνδέεται άμεσα με την αναπαράσταση του χρωμοσώματος, είναι αυτό της «επίστασης» (epistasis), ή αλλιώς η «αλληλεπίδραση» μεταξύ των γονιδίων (gene interaction). Το πρόβλημα έγκειται στο γεγονός ότι οι τιμές που δέχεται ένα γονίδιο συνδέονται σε μεγάλο βαθμό με τις τιμές κάποιου άλλου γονιδίου. Αυτό μπορεί να προκαλέσει το αρνητικό φαινόμενο κατά το οποίο η ικανοποίηση ενός περιορισμού να προκαλεί τη μη – ικανοποίηση ενός άλλου περιορισμού. Το πρόβλημα μεγαλώνει και περιπλέκεται όσο μεγαλώνει και η σχέση που πιθανόν να έχουν τα γονίδια

του χρωμοσώματος μεταξύ τους. Η ιδανική περίπτωση είναι αυτή κατά την οποία τα γονίδια είναι τελείως ανεξάρτητα μεταξύ τους και οι τιμές που δέχεται το κάθε ένα δεν εξαρτώνται και δεν επηρεάζουν τις τιμές άλλων γονιδίων. Παρά, λοιπόν, το γεγονός της σπουδαιότητας της αναπαράστασης στη συνολική λειτουργία του ΕΑ δεν υπάρχει τρόπος να ορίσουμε θεωρητικά το τί αποτελεί μία «καλή» αναπαράσταση. Οι περισσότερες αναπαραστάσεις επιλέγονται με βάση την εμπειρία και κάποια πειραματικά αποτελέσματα, παρά με κάποια θεωρητική τεκμηρίωση.

- **Αρχικός πληθυσμός.** Ο αρχικός πληθυσμός αποτελεί το σύνολο των πιθανών λύσεων του προβλήματος. Κάθε χρωμόσωμα του πληθυσμού αναπαριστά άμεσα ένα σημείο του χώρου αναζήτησης (search space). Όσο μεγαλύτερος είναι ο χώρος αναζήτησης τόσο αυξάνει η πιθανότητα να βρούμε μία καλύτερη λύση. Από την άλλη μεριά όμως αυξάνουν και οι απαιτήσεις σε υπολογιστική ισχύ, γιατί όσο μεγαλώνει ο πληθυσμός τόσο επιβραδύνεται και η διαδικασία εξέλιξης. Ο αρχικός πληθυσμός αρχικοποιείται συνήθως με τυχαίες τιμές, αν και μπορεί να χρησιμοποιηθεί παλαιότερη γνώση, αν υπάρχει. Για λόγους απλότητας και ευκολίας, συνήθως, το μέγεθος του αρχικού πληθυσμού παραμένει σταθερό κατά τη διάρκεια της εξέλιξης, αν και αυτό παραβαίνει το μοντέλο της φύσης όπου υπό φυσιολογικές συνθήκες οι πληθυσμοί τείνουν να αυξάνουν το μεγεθός τους.
- **Συνάρτηση Ποιότητας.** Η συνάρτηση ποιότητας είναι αυτή που καθορίζει το πόσο καλά είναι προσαρμοσμένο ένα άτομο στο περιβάλλον στο οποίο εξελίσσεται. Μέσα στη συνάρτηση ποιότητας μπορούμε να ενσωματώσουμε τόσο τους χαλαρούς (soft), όσο και τους σκληρούς (hard) περιορισμούς. Η συνάρτηση ποιότητας μπορεί να είναι όσο απλή, ή όσο πολύπλοκη θέλουμε εμείς. Συνήθως είναι της μορφής:

$$F(x) = \frac{1}{1 + \sum_{i=1}^n (w_i * c_i(x))}$$

όπου “**x**” είναι το χρωμόσωμα που

αξιολογείται, “**n**” είναι το πλήθος των περιορισμών, “**w<sub>i</sub>**” είναι η βαρύτητα του περιορισμού “**i**” και “**c<sub>i</sub>**” είναι η ποινή του περιορισμού “**i**”.

Η μορφή των περιορισμών “**i**” στην περίπτωση του χρονοπρογραμματισμού εξετάσεων, καθώς και απόδοσης της ποινής, μπορεί να γίνει ως εξής:

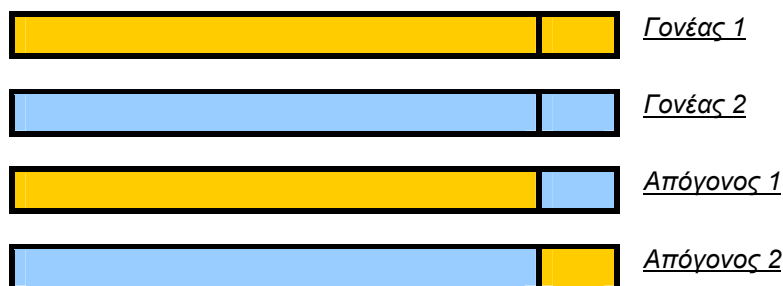
- Έστω ο περιορισμός “**i**” με τη λεκτική περιγραφή: “*αν ο φοιτητής εξετάζεται σε δύο συνεχόμενες χρονοθυρίδες, τότε πρόσθεσε στη συνολική ποινή τη μονάδα (1), αλλιώς το μηδέν (0)*”.

Τότε στη συνολική ποινή (  $\sum$  ) θα προστεθεί το γινόμενο “**1\*w<sub>i</sub>**” ή “**0\*w<sub>i</sub>**”, όπου “**w<sub>i</sub>**” είναι η βαρύτητα του περιορισμού “**i**”.

Η αξιολόγηση του πληθυσμού κατά την διάρκεια της εξέλιξης είναι από τα πιο χρονοβόρα βήματα του αλγορίθμου, ανάλογα φυσικά με την πολυπλοκότητα της συνάρτησης. Η τιμή αυτής της συνάρτησης αποτελεί την

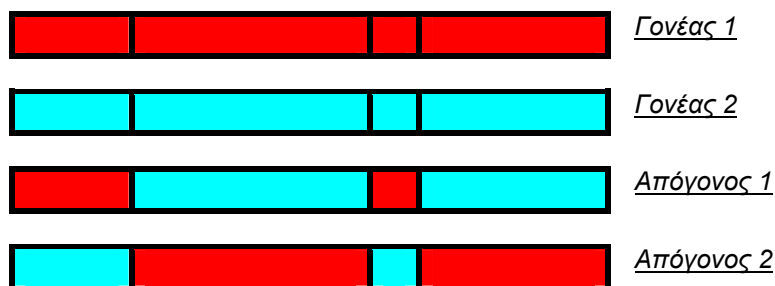
ποιότητα του χρωμοσώματος και συνήθως όσο μεγαλύτερη είναι, δηλαδή όσο λιγότερους περιορισμούς παραβιάζει το χρωμόσωμα, τόσο αυξάνει η πιθανότητα του ατόμου να ζευγαρώσει περισσότερες φορές με άλλα άτομα και να παράγει περισσότερους απογόνους.

- **Γενετικοί Τελεστές.** Σε αυτό το τμήμα του ΕΑ περιέχονται όλοι οι τελεστές οι οποίοι αναλαμβάνουν να εξελίξουν τα χρωμοσώματα του αρχικού πληθυσμού και να τα μετατρέψουν σε μονάδες πιο προσαρμοσμένες μέσα στο περιβάλλον το οποίο βρίσκονται. Στις τάξεις των γενετικών τελεστών ανήκουν οι τελεστές ανασυνδυασμού (ή διασταύρωσης στην περίπτωση των Γενετικών Αλγορίθμων) και οι τελεστές της μετάλλαξης. Υπάρχει μία πληθώρα διαφορετικών υλοποιήσεων των τελεστών και των δύο κατηγοριών, η οποία οφείλεται κυρίως στο γεγονός ότι αυτοί οι τελεστές εξαρτώνται άμεσα από την μορφή του χρωμοσώματος, δηλαδή της αναπαράστασης. Οι γενικές αρχές λειτουργίας τους όμως παραμένουν ίδιες και περιγράφονται με συνομία παρακάτω:
  - **Τελεστές Ανασυνδυασμού.** Οι τελεστές αυτοί είναι υπεύθυνοι για την ανταλλαγή τμημάτων γενετικού υλικού μεταξύ των ατόμων. Αυτό συμβαίνει κατά την φάση της αναπαραγωγής όπου, συνήθως, δύο γονείς ζευγαρώνουν για να παράξουν, συνήθως, δύο παιδιά. Οι απόγονοι (δηλαδή τα τέκνα), περιλαμβάνουν τμήματα γενετικού υλικού και από τους δύο (ή και περισσότερους) γονείς. Οι κατηγορίες που χωρίζονται αυτοί οι τελεστές είναι σε γενικές γραμμές δύο:
    - **Ανασυνδυασμός 'N' σημείων κοπής.** Η τιμή του 'N' μπορεί να είναι οποιοσδήποτε φυσικός αριθμός  $\{0, 1, \dots, n\}$ , όπου "n" πρέπει να είναι μικρότερο ή ίσο με το μήκος του χρωμοσώματος, δηλαδή του πλήθους των γονιδίων. Μία περιγραφή της λειτουργίας του τελεστή, για  $N = 1$ , δίνεται στην ενότητα 2.4.1 (και σχήμα 2.3). Όσο μεγαλύτερη τιμή έχει το 'N', τόσο περισσότερους συνδυασμούς μεταξύ των γονέων μπορούμε να δημιουργήσουμε. Από την άλλη μεριά όμως όσο αυξάνει το 'N', τόσο αυξάνει και η πιθανότητα να διαταράξουμε τα καλά σχήματα (schemata) που υπάρχουν μέσα στον πληθυσμό<sup>1</sup>.



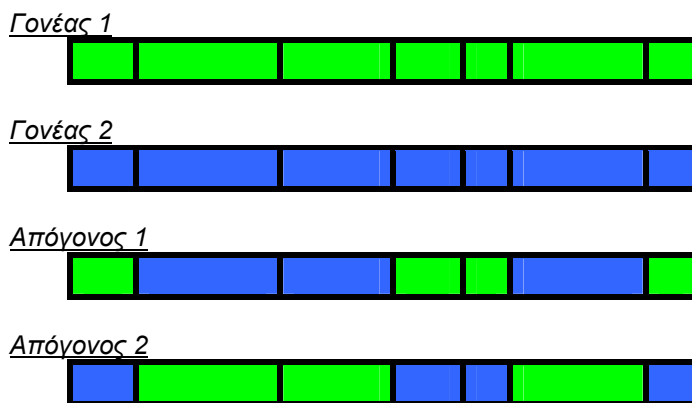
Σχήμα 4.3. Ανασυνδυασμός ενός σημείου.

<sup>1</sup> Για περισσότερες πληροφορίες σχετικά με τα σχήματα (schemata), ανατρέξτε στο σχετικό παράρτημα στο τέλος της εργασίας (Παράρτημα Α).



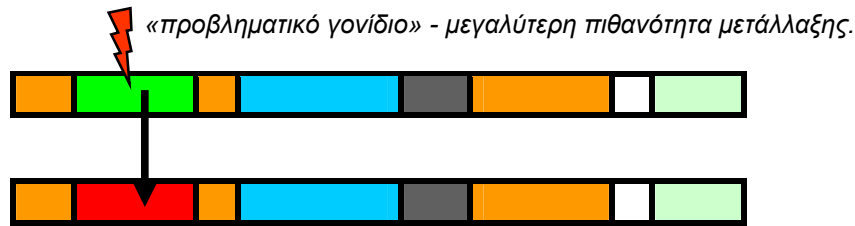
Σχήμα 4.4. Ανασυνδυασμός τεσσάρων σημείων.

- **Ομοιόμορφος Ανασυνδυασμός.** Στον ομοιόμορφο ανασυνδυασμό δεν υπάρχουν σημεία κοπής, αλλά κάθε γονίδιο έχει την ίδια πιθανότητα ανταλλαγής. Με τον τρόπο αυτό μπορούμε να επιτύχουμε όλους τους δυνατούς συνδυασμούς μεταξύ των γονέων.



Σχήμα 4.5. Ομοιόμορφος Ανασυνδυασμός.

- **Τελεστές Μετάλλαξης.** Η λειτουργία του τελεστή μετάλλαξης είναι σχετικά απλή και περιγράφεται με συντομία στην ενότητα 2.4.1 (και σχήμα 2.4). Εκείνο το οποίο πρέπει να αναφέρουμε είναι ότι εκτός από τον κλασσικό τελεστή μετάλλαξης, στον οποίο κάθε γονίδιο έχει την ίδια πιθανότητα να μεταλλαχτεί, υπάρχουν και άλλοι τρόποι εφαρμογής του σύμφωνα με τους οποίους κάθε γονίδιο έχει την δική του πιθανότητα να μεταλλαχτεί, ανάλογα με το πόσο συνεισφέρει στην συνολική ποινή του χρωμοσώματος («κατευθυνόμενη μετάλλαξη» - “directed mutation”). Γονίδια που είναι πιο «προβληματικά» έχουν μεγαλύτερη πιθανότητα να μεταλλαχτούν. Επίσης σε άλλες υλοποιήσεις η τιμή της μετάλλαξης δεν είναι και τόσο «τυχαία», όπως στην κλασσική υλοποίηση, αλλά «μεροληπτεί» και περιορίζεται σε ένα μικρότερο σύνολο αποδεκτών τιμών οι οποίες πιθανότατα να συνεισφέρουν θετικά στη συνολική ποιότητα του χρωμοσώματος.



**Σχήμα 4.6.** Παράδειγμα κατευθυνόμενης μετάλλαξης.

Εκείνο που θα πρέπει να αναφέρουμε, σχετικά με τους δύο προαναφερθέντες τελεστές (ανασυνδυασμού – μετάλλαξης), είναι ότι και οι δύο συνεισφέρουν με τον δικό τους τρόπο στη διαδικασία της εξέλιξης. Δεν μπορούμε να πούμε με βεβαιότητα ποιός από τους δύο είναι πιο σημαντικός ή πιο αποτελεσματικός, γιατί και οι δύο εξερευνούν το χώρο αναζήτησης με διαφορετικό τρόπο. Ο τελεστής ανασυνδυασμού, αναμιγνύει το ήδη υπάρχον γενετικό υλικό με σκοπό να βρεί έναν καλύτερο συνδυασμό που θα οδηγήσει σε καλύτερη ποιότητα του χρωμοσώματος. Σε καμία περίπτωση όμως δεν μπορεί να εμφανίσει σε μία θέση μία νέα τιμή για ένα γονίδιο. Αυτή είναι η δουλειά του τελεστή μετάλλαξης, η οποία μπορεί αρχικά να μή φέρει τα αναμενόμενα αποτελέσματα αλλά μακροχρόνια, με το πέρασμα των γενεών και των ανασυνδυασμών να προκύψει μία ποιοτικά καλύτερη λύση.

Εκτός από τα τέσσερα βασικά στοιχεία ενός ΕΑ που προαναφέρθηκαν, δεν θα πρέπει να παραβλέψουμε τη σημασία της Επιλογής, των γονέων, κατά την φάση της αναπαραγωγής (breeding). Η Επιλογή είναι αυτή που ευθύνεται για την καθοδήγηση της αναζήτησης σε ποιοτικά καλύτερες περιοχές του χώρου αναζήτησης (search space). Υπάρχουν τελεστές Επιλογής οι οποίοι επιλέγουν τα άτομα (τους γονείς) με βάση την απόλυτη τιμή της ποιότητάς τους και άλλοι με βάση την σχετική τιμή αυτής. Θα πρέπει να επισημάνουμε όμως ότι όταν επιλέγουμε τους γονείς με βάση την απόλυτη τιμή της ποιότητας μπορεί να φτάνουμε γρηγορότερα σε κάποιο βέλτιστο, αλλά υπάρχει μεγάλος κίνδυνος να πέσουμε σε κάποιο τοπικό βέλτιστο το οποίο να απέχει πολύ από το ολικό, που είναι συνήθως ο σκοπός μας (φαινόμενο πρόωρης σύγκλισης – premature convergence). Περισσότερες λεπτομέρειες σχετικά με διάφορους τελεστές Επιλογής θα δωθούν στο επόμενο κεφάλαιο, στο οποίο γίνεται περιγραφή του αλγορίθμου υλοποίησης, που πραγματοποιήθηκε στα πλαίσια της παρούσης πτυχιακής εργασίας.

## 4.2 Χρονοπρογραμματισμός Εξετάσεων και ΕΑ

Σε αυτήν την ενότητα θα παρουσιαστούν, περιληπτικά, μερικές από τις προσπάθειες που έγιναν κατά καιρούς, από διάφορους επιστήμονες, για την επίλυση του προβλήματος χρονοπρογραμματισμού εξετάσεων. Οι αναφορές που θα

γίνουν περιλαμβάνουν μόνο τις υλοποιήσεις που σχετίζονται με ΕΑ καθώς και με υβριδικές υλοποιήσεις αυτών.

Οι Edmund Burke, David Elliman και Rupert Weare στο [20] παρουσιάζουν ένα πρότυπο αυτοματοποιημένο σύστημα που χρησιμοποιεί Γενετικούς Αλγορίθμους (ΓΑ) και απευθύνεται τόσο στη δημιουργία προγραμμάτων εξετάσεων, όσο και στη δημιουργία προγραμμάτων διαλέξεων, σε ανώτερα εκπαιδευτικά ιδρύματα. Περιγράφουν κυρίως την πρώτη κατηγορία (δημιουργία προγραμμάτων εξετάσεων), αλλά ισχυρίζονται ότι η μέθοδος που περιγράφουν μπορεί να εφαρμοστεί εξίσου και στη δεύτερη. Η προσέγγιση που χρησιμοποίησαν ενσωματώνει γνώση που σχετίζεται με το πρόβλημα για να διασφαλίσουν ότι καμία από τις υποψήφιες λύσεις δεν είναι ανέφικτη (infeasible solution). Χρησιμοποίησαν άμεση αναπαράσταση χρωμοσώματος (direct representation) η οποία δεν περιελάμβανε μόνο το πότε, αλλά και το πού θα διεξαχθεί μία εξέταση. Επίσης αντί να χρησιμοποιήσουν ένα σταθερό πλήθος χρονικών περιόδων, που μπορεί να δημιουργούσε συγκρούσεις και να μη τοποθετούνταν κάποια μαθήματα, άφησαν το μήκος του προγράμματος μεταβλητό (πχ: από 1, μέχρι 100 χρονοθυρίδες). Φυσικά αν προέκυπτε πρόγραμμα μεγαλύτερο από το μέγιστο επιτρεπόμενο αριθμό, αυτό θα ήταν «ανεπιθύμητο». Παρ'όλ'αυτα με την μεταβλητότητα του προβλήματος δεν μπορούμε να προβλέψουμε πόσο μεγάλο θα μπορούσε να είναι το «βέλτιστο» πρόγραμμα και πιθανότατα με ένα συγκεκριμένο αριθμό χρονικών περιόδων ίσως να μην υπάρχει «βέλτιστο» πρόγραμμα. Αυτός που αναλάμβανε να δημιουργήσει τα προγράμματα είχε τη δυνατότητα να επηρεάσει το μήκος του προγράμματος διαβαθμίζοντας κατάλληλα έναν περιορισμό που σχετίζονταν με το μήκος αυτού. Όσο υψηλότερη ήταν η διαβάθμιση τόσο μικρότερο ήταν το μήκος του προγράμματος, ενώ όσο χαμηλότερη τόσο μεγάλωνε το μήκος του προγράμματος. Οι γενετικοί τελεστές της διασταύρωσης και της μετάλλαξης που χρησιμοποιήθηκαν ήταν ευριστικοί (heuristic), χρησιμοποιώντας γνώση που σχετίζονταν με το πρόβλημα. Για τελεστή Επιλογής χρησιμοποίησαν ένα πρότυπο σχήμα (standard schema), μιας και είναι ανεξάρτητος της αναπαράστασης. Ένα πλήθος από διαφορετικές συναρτήσεις αξιολόγησης ήταν προσαρμόσιμες δίνοντας έτσι τη δυνατότητα να παράγονται διαφορετικά προγράμματα τα οποία μπορούν να ικανοποιούν τις απαιτήσεις πολλών εκπαιδευτικών ιδρυμάτων. Για τη δοκιμή του συστήματος παρήγαγαν μία σειρά από δοκιμαστικά προβλήματα τα οποία περιελάμβαναν την τοποθέτηση εκατό (100) εξετάσεων σε τέσσερις (4) αίθουσες χωρητικότητας σαράντα (40), ογδόντα (80) και δύο των εκατό εξήντα (160) καθισμάτων. Ο αρχικός πληθυσμός περιελάμβανε εκατό (100), χρωμοσώματα και στις περισσότερες δοκιμές που έκαναν ο ΓΑ βελτίωνε την αρχική ποιότητα σημαντικά και μάλιστα σε ορισμένες δοκιμές η βελτίωση ήταν θεαματική. Παρ'όλ'αυτά τονίζουν ότι η ποιότητα ενός τελικού προγράμματος εξετάσεων, δηλαδή αν είναι «καλό» ή «κακό», εξαρτάται από το αν τελικά θα χρησιμοποιηθεί από το ίδρυμα ή όχι.

Οι ίδιοι ερευνητές (E.Burke, D.Elliman, R.Weare), σε μία άλλη εργασία τους [34] παρουσιάζουν έναν ισχυρό υβριδικό Γενετικό Αλγόριθμο (hybrid GA) για την επίλυση του χρονοπρογραμματισμού εξετάσεων (exam timetabling), χρησιμοποιώντας άμεση αναπαράσταση και ευριστικούς τελεστές διασταύρωσης για

να διασφαλίσουν ότι κανένας θεμελιώδης περιορισμός δε θα παραβιαστεί. Ο περιορισμός για την τοποθέτηση των εξετάσεων σε συγκεκριμένες αίθουσες, για πειραματικούς σκοπούς, αντικαταστάθηκε από τον κανόνα να μην τοποθετούνται περισσότεροι φοιτητές σε μια περίοδο, απ'όση είναι η χωρητικότητα αυτής της περιόδου. Για την αρχικοποίηση του πληθυσμού χρησιμοποιήθηκε ένα αλγόριθμος χρωματισμού γραφημάτων (graph colouring algorithm), ο οποίος λειτουργούσε ως εξής:

-Για κάθε μέλος του πληθυσμού (χρωμόσωμα).

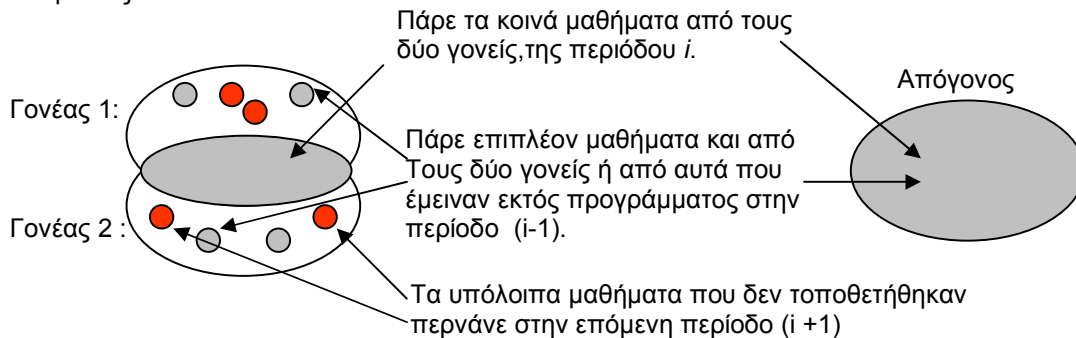
-Παρήγαγε μία τυχαία διάταξη των εξεταζομένων μαθημάτων.

-Πάρε κάθε ένα εξετ.μάθημα με την ακόλουθη σειρά:

Βρές την πρώτη χρονική περίοδο στην οποία μπορεί να τοποθετηθεί το μάθημα, χωρίς να συγκρούεται με άλλα εξεταζόμενα μαθήματα και το πλήθος των φοιτητών να μην υπερβαίνει τη συνολική χωρητικότητα αυτής της περιόδου, και τοποθέτησέ το μάθημα εκεί.

Με αυτόν τον τρόπο παρήγαγαν, γρήγορα, μεγάλους πληθυσμούς με τυχαίες τιμές, τα άτομα των οποίων, σαφώς, δεν είχαν το ελάχιστο δυνατό μήκος σε χρονική διάρκεια. Ο τελεστής διασταύρωσης, όπως αναφέραμε και παραπάνω, ήταν ευριστικός (heuristic), η λειτουργία του οποίου ήταν η εξής: αφού επιλεγούν οι γονείς, τότε για κάθε περίοδο ο τελεστής τοποθετεί τα κοινά μαθήματα των γονέων στο παιδί. Αν χρειάζεται να τοποθετηθούν επιπλέον μαθήματα χρησιμοποιεί έναν αλγόριθμο για να επιλέγει άλλα μαθήματα, χωρίς όμως να παραβιάζονται οι δύο θεμελιώδεις περιορισμοί. Δηλαδή να μην υπάρχουν συγκρούσεις μεταξύ των μαθημάτων και να μην υπερβαίνει το όριο των διαθέσιμων καθισμάτων της περιόδου.

- περίοδος "i"



Σχήμα 4.7. Λειτουργία ευριστικού τελεστή διασταύρωσης.

Οι αλγόριθμοι που χρησιμοποιήθηκαν για την επιλογή των μαθημάτων μέχρις ότου δεν μείνει κανένα μάθημα εκτός προγράμματος εξετάσεων ήταν αρκετοί και απλώς αναφέρονται ονομαστικά:

1. *Random*
2. *Largest Degree*
3. *Most Similar Exams*

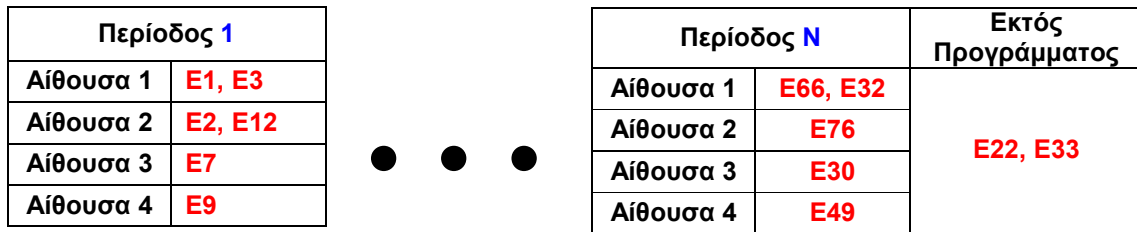


4. *Latest Scheduled in Other Parent*
5. *Least Conflicting with Previous Period*
6. *Both Late and Spread Together*

( Για περισσότερες πληροφορίες ο αναγνώστης μπορεί να ανατρέξει στη σχετική αναφορά της βιβλιογραφίας.)

Ο τελεστής της μετάλλαξης, όπως και αυτός της διασταύρωσης, θα έπρεπε να διασφαλίζει ότι μετά την εφαρμογή του, τα προγράμματα που θα παράγονταν θα ήταν εφικτά (feasible). Για να αποφευχθεί λοιπόν η δημιουργία ανέφικτων προγραμμάτων ενσωμάτωσαν τη λειτουργία της μετάλλαξης στον τελεστή της διασταύρωσης. Αυτό γινόταν με την προσθήκη μαθημάτων σε μία περίοδο από αυτά που είχαν μείνει εκτός προγράμματος στις προηγούμενες περιόδους. Για τη συνάρτηση ποιότητας δεν αναφέρουν κάποια συγκεκριμένα, αλλά αφήνουν ανοιχτή τη χρήση διαφόρων συναρτήσεων. Πάντως τονίζουν ότι μία απλή γραμμική συνάρτηση με βάρη στην απόδοση των ποινών (βλέπε ενότητα 4.1) είναι συνήθως αρκετή σε προβλήματα χρονοπρογραμματισμού. Στα πειράματα που πραγματοποίησαν χρησιμοποίησαν πληθυσμούς των διακοσίων (200) ατόμων, ενώ σαν τελεστή Επιλογής χρησιμοποίησαν εκθετική διαβάθμιση (exponential rank) με συντελεστή πίεσης (selection pressure) 0.986 . Τα αποτελέσματα που συνέλεξαν για κάθε αλγόριθμο περιγράφονται αναλυτικά στη σχετική αναφορά. Καταλήγουν ότι η χρήση υβριδικών (hybrid) τελεστών διασταύρωσης μαζί με την αρχικοποίηση του πληθυσμού από αλγορίθμους χρωματισμού γραφημάτων, μπορεί να παράξει προγράμματα εξετάσεων πολύ καλής ποιότητας ακόμα και σε προβλήματα με υψηλούς περιορισμούς. Τέλος, μερικές επεκτάσεις σε αυτόν τον αλγόριθμο παρουσιάζονται στο [17] από τους ίδιους συγγραφείς.

Οι Edmund K.Burke, James P.Newall και Rupert F.Weare, στο [16] παρουσιάζουν μία υβριδική προσέγγιση στο πρόβλημα του χρονοπρογραμματισμού των εξετάσεων, στα ανώτατα εκπαιδευτικά ιδρύματα, που περιλαμβάνει την ενσωμάτωση μεθόδων τοπικής αναζήτησης σε εξελικτικούς αλγορίθμους. Η προσέγγιση αυτή είναι γνωστή και σαν “Memetic Algorithm approach”. Η αρχική ιδέα για την χρήση αυτών των αλγορίθμων πρωτοπαρουσιάστηκε από τους Moscato και Norman, ενώ αργότερα διαμορφώθηκε από τους Radcliffe και Surrey οι οποίοι έκαναν και μία σύγκριση μεταξύ “memetic” και “genetic” αλγορίθμων. Η διαφορά μεταξύ της δομής “meme” από αυτής της “gene” (δηλαδή του γονιδίου), είναι ότι το “meme” μπορεί να παρουσιαστεί σαν μία μονάδα πληροφορίας η οποία αναπαράγει τον εαυτό του. Το “meme” καθώς μεταβιβάζεται μεταξύ των ατόμων προσαρμόζεται ανάλογα έτσι ώστε να προσαρμοστεί στο άτομο όσο το δυνατόν καλύτερα, ενώ το “gene” μεταβιβάζεται αυτούσιο. Το κύριο πλεονέκτημα των “memetic” αλγορίθμων είναι ότι μειώνουν το χώρο αναζήτησης, των πιθανών λύσεων, στον υποχώρο των τοπικών βέλτιστων (local optima). Στον αλγόριθμο που παρουσιάζουν κάνουν αποκλειστικά χρήση προγραμμάτων σταθερού μήκους, ενώ οι τεχνικές που περιλαμβάνουν είναι συνδυασμός ασθενούς (light) και ισχυρής (heavy) μετάλλαξης ακολουθούμενη από τον αλγόριθμο αναζήτησης Hill – Climbing. Η αναπαράσταση που χρησιμοποίησαν ήταν άμεση (direct) και είχε την παρακάτω μορφή:



Σχήμα 4.8. Άμεση αναπαράσταση στον “memetic” algorithm.

Όπως βλέπουμε και στην παραπάνω σχηματική αναπαράσταση, κάθε πιθανή λύση αναπαρίσταται ως ένα πλήθος από “memes”. Κάθε “meme” περιλαμβάνει πληροφορίες για την αίθουσα και για τη χρονική περίοδο που τοποθετείται κάθε συμβάν (πχ: E1, E2, E9, ...), ενώ υπάρχει και ένα τελευταίο “meme” που περιλαμβάνει όλα τα συμβάντα που δεν έχουν προγραμματιστεί ακόμα. Για την αρχικοποίηση του πληθυσμού χρησιμοποιήθηκε ένας σταθμισμένος «τροχός ρουλέτας» (weighted roulette wheel) ο οποίος τοποθετούσε τα συμβάντα (εξεταζόμενα μαθήματα) σε κατάλληλες χρονικές περιόδους. Όπως αναφέρθηκε και παραπάνω ο αλγόριθμος που παρουσιάστηκε έκανε χρήση συνδυασμού ασθενούς και ισχυρής μετάλλαξης, χωρίς την παρουσία τελεστή διασταύρωσης. Η ασθενής μετάλλαξη επέλεγε τυχαία συμβάντα μέσα στο πρόγραμμα και τα αναπρογραμμάτιζε σε κάποια άλλη αποδεκτή χρονοθυρίδα. Ενώ η ισχυρή μετάλλαξη μπορούσε να αναδιοργανώσει τα συμβάντα μίας ολόκληρης χρονικής περιόδου. Αμέσως μετά την μετάλλαξη ακολουθούσε η εφαρμογή ενός αιτιοκρατικού (deterministic) Hill – Climbing. Τέλος σαν μέθοδο Επιλογής των ατόμων που θα επιβιώσουν και θα αποτελέσουν τη νέα γενιά, χρησιμοποίησαν έναν τροχό ρουλέτας (roulette wheel selection). Οι δοκιμές που πραγματοποιήσαν έγιναν σε πραγματικά δεδομένα (Nottingham University 1994/1995 Data) τα οποία περιελάμβαναν την τοποθέτηση 805 εξετάσεων σε 32 χρονοθυρίδες, με συνολική χωρητικότητα 1550 καθισμάτων. Τα συμπεράσματά τους από τις δοκιμές τους είναι ότι ο αλγόριθμός τους παρέχει μία ενδιαφέρουσα προσέγγιση στο χώρο του εξελικτικού χρονοπρογραμματισμού (evolutionary timetabling). Επίσης η εφαρμογή του τελεστή Hill – Climbing παρόλο που προσθέτει υπολογιστικό χρόνο στη διαδικασία της εξέλιξης επιταχύνει την εύρεση καλύτερης λύσης απ’ότι αν χρησιμοποιούνταν μόνο εξελικτικοί τελεστές.

Οι Tony Wong, Pascal Côté και Paul Gely στο [19] παρουσιάζουν μία πρακτική προσέγγιση ενός Γενετικού Αλγορίθμου (ΓΑ) που επιλύει το πρόβλημα της δημιουργίας προγραμμάτων εξετάσεων, η οποία βρίσκει εφαρμογή στο πανεπιστήμιο του Quebec, “École de technologie supérieure”, από το 2001 με ενθαρρυντικά αποτελέσματα. Για την επίλυση του προβλήματός τους χωρίζουν το σύνολο των μαθημάτων, έστω  $C = \{c1, c2, c3, \dots, cn\}$ , σε μικρότερα υποσύνολα, έστω  $G_i = \{G_1 = \{c1, c4, c5\}, G_2 = \{c8, c12, c32\}, \dots\}$ , σε κάθε ένα από τα οποία να υπάρχουν φοιτητές που έχουν δηλώσει όλα τα μαθήματα του υποσυνόλου  $G_i$ . Κάθε ημέρα της εξεταστικής περιόδου τη χωρίζουν σε τρεις χρονοθυρίδες, οπότε σε ένα διάστημα  $d$  ημερών έχουν  $3*d$ , διαθέσιμες χρονοθυρίδες. Για την κωδικοποίηση του προβλήματος χρησιμοποιούν άμεση

αναπαράσταση του χρωμοσώματος, χρησιμοποιώντας ένα διάνυσμα ( $\mathbf{v}$ ) με διαστάσεις  $1 \times \mathbf{q}$ , όπου  $\mathbf{q}$  είναι το μέγιστο πλήθος των εξεταζομένων μαθημάτων. Κάθε θέση στο διάνυσμα αντιστοιχούσε σε ένα μάθημα, ενώ η τιμή που δεχόταν κάθε θέση αντιστοιχούσε σε μία χρονική περίοδο. Η συνάρτηση αξιολόγησης που χρησιμοποιήσαν ήταν μία συνηθισμένη, σε αυτού του είδους τα προβλήματα, γραμμική σταθμισμένη συνάρτηση απόδοσης ποινών η οποία ανταποκρινόταν πλήρως στους περιορισμούς που προσπαθούσε να ικανοποιήσει ο αλγόριθμος για το συγκεκριμένο εκπαιδευτικό ίδρυμα. Για την επιλογή των ατόμων που θα αναπαραχθούν χρησιμοποιήσαν τον δυαδικό “*tournament selector*”, οποίος επέλεγε τυχαία δύο άτομα του πληθυσμού και ο νικητής του τουρνουά (δηλαδή αυτός με τη μεγαλύτερη τιμή στην ποιότητα) επιλεγόταν σαν ένας από τους δύο γονείς. Κατόπιν για την παραγωγή των απογόνων χρησιμοποιούν τον τελεστή της ομοιόμορφης διασταύρωσης (uniform crossover), ενώ η μετάλλαξη γίνεται με μικρή πιθανότητα σε κάποιο γονίδιο του χρωμοσώματος αλλάζοντας την τιμή που έχει με μία εξίσου επιτρεπτή τιμή. Αυτό επειδή μπορεί να προκαλέσει συγκρούσεις μεταξύ των μαθημάτων ακολουθεί και μία τελική φάση η οποία αναλαμβάνει να σαρώσει τα χρωμοσώματα και αν εντοπίσει προβλήματα συγκρούσεων σε περισσότερα από ένα υποσύνολα ( $G_i$ ), τότε αναλαμβάνει να διορθώσει, με ευριστικές μεθόδους, εκείνο το υποσύνολο που εμφανίζει το μεγαλύτερο πρόβλημα. Αν τελικά η λύση που προέκυψε είναι χειρότερη από αυτήν που είχε προκύψει μετά από τη φάση της μετάλλαξης, τότε η διόρθωση απορρίπτεται. Το τελικό βήμα του αλγορίθμου είναι η συνολική αντικατάσταση του τρέχοντος πληθυσμού από τα νεοδημιουργηθέντα άτομα. Για να αποφασίσουν τις τιμές που θα έδιναν στις παραμέτρους των γενετικών τελεστών, πραγματοποίησαν μία σειρά πειραμάτων και κατέληξαν στις εξής τιμές:  $\alpha$ ) η πιθανότητα του τελεστή διασταύρωσης κυμαίνεται από 0.75 μέχρι 0.95 και  $\beta$ ) η πιθανότητα του τελεστή μετάλλαξης ήταν 0.01. Η φάση της διόρθωσης των λύσεων διεξαγόταν μόνο εάν εφαρμοζόταν ο τελεστής μετάλλαξης. Εν κατακλείδι συμπεραίνουν ότι τα αποτελέσματα που προκύπτουν από την εφαρμογή του συγκεκριμένου Γενετικού Αλγορίθμου επιφέρει θετικά αποτελέσματα στο πανεπιστήμιό τους.

Όπως βλέπουμε και από τα παραπάνω παραδείγματα, η εφαρμογή των Εξελικτικών Αλγορίθμων (EA) στο πρόβλημα του χρονοπρογραμματισμού εξετάσεων έχει αυξηθεί σημαντικά, ιδιαίτερα τις τελευταίες δύο δεκαετίες. Πάρα πολλές είναι οι αναφορές, από διάφορους ερευνητές, που σχετίζονται με αυτό το πρόβλημα και αφορούν είτε πειραματικά, είτε πραγματικά δεδομένα. Φαίνεται λοιπόν ότι οι EA έχουν κεντρίσει το ενδιαφέρον της επιστημονικής κοινότητας για τέτοιου είδους προβλήματα και όχι άδικα, αφού οι EA είναι αλγόριθμοι γενικού σκοπού. Αυτό σημαίνει πως οποιαδήποτε ωφέλη, ή βελτιστοποιήσεις επιτευχθούν για ένα πρόβλημα, μπορούν κάλλιστα να γενικευτούν και επωφεληθούν και άλλα προβλήματα που ανοίκουν θεωρητικά στον ίδιο χώρο. Επίσης, παρά το μειονέκτημα του υψηλού υπολογιστικού κόστους που, συνήθως, εμφανίζουν έχουν ένα σημαντικό πλεονέκτημα το οποίο μπορεί να αντισταθμίσει αυτή την κατάσταση. Αυτό δεν είναι άλλο από την παράλληλη φύση τους. Και αν αναλογιστούμε ότι τα τελευταία χρόνια η ισχύς των ηλεκτρονικών υπολογιστών αυξάνει κατακόρυφα και ότι το κόστος των παράλληλων Η/Υ πέφτει μπορούμε να

πούμε ότι το μέλλον των ΕΑ είναι αρκετά ευοίωνο στην αντιμετώπιση πολλών προβλημάτων.

# ΚΕΦΑΛΑΙΟ 5

## Υλοποίηση

Στο τρέχον κεφάλαιο γίνεται μία περιγραφή τόσο του προβλήματος που πραγματεύεται, η παρούσα πτυχιακή εργασία, όσο και του εξελικτικού αλγορίθμου που αναπτύχθηκε για την επίλυσή του. Η περιγραφή περιλαμβάνει τις ιδιαιτερότητες που εμφανίζει το πρόβλημα της δημιουργίας προγράμματος εξετάσεων στο τμήμα Πληροφορικής (Τ.Ε.Ι – Θεσσαλονίκης). Ο αλγόριθμος επίλυσης λοιπόν είχε σαν αφετηρία του να ικανοποιήσει πρώτα τους περιορισμούς που θέτει το εν λόγω εκπαιδευτικό ίδρυμα. Κατόπιν έγινε μία προσπάθεια γενίκευσης του αλγορίθμου, μέσα σε λογικά πλαίσια, έτσι ώστε να μπορεί να περιλάβει και άλλους περιορισμούς, οι οποίοι μπορεί να εμφανιστούν μελλοντικά στο τμήμα Πληροφορικής ή να ικανοποιούν τις απαιτήσεις άλλων εκπαιδευτικών ιδρυμάτων.

### 5.1 Δημιουργία Χρωμοσώματος (Αναπαράσταση Λύσης)

Ένα από τα σημαντικότερα βήματα στην εφαρμογή ενός Εξελικτικού Αλγορίθμου, για την επίλυση ενός προβλήματος, είναι η αναπαράσταση μίας πιθανής λύσης αυτού, σε μορφή κατανοητή από τον ΕΑ. Η αναπαράσταση του προβλήματος σε μορφή χρωμοσώματος είναι καθοριστική για δύο κυρίως λόγους: **α)** μια καλή αναπαράσταση μπορεί να μειώσει ή ακόμα και να εξαλείψει αρνητικές επιδράσεις όπως η επίσταση (epistasis - gene interaction) που εμφανίζεται συχνά σε προβλήματα χρονοπρογραμματισμού, και **β)** να επιταχύνει τη διαδικασία της εξέλιξης καθώς οι γενετικοί τελεστές που θα εφαρμοστούν εξαρτώνται άμεσα από την κωδικοποίηση του χρωμοσώματος, όποτε όσο πιο «γρήγοροι» είναι οι τελεστές τόσο πιο γρήγορα θα γίνεται η εξέλιξη του πληθυσμού.

Σε παρόμοια προβλήματα χρονοπρογραμματισμού, όπως αναφέρθηκαν και στο προηγούμενο κεφάλαιο (ενότητα 4.2), η τάση που έχει επικρατήσει είναι αυτή

της άμεσης αναπαράστασης (direct representation). Αυτή ήταν και η επιλογή για την παρούσα υλοποίηση. Τα στάδια που πέρασε ο σχεδιασμός του χρωμοσώματος μέχρι να πάρει την τελική του μορφή ήταν τέσσερα.

Αρχικά το χρωμόσωμα θεωρήθηκε σαν ένας δυσδιάστατος (2Δ) πίνακας (βλέπε σχήμα 5.1)  $N \times M$  ( $N$  γραμμές και  $M$  στήλες), του οποίου οι γραμμές αντιστοιχούσαν στα συμβάντα (events), δηλαδή στα μαθήματα που πρόκειται να εξεταστούν και οι στήλες αντιστοιχούσαν στις διαθέσιμες αίθουσες, όπου θα προγραμματίζονταν οι εξετάσεις. Οι τιμές που θα δεχόταν κάθε κελί του πίνακα θα ήταν ακέραιες και θα αντιστοιχούσαν στις χρονικές περιόδους (time slots).

Πρόγραμμα X	Αίθουσα 1	Αίθουσα 2	...	Αίθουσα M
<b>Μάθημα 1</b>	1	0	0	1
<b>Μάθημα 2</b>	0	12	0	0
<b>Μάθημα 3</b>	0	0	0	32
...	...	...	...	...
<b>Μάθημα N</b>	0	22	0	7

Σχήμα 5.1. Α' αναπαράσταση χρωμοσώματος.

Όπως βλέπουμε και από το παραπάνω σχήμα οι τιμές των κελιών ήταν ακέραιες και αντιστοιχούσαν στις διαθέσιμες χρονοθυρίδες. Οι χρονοθυρίδες θα ήταν προκαθορισμένες και θα ορίζονταν κάθε φορά πριν τρέξει ο αλγόριθμος. Τη χρονοθυρίδα θα μπορούσαμε να την δούμε σαν μία οντότητα η οποία είχε σαν χαρακτηριστικά της: **α)** την ημέρα διεξαγωγής της εξέτασης (πχ: 1<sup>η</sup>, 2<sup>η</sup>, 3<sup>η</sup>, ..., N<sup>η</sup>) και **β)** τις ώρες έναρξης και τερματισμού (πχ: 08:00 – 11:00). Το τμήμα Πληροφορικής συνήθως χρησιμοποιεί τέσσερις χρονοθυρίδες ανά ημέρα με σταθερή διάρκεια δύο ωρών και τριάντα λεπτά (2h:30') η κάθε μία. Αυτές είναι:

1. 08:30 – 11:00
2. 11:00 – 13:30
3. 13:30 – 16:00
4. 16:00 – 18:30

A/A	ΗΜΕΡΑ	ΩΡΑ ΕΝΑΡΞΗΣ	ΩΡΑ ΤΕΡΜΑΤΙΣΜΟΥ
1	1	08:30	11:00
2	1	11:00	13:30
3	1	13:30	16:00
4	1	16:00	18:30
5	2	08:30	11:00
6	2	11:00	13:30
7	2	13:30	16:00
8	2	16:00	18:30
...	...	...	...

Σχήμα 5.2. Πίνακας με τις διαθέσιμες χρονοθυρίδες (time slots).

Οπότε σε ένα σύνολο  $N$  ημερών που διαρκεί η εξεταστική περίοδος έχουμε  $4 \cdot N$  διαθέσιμες χρονοθυρίδες για την τοποθέτηση των συμβάντων. Φυσικά για λόγους ευελιξίας και γενικότητας τόσο το πλήθος των χρονοθυρίδων ανά ημέρα,

όσο και οι ώρες έναρξης και τερματισμού της κάθε μίας αφέθηκαν ελεύθερα στη βούληση του υπεύθυνου που θα φτιάξει το πρόγραμμα. Εκείνο όμως που αξίζει να προσέξουμε είναι ότι ο αλγόριθμος δεν έχει την «αίσθηση του χρόνου», δηλαδή δε γνωρίζει ότι η μία από την άλλη χρονοθυρίδα απέχουν «X» χρονικό διάστημα. Εκείνο που γνωρίζει και κατά συνέπεια υπολογίζει είναι η σειρά με την οποία είναι τοποθετημένες. Δηλαδή γνωρίζει ότι η 1<sup>η</sup> χρονοθυρίδα από την 4<sup>η</sup> απέχουν 2 κενές χρονοθυρίδες. Επίσης εκείνο που μπορεί να ξεχωρίσει είναι αν στο κενό αυτό διάστημα μεσολαβεί κάποιο Σαβ/Κο. Ετσι για παράδειγμα ο αλγόριθμος μπορεί να ξεχωρίσει ότι ενώ η πρώτη και η τρίτη χρονοθυρίδα, μέσα στην ίδια ημέρα, απέχουν ένα κενό διάστημα, η τελευταία χρονοθυρίδα της τελευταίας μέρας της εβδομάδας (πχ: Παρασκευή) και η πρώτη χρονοθυρίδα της πρώτης ημέρας από την ερχόμενη εβδομάδα απέχουν μεταξύ τους δύο ή τρεις ημέρες!! Τέλος, όσον αφορά τις χρονοθυρίδες θα πρέπει να αναφέρουμε ότι η πρώτη ημέρα της εξεταστικής περιόδου δεν αντιστοιχεί υποχρεωτικά στην ημέρα της εβδομάδος «Δευτέρα». Η αντιστοίχιση των ημερών της εξεταστικής περιόδου με αυτές της εβδομάδος αφέθηκαν στην κρίση του υπεύθυνου που θα φτιάξει το πρόγραμμα των εξετάσεων.

Όπως βλέπουμε όμως στο σχήμα 5.1, η συγκεκριμένη αναπαράσταση παρόλο που έχει αρκετά πλεονεκτήματα, όπως ότι περιλαμβάνει όλα τα μαθήματα ενσωματωμένα επομένως δεν υπάρχει κίνδυνος να μείνει κάποιο μάθημα εκτός προγράμματος, παρουσιάζει ένα μεγάλο μειονέκτημα. Αυτό είναι ότι κατά τη διάρκεια της εξέλιξης, με την εφαρμογή των γενετικών τελεστών, μπορεί να προκύψουν προγράμματα που δεν θα ήταν έγκυρα. Αυτό προκύπτει από την παρατήρηση ότι οι τιμές που δέχονται τα κελιά του πίνακα είναι ακέραιες από το μηδέν μέχρι και τη μέγιστη τιμή του A/A των χρονοθυρίδων (βλέπε σχήμα 5.2). Όπου υπήρχε μηδέν σήμαινε ότι το συγκεκριμένο μάθημα δεν εξετάζεται σε αυτήν την αίθουσα. Θα μπορούσε όμως κάλλιστα ένα μάθημα να προγραμματιστεί σε δύο ή και περισσότερες χρονοθυρίδες, αν σε κάθε κελί στην ίδια γραμμή έμπαινε διαφορετική χρονοθυρίδα, κάτι που δεν είναι αποδεκτό.

Για την αποφυγή αυτής της δυσάρεστης συνέπειας θα έπρεπε είτε να υλοποιηθούν πολύπλοκοι τελεστές ανασυνδυασμού και μετάλλαξης έτσι ώστε να ελέγχονται οι τιμές που τοποθετούνται στα κελιά, είτε να υπάρχει μια τρίτη φάση στην εξελικτική διαδικασία, μετά από τον ανασυνδυασμό και την μετάλλαξη, η οποία θα έλεγχε την εγκυρότητα των χρωμοσωμάτων και θα ενεργούσε ανάλογα (πχ: θα τα διόρθωνε ή θα τα απέρριπτε), είτε να εγκαταληφθεί αυτό το χρωμόσωμα και να αναζητηθεί καινούργια αναπαράσταση της λύσης. Η τρίτη επιλογή είναι φανερά πιο λογική καθώς ένας από τους στόχους της υλοποίησης του αλγορίθμου ήταν να γίνει όσο το δυνατόν πιο ταχύς. Οπότε με την αναζήτηση νέας αναπαράστασης θα γινόταν μία προσπάθεια να αποφευχθούν περιττές υπολογιστικές σπατάλες σε χρονοβόρους γενετικούς τελεστές και διορθωτικές διαδικασίες.

Η πρώτη σκέψη που έγινε ήταν να προστεθεί μία επιπλέον στήλη στον ήδη υπάρχοντα πίνακα, η οποία θα περιελάμβανε την χρονοθυρίδα στην οποία θα προγραμματιζόταν το εξεταζόμενο μάθημα. Παράλληλα οι υπόλοιπες στήλες, που αντιστοιχούσαν στις αίθουσες, θα δέχονταν μόνο τις τιμές «ένα (1)» αν το

μάθημα εξεταζόταν σε αυτή την αίθουσα και «μηδέν (0)» αν το μάθημα δεν εξεταζόταν σε αυτή την αίθουσα.

Πρόγραμμα X	Χρονο - θυρίδα	Αίθουσα 1	Αίθουσα 2	...	Αίθουσα M
<b>Μάθημα 1</b>	<b>11</b>	<b>1</b>	<b>0</b>	<b>0</b>	<b>1</b>
<b>Μάθημα 2</b>	<b>12</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>0</b>
<b>Μάθημα 3</b>	<b>32</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>
...	...	...	...	...	...
<b>Μάθημα N</b>	<b>22</b>	<b>0</b>	<b>1</b>	<b>0</b>	<b>1</b>

Σχήμα 5.3. Β' αναπαράσταση χρωμοσώματος.

Όπως βλέπουμε και στο άνωθι σχήμα το μειονέκτημα της προηγούμενης αναπαράστασης είχε πλέον εξαληφθεί καθώς κάθε μάθημα προγραμματίζεται σε μία μόνο χρονοθυρίδα, οπότε δεν υπάρχει κίνδυνος να βγει κάποιο πρόγραμμα άκυρο. Αργότερα διαπιστώθηκε ότι η συγκεκριμένη αναπαράσταση κάνει σπατάλη σε μνήμη (RAM) καθώς όπως βλέπουμε και στο σχήμα 5.3 ενώ στις στήλες που αντιστοιχούν σε αίθουσες χρησιμοποιεί μόνο δύο τιμές (0 ή 1) δεσμεύει πολύ περισσότερη μνήμη.

Η σκέψη που έγινε για την αντιμετώπιση αυτού του προβλήματος ήταν σχετικά απλή. Αντί να χρησιμοποιούνται τόσες στήλες όσες είναι οι διαθέσιμες αίθουσες θα χρησιμοποιείται μόνο μία η οποία θα περιλαμβάνει σε κωδικοποιημένη μορφή της αίθουσες που θα εξετάζονται τα μαθήματα. Η κωδικοποίηση γινόταν με μια απλή μετατροπή μίας δυαδικής συμβολοσειράς στο δεκαδικό σύστημα αρίθμησης. Η δυαδική συμβολοσειρά θα είχε μήκος ίσο με το πλήθος των διαθέσιμων αιθουσών. Επίσης κάθε αίθουσα θα κατείχε μία συγκεκριμένη θέση μέσα στη συμβολοσειρά, χωρίς όμως να έχει καμία ιδιαίτερη σημασία ποιά θα ήταν πρώτη, ποιά δεύτερη κ.ο.κ. Τέλος αν κάποιο μάθημα εξεταζόταν σε μία αίθουσα τότε στη θέση της μέσα στη συμβολοσειρά θα έμπαινε το σύμβολο «1», ενώ αν δεν εξεταζόταν σε αυτήν την αίθουσα θα έμπαινε το σύμβολο «0».

Έτσι, για παράδειγμα, αν έχουμε πέντε διαθέσιμες αίθουσες και το μάθημα προγραμματίζεται σε δύο από αυτές (πχ: στην πρώτη και στην τέταρτη), θα είχαμε τη δυαδική ακολουθία «1-0-0-1-0» που αντιστοιχεί στον ακέραιο «18» στο δεκαδικό σύστημα αρίθμησης (βλέπε σχήμα 5.4).

Πρόγραμμα X	Χρονο - θυρίδα	Αίθουσες
<b>Μάθημα 1</b>	<b>11</b>	<b>12</b>
<b>Μάθημα 2</b>	<b>12</b>	<b>10</b>
<b>Μάθημα 3</b>	<b>32</b>	<b>01</b>
...	...	...
<b>Μάθημα N</b>	<b>22</b>	<b>24</b>

→  $12_{10} = 01100_2$ , άρα το μάθημα θα εξεταστεί στις αίθουσες 2 και 3.

→  $01_{10} = 00001_2$ , άρα το μάθημα θα εξεταστεί στην αίθουσα 5.

Σχήμα 5.4. Γ' αναπαράσταση χρωμοσώματος.

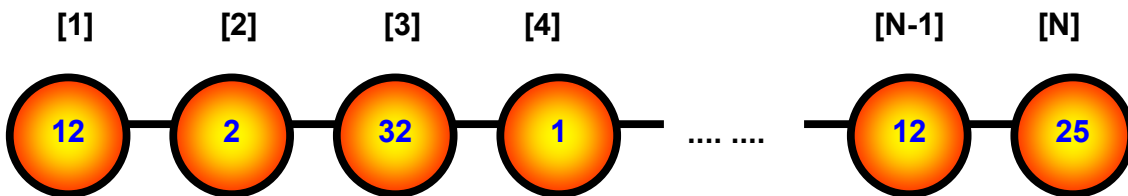


Όλες οι παραπάνω αναπαραστάσεις του προγράμματος (Α', Β' και Γ') περιελάμβαναν όχι μόνο το πότε θα διεξάγεται μία εξέταση, αλλά και το πού. Αυτό παρόλο που φαίνεται ιδανικό εγκυμονεί πολλούς κινδύνους, γιατί ενώ θα μπορούσε να βρεθεί ένα πρόγραμμα που ικανοποιεί όλους τους περιορισμούς, ακόμα και τους πιο απαιτητικούς, θα μπορούσε να έχει τοποθετήσει όλα τα μαθήματα σε μία αίθουσα και οι υπόλοιπες να είναι κενές. Για να γίνει πιο κατανοητό ας δούμε το παρακάτω παράδειγμα.

Έστω ότι έχουμε τέσσερις αίθουσες με πενήντα (50) καθίσματα η κάθε μία. Άρα η συνολική διαθέσιμη χωρητικότητα για κάθε χρονοθυρίδα είναι διακόσια καθίσματα. Έστω ότι σε μία χρονοθυρίδα τοποθετούμε με ιδανικό τρόπο (δηλαδή χωρίς να έχουμε συγκρούσεις μαθημάτων) κάποια μαθήματα των οποίων η συνολική απαίτηση σε καθίσματα είναι επίσης διακόσια. Ενώ λοιπόν θεωρητικά δεν υπερβαίνουμε το άνω όριο των συνολικών καθισμάτων, για την συγκεκριμένη χρονοθυρίδα, αν όλα τα μαθήματα τοποθετηθούν στην ίδια αίθουσα πρακτικά θα έχουμε πρόβλημα καθώς κάθε αίθουσα έχει χωρητικότητα πενήντα καθισμάτων. Οπότε θα έπρεπε γίνεται έλεγχος σε κάθε χρωμόσωμα για κάθε γενιά ώστε να αποτρέπονται τέτοιου είδους καταστάσεις. Αυτό όμως κρίθηκε υπολογιστικά ασύμφορο, καθώς όπως αναφέραμε και πριν ο αλγόριθμος θα έπρεπε να σχεδιαστεί και να υλοποιηθεί με τέτοιο τρόπο ώστε να μην υπάρχουν σπατάλες σε χρόνο εκτέλεσης.

Η λύση που βρέθηκε και εφαρμόστηκε ήταν να αποφορτιστεί ο εξελικτικός αλγόριθμος από το βάρος της κατανομής των μαθημάτων στις διάφορες αίθουσες και να ελέγχει μόνο αν παραβιάζεται η συνολική χωρητικότητα σε κάθε χρονοθυρίδα. Η κατανομή των μαθημάτων στις αίθουσες θα γινόταν μία μόνο φορά στο τέλος της εξελικτικής διαδικασίας και θα πραγματοποιείτο είτε αυτοματοποιημένα, από κάποιο απλό αλγόριθμο, είτε χειρωνακτικά από τον υπεύθυνο δημιουργίας των προγραμμάτων.

Οπότε το χρωμόσωμα (η αναπαράσταση του προγράμματος) πήρε την τελική του μορφή η οποία εφαρμόστηκε και στην πράξη. Αυτή αποτελείται από ένα διάνυσμα διαστάσεων  $1 \times N$ , όπου  $N$  είναι το πλήθος των μαθημάτων που θα συμμετέχουν στην εξεταστική διαδικασία. Κάθε κόμβος του διανύσματος θα αντιστοιχείται σε ένα μάθημα και η τιμή του κόμβου αντιστοιχείται στη χρονική περίοδο που θα εξεταστεί το συγκεκριμένο μάθημα.



Σχήμα 5.5. Τελική αναπαράσταση χρωμοσώματος.

Κάθε κόμβος αντιστοιχεί σε ένα και μόνο ένα μάθημα, ενώ το περιεχόμενο των κόμβων είναι ακέραιες τιμές από το ένα (1), μέχρι τη μέγιστη τιμή του  $A/A$

των χρονοθυρίδων (βλέπε σχήμα 5.2). Τα πλεονεκτήματα αυτής της αναπαράστασης είναι τα εξής:

1. **Απλότητα στη σχεδίαση.** Αυτό συνεπάγεται και απλότητα των γενετικών τελεστών που όπως έχουμε αναφέρει νωρίτερα εξαρτώνται άμεσα από την αναπαράσταση του χρωμοσώματος.
2. **Ταχύτερη επεξεργασία.** Αυτό προκύπτει από την απλότητα με την οποία είναι σχεδιασμένοι οι γενετικοί τελεστές (ανασυνδυασμού και μετάλλαξης) οι οποίοι δεν κάνουν περιττούς ελέγχους.
3. **Ελάχιστη δέσμευση μνήμης.** Όπως βλέπουμε και στο σχήμα 5.5 η ποσότητα της μνήμης που δεσμεύεται είναι πλέον ελάχιστη σε σχέση με τις προηγούμενες αναπαραστάσεις.
4. **Ενσωμάτωση Περιορισμών.** Ένας αυστηρός περιορισμός (hard constraint) είναι η εμφάνιση στο τελικό πρόγραμμα όλων των μαθημάτων που έχουν δηλωθεί. Αυτό επιτυγχάνεται χωρίς την παραμικρή προσπάθεια καθώς κάθε κόμβος του διανύσματος αναφέρεται σε ένα μάθημα. Άρα εκ των προτέρων δεν χρειάζεται να ελέγχουμε μήπως κάποιο μάθημα δεν έχει προγραμματιστεί. Επίσης ένας άλλος περιορισμός είναι ότι κάθε μάθημα πρέπει να εξετάζεται σε μία και μόνο μία χρονική περίοδο. Και αυτός ο περιορισμός επιτυγχάνεται απόλυτα και μάλιστα χωρίς υπολογιστικό κόστος καθώς η τιμή που δέχεται κάθε κόμβος είναι μία. Παράλληλα είναι επιθυμητό σε μία χρονοθυρίδα να προγραμματίζονται περισσότερα του ενός μαθήματα, για να γίνεται καλύτερη εκμετάλλευση των διαθέσιμων πόρων του ιδρύματος (πχ: αίθουσες). Αυτό είναι επιτρεπτό καθώς η τιμή κάθε κόμβου, που αντιστοιχεί σε μία χρονοθυρίδα, δεν είναι απαραίτητο να εμφανίζεται μία φορά μέσα στο πρόγραμμα, αλλά δύναται να εμφανίζεται περισσότερες της μίας φορές. Τέλος δεν υπάρχει κίνδυνος κάποιο γονίδιο να πάρει μία μη έγκυρη τιμή, καθώς οι τιμές που δέχονται είναι ακέραιες και όλοι οι γενετικοί τελεστές είναι παραμετροποιημένοι με τέτοιο τρόπο ώστε οι τιμές που τροφοδοτούν τα γονίδια να βρίσκονται πάντα εντός ορίων. Άρα δεν υπάρχει κίνδυνος κάποιο χρωμόσωμα να χαρακτηριστεί ανέφικτο (infeasible).

Όσον αφορά την εμφάνιση του φαινομένου της επίστασης, στην παραπάνω αναπαράσταση, είναι δύσκολο να προσδιορίσουμε τόσο το μέγεθος της, όσο και την επίδραση που θα έχει. Εκείνο που μπορούμε να πούμε με σιγουριά είναι πως από τη στιγμή που όλα τα γονίδια του χρωμοσώματος (κόμβοι του διανύσματος) κωδικοποιούν την ίδια έννοια (δηλαδή το μάθημα) το πρόβλημα της επίστασης θα περιορίζεται μόνο σε αυτή την έννοια. Δηλαδή δεν θα υπάρχουν αλληλεπιδράσεις με άλλες έννοιες όπως οι αίθουσες, οι καθηγητές, κτλ.

## 5.2 Σχεδιασμός Συνάρτησης Αξιολόγησης

Το δεύτερο βήμα στην κωδικοποίηση ενός προβλήματος για την αντιμετώπισή του από τους ΕΑ, είναι η δημιουργία της συνάρτησης αξιολόγησης (evaluation function). Η συνάρτηση αξιολόγησης είναι υπεύθυνη για την απόδοση, σε κάθε χρωμόσωμα, μίας τιμής που θα υποδεικνύει την ικανότητα του χρωμοσώματος

να επιβιώσει και να αναπαραχθεί μέσα στο συγκεκριμένο περιβάλλον στο οποίο εξελίσσεται. Στην ουσία η συνάρτηση αξιολόγησης προσομοιώνει τις ιδανικές συνθήκες του περιβάλλοντος στο οποίο υπάρχουν τα χρωμοσώματα. Χρειάζεται ιδιαίτερη μέριμνα στο σχεδιασμό της συνάρτησης αυτής, καθώς η αξιολόγηση των χρωμοσωμάτων αποτελεί, συνήθως, το πιο απαιτητικό τμήμα του αλγορίθμου σε υπολογιστικό χρόνο.

Πρίν συνεχίσουμε με την περιγραφή της συνάρτησης που υλοποιήθηκε, για τους σκοπούς της παρούσης εργασίας, θα ήταν καλό να ξεκαθαρίσουμε τη διαφορά μεταξύ της συνάρτησης **αξιολόγησης** και της συνάρτησης **ποιότητας**, γιατί συχνά επικρατεί σύγχυση και οι δύο όροι ταυτίζονται. Η συνάρτηση αξιολόγησης (evaluation function) σχετίζεται απόλυτα με την τιμή που παίρνει κάθε χρωμόσωμα, ανάλογα με το πόσο καλά είναι προσαρμοσμένο στις συνθήκες του περιβάλλοντός του. Από την άλλη μεριά η συνάρτηση ποιότητας (fitness – quality function) σχετίζεται με το πόσο καλά είναι προσαρμοσμένο κάθε χρωμόσωμα σε σχέση με το μέσο όρο της ποιότητας του συνολικού πληθυσμού στον οποίο βρίσκεται. Αν υποθέσουμε ότι η συνάρτηση αξιολόγησης είναι η  $f(x)$

τότε η συνάρτηση ποιότητας ορίζεται ως:  $g(x) = \frac{f(x)}{\bar{f}}$ , όπου  $\bar{f}$  είναι ο μέσος

όρος της  $f(x)$  όλου του πληθυσμού. Σε γενικές γραμμές η συνάρτηση ποιότητας ορίζει το πόσους απογόνους μπορεί να αποκτήσει κάθε χρωμόσωμα με δεδομένη την τιμή της συνάρτησης αξιολόγησης. Έτσι, αν ένα χρωμόσωμα έχει  $f(x) = 0.85$  και ισχύει  $\bar{f} = 0.45$ , τότε σημαίνει ότι αυτό το χρωμόσωμα μπορεί να αποκτήσει περίπου 1.88 απογόνους.

### 5.2.1 Περιορισμοί ικανοποίησης

Όπως αναφέρθηκε και νωρίτερα ο αλγόριθμος που αναπτύχθηκε είχε σαν αρχικό στόχο να ικανοποιήσει τις απαιτήσεις του τμήματος Πληροφορικής στη δημιουργία του προγράμματος εξετάσεων. Παρ'όλ'αυτά το σύνολο των περιορισμών που δημιουργήθηκε είναι λίγο πιο διευρυμένο έτσι ώστε να μπορεί να καλύπτει και μελλοντικές ανάγκες που πιθανά να προκύψουν τα επόμενα χρόνια.

Ήδη με την κωδικοποίηση του χρωμοσώματος καλύφθηκαν μερικοί περιορισμοί οι οποίοι θα μπορούσαν να χαρακτηρίσουν ένα πρόγραμμα ως ανέφικτο ή άκυρο. Οι περιορισμοί που αναπτύχθηκαν στη συνέχεια είχαν όλοι τη μορφή χαλαρών περιορισμών (soft constraints) με την έννοια ότι αν κάποιος παραβιαζόταν το πρόγραμμα δεν ακυρώνεται αλλά έχει επίπτωση στη συνολική του ποιότητα. Οι περιορισμοί που συμπεριλήφθηκαν είναι οι εξής:

1. Έλεγχος χωρητικότητας των διαθέσιμων εξεταζόμενων θέσεων.
2. Έλεγχος για μαθήματα που συνδέονται.
3. Δυνατότητα δέσμευσης μίας χρονικής περιόδου από κάποιο μάθημα.
4. Τα μαθήματα με το μεγαλύτερο πλήθος δηλώσεων να εξετάζονται μετά απο κενό (πχ: Σαβ/Κο ή αργία).
5. Βέλτιστη απόσταση μεταξύ των εξεταζόμενων μαθημάτων του φοιτητή.
6. Μαθήματα του ίδιου εξαμήνου, να μην εξετάζονται την ίδια μέρα.
7. Μαθήματα του ίδιου φοιτητή, να μην εξετάζονται την ίδια περίοδο.

Ο πρώτος περιορισμός παρόλο που σε γενικές γραμμές θεωρείται ως ένας από τους δύο θεμελιώδεις περιορισμούς που δεν πρέπει ποτέ να παραβιάζονται, στην περίπτωση του τμήματος Πληροφορικής μετατράπηκε σε χαλαρό περιορισμό εξαιτίας των περιορισμένων πόρων του ιδρύματος σε διαθέσιμα καθίσματα. Σε αυτόν τον έλεγχο ο αλγόριθμος ελέγχει αν το πλήθος των δηλώσεων των μαθημάτων που έχουν προγραμματιστεί σε μία χρονοθυρίδα υπερβαίνει το ανώτατο όριο της συνολικής χωρητικότητας. Σε αυτό το σημείο θα πρέπει να αναφέρουμε ότι ένα πρόβλημα που εμφανίζεται συχνά στα περισσότερα, αν όχι σε όλα, τα ανώτατα ιδρύματα είναι αυτό του «υπερπληθυσμού» (overpopulation - overcrowding) ορισμένων μαθημάτων. Αυτό προκαλείται από το υπερβολικά μεγάλο πλήθος δηλώσεων ορισμένων μαθημάτων τα οποία για να χωρέσουν μέσα στο πρόγραμμα συχνά δεσμεύουν όλη την χωρητικότητα μίας χρονοθυρίδας.

Για την αντιμετώπιση αυτού του προβλήματος ο αλγόριθμος παρέχει τρεις επιλογές από τις οποίες ο υπεύθυνος δημιουργίας των προγραμμάτων καλείται να επιλέξει μία από αυτές. Η πρώτη και πιο απλή είναι να μη λάβει καμία ιδιαίτερη πρόνοια. Δηλαδή το πρόβλημα δεν αντιμετωπίζεται με καμία μέθοδο και απλώς όπου εμφανίζεται θα μειώνει τη συνολική ποιότητα του προγράμματος, ανάλογα με το μέγεθος του προβλήματος. Η δεύτερη επιλογή είναι η χρήση των εργαστηριακών αιθουσών. Σε εκπαιδευτικά ιδρύματα τεχνολογικού χαρακτήρα, όπως είναι και το τμήμα Πληροφορικής (Τ.Ε.Ι - Θεσσαλονίκης), γίνεται σαφής διάκριση μεταξύ αιθουσών που προορίζονται για τη διδασκαλία θεωρητικών μαθημάτων και αυτών που γίνεται διδασκαλία εργαστηριακών μαθημάτων. Στη δεύτερη περίπτωση έχει κυριαρχήσει άτυπα ο όρος «Εργαστήριο». Συνηθίζεται λοιπόν οι εξετάσεις των εργαστηριακών μαθημάτων να διεξάγονται σε εργαστηριακές αίθουσες μίας και εκεί υπάρχει η απαραίτητη υλικοτεχνική υποδομή (Η/Υ, δίκτυα, κτλ) και οι εξετάσεις των θεωρητικών μαθημάτων να διεξάγονται στις αίθουσες «θεωρίας». Η εξέταση των εργαστηριακών μαθημάτων, στο τμήμα Πληροφορικής, διεξάγεται στις χρονικές περιόδους των δύο τελευταίων διαλέξεων όπως αυτές είχαν οριστεί από το πρόγραμμα μαθημάτων. Αυτό σημαίνει πως η δημιουργία των προγραμμάτων εξετάσεων δεν περιλαμβάνει τα εργαστηριακά μαθήματα, παρά μόνο τα θεωρητικά. Παρ'όλα αυτά ο αλγόριθμος με κατάλληλη παραμετροποίηση μπορεί είτε μεμονωμένα είτε ταυτόχρονα με τα θεωρητικά να παράγει και πρόγραμμα εξετάσεων για τα εργαστηριακά μαθήματα.

Στη συνήθη περίπτωση όμως της δημιουργίας των προγραμμάτων εξετάσεων οι εργαστηριακές αίθουσες είναι κενές. Αυτό λοιπόν προσπαθεί να το εκμεταλλευτεί ο αλγόριθμος και αν κατά την τοποθέτηση των μαθημάτων στις χρονοθυρίδες υπάρχει ανάγκη για περισσότερη χωρητικότητα τότε λαμβάνονται υπόψιν και τα καθίσματα των εργαστηριακών αιθουσών.

Η τρίτη επιλογή που παρέχει ο αλγόριθμος είναι αυτή των «συνεχόμενων περιόδων (χρονοθυρίδων)». Η έννοια των συνεχόμενων χρονοθυρίδων είναι σχετικά απλή και περιλαμβάνει την λογική, κυρίως, συνέχεια δύο ή και περισσότερων χρονοθυρίδων οι οποίες ανήκουν στην ίδια ημέρα της εξεταστικής περιόδου. Με αυτόν τον τρόπο εμφανίζεται μόνο μία χρονοθυρίδα, από αυτές που είναι συνεχόμενες, μέσα στο πρόγραμμα εξετάσεων η οποία έχει

πολλαπλάσια χωρητικότητα ανάλογα με το πόσες χρονοθυρίδες συνδέει. Για να επιτευχθεί αυτή η «συνέχεια» προστέθηκε μία επιπλέον στήλη στον πίνακα των χρονοθυρίδων (βλέπε σχήμα 5.2) η οποία λειτουργεί σαν συνδετικός κρίκος (ή σαν δείκτης, σε κάποια ανώτερη γλώσσα προγραμματισμού) και περιέχει τον A/A της χρονοθυρίδας με την οποία επιθυμούμε να συνδέσουμε την επιλεγμένη χρονοθυρίδα.

A/A	ΗΜΕΡΑ	ΩΡΑ ΕΝΑΡΞΗΣ	ΩΡΑ ΤΕΡΜΑΤΙΣΜΟΥ	ΣΥΝΔΕΣΗ A/A
1	1	08:30	11:00	0
2	1	11:00	13:30	1
3	1	13:30	16:00	1
4	1	16:00	18:30	0
5	2	08:30	11:00	0
6	2	11:00	13:30	0
7	2	13:30	16:00	0
8	2	16:00	18:30	0
...	...	...	...	...

**Σχήμα 5.6.** Παράδειγμα συνδεδεμένων χρονοθυρίδων.

Όπως βλέπουμε και στο παραπάνω σχήμα οι χρονοθυρίδες με A/A '2' και '3' έχουν στη στήλη «ΣΥΝΔΕΣΗ A/A» την τιμή '1'. Αυτό σημαίνει ότι αυτές οι δύο χρονοθυρίδες συνδέονται, λογικά, με τη χρονοθυρίδα που έχει A/A την τιμή '1'. Πρακτικά λοιπόν οι χρονοθυρίδες '2' και '3' δεν θα εμφανίζονται μέσα στο πρόγραμμα αλλά η χωρητικότητα τους θα προστεθεί σε αυτήν της '1'. Αυτό έχει σαν συνέπεια η χωρητικότητα της '1' να τριπλασιαστεί, με την προϋπόθεση ότι όλες οι χρονοθυρίδες έχουν την ίδια χωρητικότητα σε καθίσματα. Όπου υπάρχει η τιμή '0', στη στήλη «ΣΥΝΔΕΣΗ A/A» σημαίνει ότι η εν λόγω χρονοθυρίδα δεν συνδέεται με καμία άλλη χρονοθυρίδα.

Όπως καταλαβαίνουμε για να μπορέσει να γίνει χρήση αυτής της επιλογής θα πρέπει ο υπεύθυνος δημιουργίας των προγραμμάτων εξετάσεων να έχει προνοήσει και να έχει δημιουργήσει την/τις απαραίτητες συνεχόμενες χρονοθυρίδες. Αυτό είναι μία απλή διαδικασία καθώς ολόκληρη η παραμετροποίηση του ΕΑ γίνεται από ένα εύχρηστο και απλό γραφικό περιβάλλον.<sup>2</sup>

Ο δεύτερος περιορισμός δίνει τη δυνατότητα σε δύο μαθήματα να τοποθετηθούν σε διαδοχικές χρονοθυρίδες. Με αυτόν τον τρόπο παρέχεται και μία ακόμα έμμεση λύση στο πρόβλημα του υπερπληθυσμού, καθώς αν ένα μάθημα έχει πολύ μεγάλο πλήθος δηλώσεων μπορούμε να το χωρίσουμε, λογικά, σε δύο τμήματα και αυτά τα δύο τμήματα να συμμετέχουν σαν ξεχωριστά συμβάντα μέσα στη δημιουργία του προγράμματος. Έτσι τοποθετώντας τα σε διαδοχικές χρονοθυρίδες μειώνουμε ελαφρώς την «αδικία» που διαπράτουμε καθώς επιτρέπουμε στο ένα τμήμα να έχει περισσότερο χρόνο για μελέτη. Εκτός όμως από αυτή την έμμεση λύση ο δεύτερος περιορισμός καλύπτει διάφορες ανάγκες για την διαδοχική εξέταση μαθημάτων που ίσως να έχουν συναφές περιεχόμενο και είναι επιθυμητή η εξέτασή τους με το ελάχιστο δυνατό κενό διάστημα. Θα

<sup>2</sup> Για περισσότερες πληροφορίες σχετικά με τον τρόπο λειτουργίας της εφαρμογής που αναπτύχθηκε ανατρέξτε στο σχετικό εγχειρίδιο χρήστη (GA-FETP User's Manual).

πρέπει όμως να επισημάνουμε ότι παρόλο που δίνουμε τη δυνατότητα σε δυάδες μαθημάτων να εξετάζονται σε διαδοχικές περιόδους, δεν έχουμε τη δυνατότητα να ορίσουμε επακριβώς σε ποιές χρονοθυρίδες θα τοποθετηθούν αυτά τα μαθήματα. Ο αλγόριθμος έχει πλήρη ελευθερία να τοποθετήσει τα μαθήματα όπου αυτά εμφανίζουν την μικρότερη ποινή στη συνολική ποιότητα. Επίσης αν ο αλγόριθμος δεν μπορέσει να βρεί μία καλή τοποθέτηση για τα μαθήματα αυτά τότε ενδέχεται να μη τοποθετηθούν σε διαδοχικές χρονοθυρίδες.

Ο τρίτος περιορισμός προσπαθεί να τοποθετήσει ορισμένα μαθήματα σε συγκεκριμένες χρονοθυρίδες μέσα στο πρόγραμμα. Η ανάγκη που ώθησε στη δημιουργία αυτού του περιορισμού ήταν να δώσουμε τη δυνατότητα στους φοιτητές που εργάζονται να δίνουν ορισμένα μαθήματα σε ώρες εκτός εργασίας έτσι ώστε να μην έχουν αρνητικές επιπτώσεις. Εκτός όμως από αυτή την ανάγκη ο περιορισμός αυτός μπορεί να χρησιμοποιηθεί ακόμα και στην περίπτωση που κάποιος καθηγητής, για οποιοδήποτε λόγο, επιθυμεί να εξεταστεί το μάθημά του σε μία συγκεκριμένη χρονοθυρίδα. Θα πρέπει και εδώ να επισημάνουμε ότι ανάλογα με το πλήθος των μαθημάτων που πρέπει να προγραμματιστούν και το πλήθος των διαθέσιμων χρονοθυρίδων ενδέχεται η τοποθέτηση ενός μαθήματος σε μία συγκεκριμένη χρονοθυρίδα να μην είναι εφικτή. Ο αλγόριθμος όμως εγκυιάται ότι η τοποθέτηση του μαθήματος αν δεν είναι στη συγκεκριμένη χρονοθυρίδα θα είναι όσο γίνεται πιο κοντά.

Ο τέταρτος περιορισμός προσπαθεί να προγραμματίσει τα μαθήματα με το μεγαλύτερο πλήθος δηλώσεων είτε στην αρχή της εξεταστικής περιόδου, είτε μετά από κάποιο κενό διάστημα, όπως είναι τα Σαβ/κα, οι αργίες, κτλ. Σε αρχικό στάδιο του αλγορίθμου ελέγχονται όλες οι χρονοθυρίδες και ορισμένες μέρες της εξεταστικής περιόδου χαρακτηρίζονται ως «προνομιούχες». Αυτές είναι: **α)** η πρώτη μέρα της εξεταστικής περιόδου, **β)** η πρώτη μέρα μετά από το Σαβ/κο, **γ)** η πρώτη μέρα μετά από κάποιο κενό, πιθανά από κάποια αργία, κτλ. Σε αυτές τις ημέρες λοιπόν γίνεται μία προσπάθεια να τοποθετηθούν τα πιο απαιτητικά μαθήματα από άποψη δηλώσεων. Θα πρέπει να προσέξουμε όμως ότι ο αλγόριθμος τοποθετεί μονάχα ένα τέτοιο μάθημα σε κάθε «προνομιούχο» ημέρα. Δηλαδή ταξινομεί τα μαθήματα ανάλογα με το πλήθος των δηλώσεών τους και αν οι μέρες που θεωρούνται προνομιούχες είναι πχ: τρείς, τότε ευνοεί μόνο τα τρία πρώτα μαθήματα. Μάλιστα για την αποφυγή της τοποθέτησης δύο ή και περισσότερων μαθημάτων, με μεγάλο πλήθος δηλώσεων, στην ίδια ημέρα ο αλγόριθμος αποδίδει επιπλέον ποινή.

Ο πέμπτος περιορισμός καλύπτει τις συγκρούσεις δεύτερου, τρίτου, κτλ βαθμού (second, third, ... order conflicts). Με αυτόν τον περιορισμό ο αλγόριθμος προσπαθεί να «απλώσει» όσο το δυνατόν μακρύτερα τα μαθήματα που εξετάζεται κάθε φοιτητής. Για να το πετύχει αυτό υπολογίζει για κάθε φοιτητή το ποιλίκο του πλήθους των χρονοθυρίδων με το πλήθος των μαθημάτων στα οποία θα εξεταστεί και ανάλογα προσπαθεί να τα τοποθετεί σε ακέραια πολλαπλάσια του ποιλίκου αυτού. Για παράδειγμα, αν έχουμε διαθέσιμες σαράντα (40) χρονοθυρίδες και κάποιος φοιτητής εξετάζεται σε πέντε (5) μαθήματα, τότε το ιδανικό για αυτόν τον φοιτητή είναι να εξετάζεται κάθε 8 χρονοθυρίδες ( $40 / 5 = 8$ ). Θα πρέπει να αναφέρουμε ότι αυτός ο περιορισμός είναι από τους πιο δύσκολα ικανοποιήσιμους καθώς η προσπάθεια να ευνοηθεί ένας φοιτητής ή ένα

σύνολο φοιτητών με όμοιες δηλώσεις μπορεί να προκαλεί τη μή εύνοια ενός άλλου συνόλου φοιτητών με άλλες δηλώσεις μαθημάτων. Επειδή όπως είναι αντιληπτό και από τα παραπάνω ο αλγόριθμος πρέπει να υπολογίζει τις δηλώσεις κάθε φοιτητή, για την επιτάχυνση της διαδικασίας εισήχθηκε ένα είδος βελτιστοποίησης κατά την οποία σε αρχικό στάδιο ο αλγόριθμος βρίσκει όλες τις ομοιότητες μεταξύ των δηλώσεων των φοιτητών και κρατάει μόνο ένα αντίγραφο αυτών. Έτσι όταν ελέγχει έναν φοιτητή στην πραγματικότητα ελέγχει και όσους έχουν ακριβώς όμοιες δηλώσεις με αυτόν. Η βελτιστοποίηση αυτή αναμένεται να έχει θετικά αποτελέσματα κυρίως στους φοιτητές των πρώτων εξαμήνων καθώς εκεί παρατηρείται συχνά το φαινόμενο μεγάλες ομάδες ατόμων να δηλώνουν ακριβώς τα ίδια μαθήματα.

Ο έκτος, εν σειρά, περιορισμός προσπαθεί να τοποθετεί τα μαθήματα που ανήκουν στο ίδιο ακαδημαϊκό εξάμηνο σε διαφορετικές ημέρες της εξεταστικής περιόδου. Αυτό είναι σχετικά εύκολο να γίνει εκτός από την περίπτωση όπου το σύνολο των μαθημάτων ενός εξαμήνου υπερβαίνει το σύνολο των ημερών της εξεταστικής περιόδου. Αυτό πρακτικά συμβαίνει με τα μαθήματα του εβδόμου (Ζ) εξαμήνου τα οποία είναι συνήθως Μαθήματα Επιλογής (ΜΕ) και το πλήθος τους συχνά υπερβαίνει το πλήθος των ημερών που διαρκεί η εξεταστική περίοδος.

Ο έβδομος περιορισμός, αναφέρεται στη σχετική βιβλιογραφία σαν ένας από τους δύο θεμελιώδεις περιορισμούς που πρέπει κάθε πρόγραμμα να ικανοποιεί. Είναι αυτός που περιλαμβάνει τις συγκρούσεις πρώτου βαθμού (first order conflicts) και πρέπει κατά την ικανοποίησή του να μην υπάρχει κανένας φοιτητής που να απαιτείται να βρίσκεται σε δύο διαφορετικά μέρη ταυτόχρονα. Δηλαδή κανένας φοιτητής να μην εξετάζεται σε δύο διαφορετικά μαθήματα την ίδια χρονική στιγμή. Οι μέχρι τώρα δοκιμές που έγιναν, με τον τρέχον αλγόριθμο, σε πειραματικά δεδομένα έδειξαν ότι αυτός ο περιορισμός ικανοποιείται απόλυτα και μάλιστα σε πολύ σύντομο χρονικό διάστημα. Φυσικά με πραγματικά δεδομένα και μάλιστα του τμήματος Πληροφορικής ενδέχεται αυτός ο περιορισμός να μην ικανοποιείται απόλυτα. Αυτό οφείλεται κυρίως στους περιορισμένους πόρους του ιδρύματος σε διαθέσιμα καθίσματα. Αυτός ήταν και ο κύριος λόγος που ένας τόσο σοβαρός περιορισμός μετατράπηκε σε χαλαρό.

## 5.2.2 Απόδοση Ποινών

Η συνάρτηση αξιολόγησης που υλοποιήθηκε είναι μία σταθμισμένη γραμμική συνάρτηση απόδοσης ποινών, που συνηθίζεται σε αυτήν την κατηγορία προβλημάτων. Ο γενικός τύπος της συνάρτησης αυτής είναι ο εξής:

$$f(x) = \frac{1}{1 + \sum_{i=0}^n (w_i * c_i)}, \text{ όπου } \sum_{i=0}^n (w_i * c_i) \text{ είναι το άθροισμα όλων των}$$

ποινών που αποδίδονται στο χρωμόσωμα (x). Το (C<sub>i</sub>) είναι η ποινή του περιορισμού (i), το (W<sub>i</sub>) είναι η βαρύτητα του περιορισμού (i) στη συνολική ποιότητα ενώ (n) είναι το πλήθος των περιορισμών.

Η τιμή που δέχεται η μεταβλητή (W<sub>i</sub>) είναι ένας πραγματικός αριθμός (κινητής υποδιαστολής) από το μηδέν (0.0) μέχρι και το ένα (1.0). Όσο μεγαλύτερη είναι

αυτή η τιμή τόσο πιο αυστηρός θεωρείται ο περιορισμός (i). Η τιμή αυτή καθορίζεται χειρωνακτικά από τον υπεύθυνο δημιουργίας των προγραμμάτων εξετάσεων λίγο πριν τρέξει ο αλγόριθμος και από την στιγμή που θα οριστούν δεν μπορούν να μεταβληθούν κατά την διάρκεια που ο αλγόριθμος θα βρίσκεται σε εξέλιξη. Από την άλλη μεριά το (c<sub>i</sub>) είναι επίσης ένας πραγματικός αριθμός, αλλά αποδίδεται με διαφορετικό τρόπο σε κάθε περιορισμό.

Στην περίπτωση του πρώτου περιορισμού το (c<sub>1</sub>) είναι '1' για κάθε θέση που υπερβαίνει το ανώτατο όριο των θέσεων κάθε χρονοθυρίδας, διαφορετικά είναι '0'. Δηλαδή αν σε μία χρονοθυρίδα η διαθέσιμη χωρητικότητα είναι πχ: 200 καθίσματα και τα μαθήματα που προγραμματίστηκαν απαιτούν 225 καθίσματα, τότε το (c<sub>1</sub>) θα ισούται με 25 (225 – 200 = 25) και η συνολική ποινή που θα προστεθεί θα είναι [ 25\*(w<sub>1</sub>) ].

Στο δεύτερο περιορισμό το (c<sub>2</sub>) δεν αποδίδεται τόσο απλά όσο το (c<sub>1</sub>). Ο υπολογισμός του δεν είναι σταθερός και εξαρτάται κυρίως από δύο συνιστώσες. Η πρώτη συνιστώσα είναι η πραγματική απόσταση (dtP) μεταξύ των δύο μαθημάτων που επιθυμούμε να εξεταστούν σε διαδοχικές χρονοθυρίδες. Η επιθυμητή τιμή της μεταβλητής αυτής είναι, φυσικά, η μονάδα. Οποιαδήποτε άλλη τιμή συμβάλει θετικά στην αύξηση του συνόλου της ποινής του χρωμοσώματος. Η δεύτερη συνιστώσα είναι ο μέσος όρος (M.O) της «βαρύτητας» των δύο μαθημάτων. Η έννοια της βαρύτητας ενός μαθήματος περιλήφθηκε για δικαιότερη απονομή στη ποινή του χρωμοσώματος. Η βαρύτητα ενός μαθήματος εξαρτάται, με τη σειρά της, από δύο παράγοντες. Ο πρώτος είναι μία ακέραια τιμή που ορίστηκε από το ένα (1) μέχρι και το πέντε (5). Όσο αυξάνει η τιμή, τόσο αυξάνει, θεωρητικά, και η δυσκολία ενός μαθήματος. Η τιμή αυτή αποδίδεται από τον υπεύθυνο δημιουργίας των προγραμμάτων και ορίζεται είτε από στατιστικά στοιχεία παλιότερων εξεταστικών περιόδων, για να βρούμε το ποσοστό των επιτυγχόντων σε κάθε μάθημα, είτε από κάποια έρευνα που μπορεί να γίνει εσωτερικά στο τμήμα για να απαντήσουν οι ίδιοι οι φοιτητές ποιά μαθήματα θεωρούν δύσκολα, είτε από τον καθηγητή που διδάσκει κάθε μάθημα, είτε από όλα τα προηγούμενα μαζί. Ο δεύτερος παράγοντας είναι το πλήθος των φοιτητών που έχει δηλωθεί σε κάθε μάθημα. Από τα δύο προηγούμενα προκύπτει κάθε φορά δυναμικά η βαρύτητα κάθε μαθήματος ως εξής:

$$Lweight(i) = \frac{Lrate(i) * declarations(i)}{\sum_{i=1}^K declarations(i)}, \text{ όπου:}$$

- $Lrate(i)$  είναι η διαβάθμιση του μαθήματος (i), [1, ... ,5]
- $declarations(i)$  είναι το πλήθος των δηλώσεων του μαθήματος (i)
- $\sum_{i=1}^k declarations(i)$  είναι το πλήθος όλων των δηλώσεων, και
- $k$ , είναι το πλήθος των μαθημάτων που συμμετέχουν στο πρόγραμμα.

Από τα παραπάνω προκύπτει ότι το (c<sub>2</sub>) ισούται με:

$$(c_2) = dtP * M.O (Lweight1, Lweight2)$$



Στην περίπτωση που τα δύο μαθήματα τοποθετούνταν στην ίδια χρονοθυρίδα, τότε το  $dtP$  έπαιρνε την τιμή ένα (1) και όχι μηδέν (0).

Για την απόδοση ποινής στον τρίτο περιορισμό τα πράγματα κινούνται περίπου στην ίδια βάση και η ποινή ( $c3$ ) ορίζεται από το γινόμενο:

$(c3) = dtP2 * Lweight$ , όπου ( $dtP2$ ) είναι η απόσταση μεταξύ της επιθυμητής χρονοθυρίδας και της χρονοθυρίδας στην οποία είναι τοποθετημένο το μάθημα στο τρέχων πρόγραμμα. Και ( $Lweight$ ) είναι η βαρύτητα του συγκεκριμένου μαθήματος, όπως αυτή ορίστηκε νωρίτερα.

Στον τέταρτο περιορισμό η απόδοση ποινής απλοποιείται και στην περίπτωση που κάποιο από τα μαθήματα, με το μεγαλύτερο πλήθος δηλώσεων, βρίσκεται σε «ευνοϊκή» μέρα τότε το ( $c4$ ) ισούται με μηδέν (0). Σε αντίθετη περίπτωση ταυτίζεται με την βαρύτητα του συγκεκριμένου μαθήματος ( $Lweight$ ). Τέλος αν έχει τοποθετηθεί σε μία «ευνοϊκή» μέρα αλλά στην ίδια ημέρα έχει τοποθετηθεί νωρίτερα κάποιο άλλο απαιτητικό μάθημα, τότε το ( $c4$ ) ισούται με τη μισή, περίπου, βαρύτητα του μαθήματος. Αυτό έγινε για να αποτρέπεται η τοποθέτηση μαθημάτων με πολλές δηλώσεις την ίδια ημέρα της εξεταστικής περιόδου.

Η απόδοση ποινής στον πέμπτο περιορισμό γίνεται επίσης με ευριστικά κριτήρια καθώς για κάθε φοιτητή ξεχωριστά, υπολογίζεται η βέλτιστη απόσταση στην οποία μπορούν να τοποθετηθούν τα μαθήματα στα οποία εξετάζεται. Κατόπιν ο αλγόριθμος «βλέπει» σε ποιές χρονοθυρίδες είναι τοποθετημένα τα μαθήματά του και ανάλογα με τις διαφορές στα κενά μεταξύ όλων των μαθημάτων αποδίδει την ποινή συνυπολογίζοντας και την βαρύτητα των μαθημάτων. Για να γίνει αυτό πιο κατανοητό, μπορούμε να δούμε το παρακάτω παράδειγμα:

Έστω ότι ένας φοιτητής έχει δηλώσει τέσσερα μαθήματα με βαρύτητες  $Lw1$ ,  $Lw2$ ,  $Lw3$  και  $Lw4$ . Επίσης θεωρούμε ότι το σύνολο των χρονοθυρίδων είναι σαράντα (40). Με μία απλή (ακέραια) διαίρεση υπολογίζουμε την βέλτιστη απόσταση για τον φοιτητή. Αυτή είναι  $BestD = 40 / 4 = 10$ . Τότε παίρνουμε τα μαθήματα με την σειρά που εμφανίζονται στο πρόγραμμα και έστω ότι αυτά έχουν τοποθετηθεί στις εξής χρονοθυρίδες:  $ts1 = 2$ ,  $ts2 = 9$ ,  $ts3 = 30$ ,  $ts4 = 40$ .

Αυτό σημαίνει ότι το πρώτο μάθημα εξετάζεται στη δεύτερη χρονοθυρίδα ( $ts1 = 2$ ), το δεύτερο μάθημα στην ένατη ( $ts2 = 9$ ), το τρίτο στην τριακοστή ( $ts3 = 30$ ) και τέλος το τέταρτο στην τεσσαρακοστή ( $ts4 = 40$ ). Έπειτα βλέπουμε τις διαφορές που υπάρχουν μεταξύ των διαδοχικών χρονοθυρίδων. Αυτές είναι:  $d1 = ts2 - ts1 = 9 - 2 = 7$ ,  $d2 = ts3 - ts2 = 30 - 9 = 21$ ,  $d3 = ts4 - ts3 = 40 - 30 = 10$ . Όπως βλέπουμε και από τα παραπάνω η μόνη διαφορά ( $d_i$ ) που είναι «αξιόποινη» είναι η πρώτη ( $d1 = 7$ ). Οι υπόλοιπες υπερβαίνουν το όριο οπότε δεν αποδίδουμε καμία ποινή. Ακόμη όμως και στην περίπτωση που η απόσταση είναι κάτω του επιθυμητού ορίου ελέγχουμε μήπως μεταξύ των δύο χρονοθυρίδων μεσολαβεί Σαβ/κο. Κάτι τέτοιο αποτρέπει την απόδοση ποινής καθώς μπορεί τα μαθήματα να εξετάζονται σε μικρότερο χρονικό διάστημα, αλλά όπως αναφέραμε και νωρίτερα ο αλγόριθμος δεν έχει την «αίσθηση» του χρόνου, οπότε πρέπει να φροντίσουμε στο να μη γίνει «άδικος» και αποδίδει ποινή εκεί που δεν πρέπει. Αν δεχτούμε ότι τελικά πρέπει να αποδοθεί ποινή τότε αυτή θα ισούται με:

$$(c5) = \frac{1}{1 + (day2 - day1)} * (BestD - d1) * M.O(Lw1, Lw2) / SumStudents,$$

όπου:

*SumStudents* είναι το άθροισμα όλων των φοιτητών που θα εξεταστούν. Το *day2* είναι η μέρα της δεύτερης χρονοθυρίδας (*ts2*), ενώ το *day1* είναι η ημέρα της πρώτης χρονοθυρίδας (*ts1*). Αυτό γίνεται για να δώσουμε μεγαλύτερη ποινή όταν οι δύο χρονοθυρίδες βρίσκονται και οι δύο στην ίδια μέρα ( $c5 = 1 * \dots$ ), ενώ αν βρίσκονται σε διαφορετικές μέρες η ποινή είναι μικρότερη ( $c5 = 0.5 * \dots$ ,  $c5 = 0.33 * \dots$ , κτλ).

Ο έκτος περιορισμός ελέγχει το πλήθος των μαθημάτων, ανά εξαμήνο, που έχουν τοποθετηθεί σε κάθε ημέρα της εξεταστικής περιόδου και αν προκύψει σε μία μέρα να έχουν τοποθετηθεί περισσότερα του ενός μαθήματα τότε αποδίδει ποινή ίση με το άθροισμα της βαρύτητας των μαθημάτων. Δηλαδή:

$$(c6) = \Sigma(LwA) + \Sigma(LwB) + \Sigma(Lw\Gamma) + \dots, \text{ όπου}$$

$\Sigma(LwA)$  είναι το άθροισμα της βαρύτητας των μαθημάτων του Α εξαμήνου, που τοποθετήθηκαν την ίδια μέρα, κ.ο.κ.

Ο έβδομος και τελευταίος περιορισμός αφορά τις γνωστές συγκρούσεις πρώτης τάξης (*first order conflict*). Σύμφωνα με αυτόν τον περιορισμό κανένας φοιτητής δεν πρέπει να εξετάζεται την ίδια χρονική στιγμή σε δύο ή και περισσότερα μαθήματα. Κάθε φορά που παραβιάζεται αυτός ο περιορισμός η ποινή που αποδίδεται ισούται με το άθροισμα των *M.O* της βαρύτητας των μαθημάτων που ο φοιτητής εξετάζεται ταυτόχρονα, ανά δυάδες μαθημάτων. Δηλαδή αν ο φοιτητής εξετάζεται ταυτόχρονα σε δύο μαθήματα, τότε  $(c7) = M.O(Lw1, Lw2)$ . Αν όμως ο φοιτητής εξετάζεται ταυτόχρονα σε τρία μαθήματα, τότε  $(c7) = M.O(Lw1, Lw2) + M.O(Lw2, Lw3)$ , κ.ο.κ.

Θα πρέπει εν τέλει να αναφέρουμε ότι για την επιτάχυνση της διαδικασίας αξιολόγησης ο αλγόριθμος παρακάμπτει τους περιορισμούς που έχουν τιμή μηδέν (0.0) στην παράμετρο ( $w_i$ ). Δηλαδή αν δεν ενδιαφερόμαστε για την ικανοποίηση ενός περιορισμού ή από την εμπειρία μας γνωρίζουμε ότι η σημαντικότητα ενός περιορισμού είναι τόσο μικρή που δεν αξίζει να χαλάσει την συνολική ποιότητα των προγραμμάτων, τότε απλώς θέτουμε τιμή μηδέν (0.0) στην παράμετρο ( $w_i$ ) του συγκεκριμένου περιορισμού.

### 5.3 Τελεστές

Οι τελεστές που εφαρμόζονται στους Εξελικτικούς Αλγορίθμους (EA) χωρίζονται σε δύο κύριες κατηγορίες: **α)** σε αυτούς που εξαρτώνται άμεσα από την αναπαράσταση του χρωμοσώματος (*representation depended*) και **β)** σε αυτούς που δεν εξαρτώνται από την αναπαράσταση του χρωμοσώματος (*representation independed*). Στην πρώτη κατηγορία ανήκουν κυρίως οι γενετικοί τελεστές του ανασυνδυασμού (διασταύρωσης στην περίπτωση των ΓΑ) και της μετάλλαξης, ενώ στη δεύτερη κατηγορία ανήκουν τελεστές επιλογής – επιβίωσης. Στις ενότητες που ακολουθούν γίνεται τόσο αναφορά των τελεστών που μπορεί να εφαρμόσει ο αλγόριθμος που αναπτύχθηκε, όσο και ο τρόπος λειτουργίας του κάθε τελεστή.

### 5.3.1 Τελεστές Ανασυνδυασμού

Οι τελεστές ανασυνδυασμού είναι υπεύθυνοι για την μεταφορά τμημάτων γενετικού υλικού από τους γονείς στα παιδιά. Τόσο οι κατηγορίες στις οποίες χωρίζονται, όσο και ο γενικός τρόπος λειτουργίας των τελεστών αυτών περιγράφονται σε προηγούμενο κεφάλαιο (4.1). Στον αλγόριθμο που υλοποιήθηκε, στα πλαίσια της παρούσης πτυχιακής εργασίας, δώθηκε η δυνατότητα να χρησιμοποιηθεί μία πληθώρα τελεστών διασταύρωσης. Αυτοί είναι οι εξής:

1. **Διασταύρωση 'N' σημείων κοπής (List Crossover).**
2. **Ομοιόμορφη Διασταύρωση (Uniform Crossover).**
3. **Τυχαία Επιλογή Μεθόδου (Random Crossover).**

Η λειτουργία των δύο πρώτων τελεστών έχει περιγραφεί, όπως αναφέραμε νωρίτερα, σε προηγούμενη ενότητα. Η τρίτη από τις παραπάνω επιλογές δεν εκτελεί κάποιο νέο είδος διασταύρωσης αλλά με τυχαία τρόπο επιλέγει μία από τις δύο παραπάνω μεθόδους κάθε φορά που πρόκειται να εφαρμοστεί διασταύρωση στα χρωμοσώματα του πληθυσμού. Επειδή η συνάρτηση που επιλέγει ποιά μέθοδος θα εφαρμοστεί, επιλέγει με ίση πιθανότητα αναμένεται στο τέλος της εξελικτικής διαδικασίας οι δύο παραπάνω τελεστές να έχουν εφαρμοστεί περίπου ίσες φορές.

Όσον αφορά τις παραμέτρους με τις οποίες μπορούμε να τροφοδοτήσουμε τον αλγόριθμο, αυτές αφορούν το πλήθος των σημείων κοπής (δηλαδή το 'N'), με προεπιλεγμένη τιμή τη μονάδα (1), καθώς και την πιθανότητα διασταύρωσης ( $P_{cross}$ ) με προεπιλεγμένη τιμή 90% ή 0.9. Οι τιμές αυτές μπορούν να αλλάξουν κατά την διάρκεια της εξελικτικής διαδικασίας μόνο αν ο αλγόριθμος τεθεί σε κατάσταση παύσης (paused). Επίσης για λόγους ασφάλειας αν η τιμή που δώσουμε για το πλήθος των σημείων κοπής ξεπερνά το μέγεθος του χρωμοσώματος (δηλαδή το πλήθος των γονιδίων), τότε σαν προεπιλεγμένη τιμή τοποθετείται η μονάδα (1). Θα πρέπει να θυμόμαστε ότι το χρωμόσωμα δημιουργείται με δυναμικό τρόπο κάθε φορά που ετοιμάζεται να τρέξει ο αλγόριθμος και εξαρτάται αποκλειστικά από τα δεδομένα με τα οποία θα αρχικοποιηθεί. Χονδρικά το πλήθος των γονιδίων (άρα το μέγιστο πλήθος των σημείων κοπής) ισούται με το πλήθος των μαθημάτων που θα τοποθετηθούν στο πρόγραμμα εξετάσεων.

### 5.3.2 Τελεστές Μετάλλαξης

Οι τελεστές μετάλλαξης είναι υπεύθυνοι για την εισαγωγή νέου γενετικού υλικού που πιο πριν δεν υπήρχε. Η εφαρμογή του τελεστή της μετάλλαξης έχει υποστεί ιδιαίτερη κριτική κυρίως από τους υποστηρικτές των Γενετικών Αλγορίθμων (ΓΑ), όπου εκεί η μετάλλαξη λειτουργεί σαν δευτερεύων τελεστής στο παρασκήνιο. Από την άλλη μεριά οι υποστηρικτές των Εξελικτικών Στρατηγικών και του Εξελικτικού Προγραμματισμού υποστηρίζουν ότι η μετάλλαξη από μόνη της μπορεί να παρέχει ένα ισχυρό εργαλείο αναζήτησης. Στην παρούσα υλοποίηση του αλγορίθμου δίνεται η δυνατότητα σε αυτόν που θα παραμετροποιήσει τον ΕΑ να χρησιμοποιήσει την μετάλλαξη με όποιο τρόπο

επιθυμεί. Είτε από μόνη της, είτε σε συνδυασμό με τη διασταύρωση, είτε να την απορρίψει τελείως. Οι τελεστές που παρέχονται και εδώ είναι αρκετοί και περιγράφονται παρακάτω<sup>3</sup>:

1. **IntMutator**. Η λειτουργία αυτού του τελεστή είναι η συνήθης λειτουργία της μετάλλαξης. Δηλαδή σαρώνεται κάθε χρωμόσωμα και με πολύ μικρή πιθανότητα (συνήθως 0.001 στην περίπτωση των ΓΑ) αλλάζει η τιμή ενός γονιδίου με μία άλλη τυχαία, αλλά έγκυρη τιμή.
2. **IntIncMutator**. Η λειτουργία αυτού του τελεστή είναι όμοια με του προηγούμενου με την εξής διαφορά. Όταν επιλέγεται μία θέση για μετάλλαξη, τότε ο αλγόριθμος δεν τοποθετεί μια τυχαία τιμή αλλά λειτουργεί ως εξής: χρησιμοποιεί μία πιθανότητα, που ορίζεται από το χρήστη και αν αυτή η πιθανότητα είναι αρκετά μεγάλη τότε ο αλγόριθμος αυξάνει την τιμή που περιέχεται στο γονίδιο κατά μία ποσότητα (πχ: 1), η οποία ορίζεται επίσης από το χρήστη. Αντιθέτως αν η πιθανότητα κριθεί μικρή τότε η τιμή του γονιδίου μειώνεται κατά την ίδια προαναφερθήσα ποσότητα.
3. **ExchangeMutator**. Ο τελεστής αυτός είναι πιο απλός από τους δύο παραπάνω και η λειτουργία του είναι η εξής: επιλέγει με τυχαίο τρόπο δύο θέσεις μέσα στο χρωμόσωμα και απλώς ανταλλάσει τις τιμές των γονιδίων.
4. **RandomMutator**. Ο τελεστής αυτός χρησιμοποιεί μία συνάρτηση η οποία επιλέγει ισοπίθανα έναν από τους τρεις τελεστές που προαναφέρθηκαν και εκτελεί κάθε φορά την συγκεκριμένη λειτουργία μετάλλαξης.

Οι παράμετροι των τελεστών που μπορούμε να ορίσουμε είναι, αρχικά η πιθανότητα μετάλλαξης ( $P_{mut}$ ) με αρχική τιμή 0.01, η πιθανότητα αύξησης / μείωσης της τιμής του γονιδίου με προεπιλεγμένη τιμή το 50% ή 0.5 και τέλος η ποσότητα αύξησης/μείωσης της τιμής του γονιδίου με αρχική τιμή τη μονάδα (1).

### 5.3.3 Τελεστές Επιλογής

Οι τελεστές επιλογής είναι αυτοί που κατευθύνουν την αναζήτηση σε ποιοτικά καλύτερες περιοχές του χώρου αναζήτησης (search space). Η φάση της επιλογής είναι αυτή που επιλέγει τους γονείς που θα ζευγαρώσουν και θα παράξουν τους απογόνους, πολλοί από τους οποίους (ίσως και όλοι) θα αποτελέσουν τη νέα γενιά. Υπάρχουν τελεστές που επιλέγουν τους γονείς με βάση την απόλυτη τιμή της συνάρτησης αξιολόγησης (evaluation function) και άλλοι που χρησιμοποιούν περισσότερο την ποιότητα των χρωμοσωμάτων, δηλαδή την τιμή που προκύπτει από τη συνάρτηση ποιότητας (fitness function). Το πλήθος των τελεστών που μπορούν να εφαρμοστούν, στον τρέχοντα αλγόριθμο, είναι ιδιαίτερα μεγάλο, πράγμα που του δίνει μεγάλες δυνατότητες και του επιτρέπει να πειραματιστεί με πολλούς και διάφορους τρόπους.

---

<sup>3</sup> Τα ονόματα των τελεστών ταυτίζονται με αυτά που εμφανίζονται και στην υλοποίηση του αλγορίθμου, στην εφαρμογή λογισμικού.

1. **TournamentSelector.** Η λειτουργία αυτού του τελεστή είναι η διεξαγωγή ενός «τουρνουά», κάθε φορά που χρειάζεται να επιλεγεί ένας γονέας. Συλλέγει με τυχαίο τρόπο ένα πλήθος ατόμων από τον πληθυσμό και το καλύτερο άτομο από αυτά (δηλαδή αυτός με την μεγαλύτερη τιμή στην συνάρτηση αξιολόγησης) επιλέγεται ως νικητής. Αυτό μπορεί να παραμετροποιηθεί με μία πιθανότητα, έτσι ώστε να δίνεται η δυνατότητα και σε κάποιο άτομο με μικρότερη ποιότητα να επιλέγεται ως νικητής. Η πρώτη βασική παράμετρος που δέχεται είναι το πλήθος των συμμετεχόντων στο τουρνουά. Αυτό συνήθως είναι δύο (2) αλλά μπορεί να είναι και μεγαλύτερο. Επίσης μία δεύτερη παράμετρος είναι η πιθανότητα με την οποία θα επιλέγεται ως νικητής το άτομο με την μεγαλύτερη ποιότητα. Συνήθως είναι μεγάλη (πχ: 0.9 ή 90%), αλλά εξαρτάται από αυτόν που παραμετροποιεί τον αλγόριθμο το πόσο πολωμένος (biased) θέλει να είναι ο τελεστής.
2. **SequentialSelector.** Η λειτουργία αυτού του τελεστή είναι απλή και δίνει την δυνατότητα σε όλα τα άτομα του πληθυσμού να επιλεγούν σαν γονείς. Σε πρώτη φάση ταξινομεί τα άτομα του πληθυσμού σε φθίνουσα τάξη και έπειτα κάθε φορά που απαιτείται ένας γονέας επιλέγει με την σειρά από τον καλύτερο προς το χειρότερο. Αν φτάσει στο τέλος και απαιτηθεί νέο άτομο, τότε ξεκινάει πάλι από την αρχή.
3. **BestNSelector.** Η λειτουργία του τελεστή “BestN”, είναι παρόμοια με αυτήν του προηγούμενου τελεστή (“Sequential”), με την διαφορά ότι εδώ ορίζουμε το μέγιστο όριο επιλογής ατόμων από την ταξινομημένη λίστα. Δηλαδή η παράμετρος που δίνουμε στη μεταβλητή ‘N’, δηλώνει ότι κάθε φορά θα επιλέγουμε έναν γονέα από τους ‘N’ καλύτερους της λίστας. Το ‘N’ μπορεί να ισούται από ένα (1), μέχρι και το πλήθος του συνόλου των ατόμων από τα οποία επιλέγουμε. Όσο μικρότερη είναι η τιμή του ‘N’ τόσο πιο πολωμένη είναι η επιλογή και υπάρχει κίνδυνος η αναζήτηση να κολλήσει σε κάποιο τοπικό βέλτιστο. Από την άλλη μεριά αν το ‘N’ ισούται με το μέγεθος του πληθυσμού, τότε η λειτουργία του τελεστή θα ταυτιστεί με αυτήν του τελεστή “SequentialSelector”.
4. **RankingSelector.** Ο τελεστής “Ranking” χρησιμοποιεί τη θέση κάθε ατόμου μέσα σε μία φθίνουσα ταξινομημένη λίστα και υπολογίζει την πιθανότητα που έχει κάθε άτομο να επιλεγεί ως γονέας. Κατόπιν με το γύρισμα του τροχού μιάς ρουλέτας επιλέγει ένα άτομο κάθε φορά. Για να γίνει πιο κατανοητή η λειτουργία του ας δούμε το παρακάτω παράδειγμα:
  - a. Έστω ότι έχουμε έναν πληθυσμό με πέντε (5) άτομα. Ταξινομούμε αυτά τα άτομα σε φθίνουσα σειρά με βάση την τιμή της συνάρτησης αξιολόγησης. Η πιθανότητα που έχει κάθε άτομο να επιλεγεί ισούται

$$\text{με: } p(i) = \frac{\text{populationSize} - i}{\sum_{i=0}^{\text{populationSize}} i}, \text{ όπου}$$

- i. “i” είναι η θέση του μέσα στη λίστα,
- ii. “populationSize” είναι το μέγεθος του πληθυσμού.

Δηλαδή η πιθανότητα να επιλεγεί το πρώτο άτομο της λίστας είναι:

$$p(0) = \frac{5 - 0}{(0 + 1 + 2 + 3 + 4)} = \frac{5}{10} = 0.5. \text{ Αφού υπολογιστούν οι}$$

πιθανότητες επιλογής όλων των ατόμων εφαρμόζεται η τεχνική του γυρίσματος της ρουλέτας για την επιλογή του ατόμου από τον πληθυσμό.

5. **RandomSelector.** Η λειτουργία του τελεστή “Random” δεν απαιτεί καμία ιδιαίτερη γνώση που να αφορά την ποιότητα των ατόμων, ούτε εκτελεί περίπλοκες διαδικασίες. Απλώς κάθε φορά που εκτελείται επιλέγει τυχαία , με ίση πιθανότητα, ένα άτομα από ολόκληρο τον πληθυσμό.
6. **RouletteWheelSelector.** Η λειτουργία αυτού του τελεστή κάνει χρήση όχι της απόλυτης τιμής της συνάρτησης αξιολόγησης, αλλά της ποιότητας κάθε ατόμου. Δηλαδή ανάλογα με την ποιότητα κάθε ατόμου του αντιστοιχεί και ένα «κομμάτι» από τον τροχό μίας ρουλέτας. Κατόπιν περιστρέφεται ο τροχός και στο σημείο που θα σταματήσει επιλέγεται εκείνο το άτομο στον οποίο είναι ανατεθεί αυτό το τμήμα του τροχού. Με αυτό τον τρόπο όσο μεγαλύτερη ποιότητα έχει ένα άτομο τόσο μεγαλύτερη πιθανότητα έχει να επιλεγεί σαν γονέας. Παράλληλα αφήνεται ανοιχτό το ενδεχόμενο να επιλεγούν σαν γονείς και άτομα που δεν έχουν καλή ποιότητα. Με αυτόν τον τρόπο μπορούμε να διατηρήσουμε μία ανομοιοτητα (diversity) στα άτομα του πληθυσμού που μας βοηθάει στο να εξερευνούμε καλύτερα τον χώρο αναζήτησης.
7. **WorstNSelector.** Η λειτουργία του τελεστή “WorstN”, όπως μπορεί κανείς εύκολα να συμπεράνει ταυτίζεται με αυτήν του τελεστή “BestN”. Η μόνη διαφορά είναι ότι αυτός ο τελεστής ταξινομεί ανάποδα τη λίστα με τα χρωμοσώματα του πληθυσμού (δηλαδή σε αύξουσα σειρά) και επιλέγει κάθε φορά από τα ‘N’ χειρότερα άτομα του πληθυσμού.
8. **GreedyOverSelector.** Ο τελεστής επιλογής “GreedyOverSelector” εφαρμόζεται κυρίως σε μεγάλους πληθυσμούς. Για την λειτουργία του απαιτείται ο διαχωρισμός του πληθυσμού σε δύο υποπληθυσμούς και η επιλογή κάθε ατόμου γίνεται ανάλογα με μία πιθανότητα είτε από τον έναν υποπληθυσμό, είτε από τον άλλο. Αξίζει να σημειωθεί ότι σε κάθε υποπληθυσμό εφαρμόζεται ένας διαφορετικός τελεστής επιλογής. Οι τελεστές που μπορούν να εφαρμοστούν σε κάθε υποπληθυσμό είναι οι εξής:
  - a. TournamentSelector
  - b. SequentialSelector
  - c. BestNSelector
  - d. RankingSelector
  - e. RandomSelector
  - f. RouletteWheelSelector
  - g. WorstNSelector

Ο διαχωρισμός των ατόμων γίνεται με τυχαίο τρόπο και με βάση κάποιο ποσοστό. Δηλαδή ο ένας υποπληθυσμός παίρνει το ‘X%’ (πχ:30%) επί του συνόλου των ατόμων και ο δεύτερος το ‘(1-X)%’ (πχ: 70%). Κατόπιν η επιλογή γίνεται πάλι με κάποια πιθανότητα. Δηλαδή η πιθανότητα να

- επιλεγεί άτομο από τον πρώτο υποπληθυσμό είναι, έστω,  $P_x$  και η πιθανότητα να επιλεγεί το άτομο από το δεύτερο είναι  $1-P_x$ .
9. **FullStarSelector (NeighbourhoodSelector)**. Ο τελεστής αυτός, καθώς και οι επόμενοι δύο που ακολουθούν (10, 11) επιλέγουν τα άτομα από ένα πλέγμα δύο διαστάσεων ( $2D - 2D$ ). Το δυσδιάστατο αυτό πλέγμα έχει διαστάσεις  $N \times N$  και τα άτομα που τοποθετούνται μέσα σε αυτό έχουν τη δυνατότητα να μπούν είτε ταξινομημένα (με βάση την τιμή της συνάρτησης αξιολόγησης) είτε αταξινόμητα, δηλαδή με τυχαίο τρόπο. Ο τελεστής αυτός με βάση μία παράμετρο που καλείται ακτίνα (radius), δημιουργεί μία μικρή γειτονιά ατόμων στο  $2D$  πλέγμα, με επίκεντρο μία θέση η οποία αρχικά επιλέγεται τυχαία και κατόπιν επιλέγεται από την θέση του ατόμου που είχε επιλεγεί την προηγούμενη φορά (βλέπε σχήμα 5.7 που ακολουθεί). Έπειτα τα άτομα αυτής της γειτονιάς συλλέγονται και εφαρμόζεται πάνω τους η λειτουργία ενός άλλου τελεστή επιλογής. Οι συντεταγμένες του ατόμου που θα επιλεγεί θα αποτελέσουν το σημείο εκκίνησης της δημιουργίας της επόμενης γειτονιάς. Οι τελεστές που μπορούν να εφαρμοστούν είναι οι ίδιοι που χρησιμοποιεί και ο τελεστής "Greedy-OverSelector" (8).
  10. **FullCrossSelector**. Η λειτουργία αυτού του τελεστή μοιάζει αρκετά με του προηγούμενου τελεστή. Μόνο που σε αυτήν την περίπτωση πάνω στο  $2D$  πλέγμα δεν σχηματίζεται γειτονιά, αλλά ένας «σταυρός» (βλέπε σχήμα 5.7 που ακολουθεί). Κατόπιν τα άτομα επιλέγονται και με βάση την εφαρμογή ενός άλλου τελεστή επιλογής επιλέγεται το άτομο – γονέας. Το επίκεντρο του νέου σταυρού θα είναι πάλι οι συντεταγμένες του ατόμου που επιλέχτηκε τελευταίο. Και σε αυτήν την περίπτωση έχουμε την δυνατότητα να τοποθετούμε τα άτομα μέσα στο πλέγμα είτε ταξινομημένα, είτε τυχαία.
  11. **RandomWalkSelector**. Ο τελεστής αυτός παρόλο που χρησιμοποιεί ένα  $2D$  πλέγμα εργάζεται διαφορετικά από τους άλλους δύο καθώς δεν περιέχει την έννοια της γειτονιάς (βλέπε σχήμα 5.7 που ακολουθεί). Αρχικά επιλέγεται τυχαία μία θέση και από εκεί ξεκινάει έναν «τυχαίο περίπατο» συλλέγοντας σε κάθε στάση το άτομο που βρίσκεται σε αυτή τη θέση. Το μήκος του περιπάτου είναι προκαθορισμένο (πχ: 8 βήματα) αλλά οι κατευθύνσεις των βημάτων επιλέγονται τυχαία την ώρα που ο αλγόριθμος εξελίσσεται. Το άτομο που θα επιλεγεί, ως γονέας, είναι αυτό που θα έχει τη μεγαλύτερη τιμή στη συνάρτηση αξιολόγησης. Κατόπιν η θέση αυτού του ατόμου θα αποτελέσει τη νέα θέση εκκίνησης του περιπάτου.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0	$x+2, y-1$	$x+2, y$	$x+2, y+1$	$x+2, y+2$	$x+2, y+3$														$x+2, y-3$	$x+2, y-2$
1	$x+3, y-1$	$x+3, y$	$x+3, y+1$	$x+3, y+2$	$x+3, y+3$														$x+3, y-3$	$x+3, y-2$
2																				
3																$x-2, y$				
4																$x-1, y$				
5				$x-1, y-1$	$x-1, y$	$x-1, y+1$							$x, y-2$	$x, y-1$	$x, y$	$x, y+1$	$x, y+2$			
6				$x, y-1$	$x, y$	$x, y+1$										$x+1, y$				
7				$x+1, y-1$	$x+1, y$	$x+1, y+1$										$x+2, y$				
8																				
9																				
10																				
11												$x-4, y-2$								
12											$x-3, y-3$	$x-3, y-2$								
13											$x-2, y-3$									
14											$x-1, y-3$	$x-1, y-2$								
15	$x-3, y-1$	$x-3, y$	$x-3, y+1$	$x-3, y+2$	$x-3, y+3$							$x, y-2$	$x, y-1$	$x, y$					$x-3, y-3$	$x-3, y-2$
16	$x-2, y-1$	$x-2, y$	$x-2, y+1$	$x-2, y+2$	$x-2, y+3$														$x-2, y-3$	$x-2, y-2$
17	$x-1, y-1$	$x-1, y$	$x-1, y+1$	$x-1, y+2$	$x-1, y+3$														$x-1, y-3$	$x-1, y-2$
18	$x, y-1$	$x, y$	$x, y+1$	$x, y+2$	$x, y+3$														$x, y-3$	$x, y-2$
19	$x+1, y-1$	$x+1, y$	$x+1, y+1$	$x+1, y+2$	$x+1, y+3$														$x+1, y-3$	$x+1, y-2$

Σχήμα 5.7. Αναπαράσταση τελεστών επιλογής σε 2Δ πλέγμα (9, 10, 11).

Όπως βλέπουμε και στο παραπάνω σχήμα με κόκκινο χρώμα σημειώνεται το επίκεντρο της εφαρμογής κάθε τελεστή.



Στο σχήμα με το κίτρινο χρώμα βλέπουμε την γειτονιά που σχηματίζει ο τελεστής “NeighbourhoodSelector” (9) με ακτίνα τη μονάδα. Όμοια ο ίδιος τελεστής σχηματίζει και την γειτονιά με το μπλέ χρώμα που έχει ακτίνα τρία (3).

Με πράσινο χρώμα εμφανίζεται ο σταυρός που δημιουργεί ο τελεστής “FullCrossSelector” (10) με ακτίνα δύο (2) μονάδες, ενώ τέλος με χρυσαφί χρώμα εμφανίζεται ο «περίπατος» μήκους οχτώ (8) βημάτων από τον τελεστή “RandomWalkSelector” (11). Τα βήματα που κάνει, από το επίκεντρο, είναι διαδοχικά :

1. αριστερά (left)
2. αριστερά (left)
3. πάνω (up)
4. αριστερά (left)
5. πάνω (up)
6. πάνω (up)
7. δεξιά (right)
8. πάνω (up)

Στις παραπάνω κινήσεις δεν εμφανίζεται άλλη μία κίνηση, η «κάτω (down)» η οποία εμφανίζεται επί ίσης όροις με τις άλλες τρεις κινήσεις. Απλώς στο παραπάνω παράδειγμα δεν έτυχε να εμφανιστεί.

### 5.3.4 Τελεστές Επιβίωσης

Η φάση της επιβίωσης αποτελεί το τελικό στάδιο στο συνήθη εξελικτικό κύκλο καθώς μετά την ολοκλήρωση αυτής, δημιουργείται ο νέος πληθυσμός και ξεκινάει απ’την αρχή η συνήθης εξελικτική διαδικασία. Κατά πολλούς ερευνητές η φάση της επιβίωσης των ατόμων, που θα αποτελέσουν τη νέα γενιά, αποτελεί το δεύτερο ήμισυ της φάσης της Επιλογής, θεωρώντας ως πρώτο ήμισυ την επιλογή των γονέων που θα παράξουν τα παιδιά. Σε αυτήν την ενότητα θα αναφερθούν οι διάφοροι τρόποι που παρέχει ο αλγόριθμος για την επιβίωση των ατόμων, αλλά όσον αφορά τους τελεστές επιλογής που θα εμφανισθούν θα γίνει απλή αναφορά του ονόματός τους, καθώς η λειτουργία τους περιγράφηκε στην αμέσως προηγούμενη ενότητα (5.3.3). Οι επιλογές που παρέχονται είναι τέσσερις και είναι οι εξής:

1. **Επιβίωση μόνο των Απογόνων.** Με αυτήν την επιλογή η νέα γενιά θα αποτελείται αποκλειστικά από τα παιδιά που παράχθηκαν κατά την φάση της αναπαραγωγής. Δε γίνεται κανένας έλεγχος σχετικά με την ποιότητα των ατόμων.
2. **Επιβίωση του Καλύτερου.** Με αυτήν την επιλογή τα άτομα που θα συμμετέχουν στην επόμενη γενιά θα είναι αυτά που θα έχουν καλύτερη ποιότητα από τον γονέα τους. Σε περίπτωση που ο γονέας έχει καλύτερη ποιότητα από το τέκνο του, τότε επιβιώνει εκείνος ενώ ο απόγονός του χάνεται. Με αυτήν την επιλογή δίνουμε τη δυνατότητα σε γονείς που έχουν καλή ποιότητα να επιβιώσουν αρκετές γενιές και να παράγουν περισσότερους απογόνους.

3. **Επιλεγμένη Επιβίωση.** Με αυτή την επιλογή μπορούμε να εφαρμόσουμε τη λειτουργία ενός τελεστή επιλογής (ενότητα 5.3.3) για να επιλέξει εκείνος τα άτομα της νέας γενιάς. Οι διαθέσιμοι τελεστές είναι :
  - a. TournamentSelector
  - b. SequentialSelector
  - c. BestNSelector
  - d. RankingSelector
  - e. RandomSelector
  - f. RouletteWheelSelector
  - g. WorstNSelector
  - h. GreedyOverSelectorΗ λειτουργία των παραπάνω τελεστών περιγράφεται με λεπτομέρειες στην προηγούμενη ενότητα, γι'αυτό εδώ παραλείπεται.
4. **Τυχαία Επιλογή μεθόδου.** Με αυτήν την επιλογή επιλέγεται κάθε φορά με ίση πιθανότητα μία από τις παραπάνω τρεις μεθόδους. Στο τέλος της εξελικτικής διαδικασίας αναμένεται κάθε μέθοδος να έχει εφαρμοστεί σε ποσοστό 33.3%.

## 5.4 Ελιτισμός

Η έννοια του «ελιτισμού» (elitism) έχει εισαχθεί στους εξελικτικούς αλγόριθμους για τη διατήρηση των ατόμων με υψηλή ποιότητα από γενιά σε γενιά. Με την εφαρμογή του ελιτισμού τα άτομα με τη μεγαλύτερη τιμή στη συνάρτηση αξιολόγησης επιβιώνουν πάντα στην επόμενη γενιά. Ο λόγος, όπως αναφέραμε και νωρίτερα, είναι να διατηρούνται τα άτομα που περιέχουν σχήματα υψηλής ποιότητας από γενιά σε γενιά, δίνοντάς τους έτσι τη δυνατότητα να αποκτούν περισσότερους απογόνους. Αυτό σημαίνει πρακτικά πως στη διάρκεια της εξελικτικής διαδικασίας η μέγιστη τιμή της ποιότητας του πληθυσμού δεν μπορεί να οπισθοδρομήσει. Το μέγεθος του ελιτισμού μπορεί να κυμαίνεται αλλά είναι καλό να βρίσκεται σε μικρές τιμές ή χαμηλά ποσοστά, καθώς αν εφαρμόσουμε ελιτισμό σε μεγάλο ποσοστό τότε θα παρατηρηθεί το φαινόμενο της πρόωρης σύγκλισης καθώς τα νέα άτομα που θα παράγονται θα μοιάζουν όλο και περισσότερο σε αυτά που επιβιώνουν λόγω της υπεροχής τους στη συνάρτηση αξιολόγησης με αποτέλεσμα ο πληθυσμός να εκφυλιστεί σύντομα και όλα τα άτομα να είναι ίδια.

Στο αλγόριθμο που παρουσιάζουμε ο ελιτισμός μπορεί να εφαρμοστεί με δύο τρόπους. Ο πρώτος αφορά ποσοστιαία επιλογή των ατόμων που θα αποτελούν την «ελίτ» και ο δεύτερος αριθμητική. Δηλαδή στην πρώτη περίπτωση ορίζουμε ένα ποσοστό  $p\%$ : 2% και κάθε φορά το 2% του πληθυσμού με την υψηλότερη τιμή στη συνάρτηση αξιολόγησης επιβιώνει και μεταφέρεται αυτόματα στην επόμενη γενιά. Στη δεύτερη περίπτωση ορίζουμε το όριο της «ελίτ» στα  $p\%$ : 2 άτομα και κάθε φορά τα δύο καλύτερα άτομα του πληθυσμού επιβιώνουν στην επόμενη γενιά.

Έχει αποδειχτεί πειραματικά ότι η εφαρμογή ελιτισμού στους ΓΑ αποφέρει καλύτερα αποτελέσματα απ'ότι η μη εφαρμογή αυτού.

## 5.5 Αναπροσαρμογή Πιθανοτήτων

Όπως έχουμε παρατηρήσει από τα παραπάνω για την εφαρμογή ενός ΕΑ απαιτείται η παραμετροποίηση ενός πλήθους τελεστών (πχ: γενετικοί, επιλογής, κτλ) πολλοί εκ των οποίων απαιτούν με τη σειρά τους την παραμετροποίηση ενός άλλου συνόλου τελεστών. Τόσο η σωστή επιλογή των τελεστών που θα εφαρμοστούν, όσο και η σωστή παραμετροποίηση αυτών των τελεστών αποτελεί ακόμα «μαύρο κουτί» για τους ερευνητές που ασχολούνται με τους ΕΑ. Η επιλογή και η παραμετροποίηση των τελεστών είναι κυρίως προσανατολισμένη γύρω από κάθε πρόβλημα και για το λόγο αυτό πραγματοποιείται περισσότερο εμπειρικά, διαισθητικά ή βασισμένη σε στατιστικά δεδομένα προηγούμενων πειραμάτων.

Ένα από τα προβλήματα που εμφανίζεται συχνά στο χώρο των ΓΑ είναι και η ποσοστιαία αναλογία της εφαρμογής των τελεστών διασταύρωσης και μετάλλαξης. Όπως έχουμε αναφέρει και νωρίτερα στους ΓΑ ο κύριος γενετικός τελεστής είναι αυτός της διασταύρωσης, ενώ ο τελεστής της μετάλλαξης λειτουργεί στο παρασκήνιο. Παρ'όλ'αυτά όπως γνωρίζουμε οι δύο τελεστές εξερευνούν το χώρο αναζήτησης με διαφορετικό τρόπο, ενώ όπως τονίζουν αρκετοί επιστήμονες ο ρόλος της μετάλλαξης στους ΓΑ έχει υποτιμηθεί αρκετά τις τελευταίες δεκαετίες. Η λύση συνήθως βρίσκεται κάπου στη μέση. Δηλαδή ούτε πρέπει να απορρίψουμε την εφαρμογή της μετάλλαξης στην αναζήτηση, αλλά ούτε και να την υπερτιμήσουμε.

Για το σκοπό αυτό οι τελεστές διασταύρωσης και μετάλλαξης λειτουργούν συμπληρωματικά. Δηλαδή αν η πιθανότητα διασταύρωσης δύο χρωμοσωμάτων είναι  $X\%$  (πχ:  $90\%$ ), τότε η πιθανότητα μετάλλαξης κάθε χρωμοσώματος είναι  $(1-X)\%$  (πχ:  $10\%$ ). Ακόμα όμως και στην περίπτωση που δεχτούμε τη συμπληρωματική εφαρμογή των δύο γενετικών τελεστών είναι άγνωστο, θεωρητικά, ποιές θα αποτελούσαν καλές πιθανότητες για την εφαρμογή τους. Για το λόγο αυτό αναπτύχθηκαν διάφορες μεθοδολογίες οι οποίες κατά τη διάρκεια της εξέλιξης του αλγόριθμου αναπροσαρμόζουν περιοδικά αυτές τις πιθανότητες βάση, κυρίως, στοιχείων που προκύπτουν από το ποσοστό σύγκλισης του πληθυσμού.

Στον αλγόριθμο που αναπτύχθηκε ενσωματώθηκαν μερικές τέτοιες μεθοδολογίες οι οποίες κάνουν χρήση, κυρίως, στατιστικών στοιχείων που έχουν προκύψει από την μέχρι τώρα εξέλιξη του πληθυσμού. Οι μεθοδολογίες αυτές είναι:

1. **Χωρίς Προσαρμογή.** Στην ουσία αυτή η επιλογή δεν αποτελεί κάποια μεθοδολογία, αλλά όπως θα δούμε παρακάτω λόγω της δυνατότητας παράλληλης εξέλιξης πληθυσμών, με διαφορετικές παραμετροποιήσεις, ενδέχεται να μην επιθυμούμε την αναπροσαρμογή των πιθανοτήτων ενός ή περισσότερων πληθυσμών, οπότε τους αναθέτουμε την παραπάνω «μεθοδολογία».
2. **Χρήση Στατιστικών Στοιχείων.** Αυτή η μεθοδολογία χρησιμοποιεί στατιστικά δεδομένα που έχουν προκύψει από τις προηγούμενες γενεές και αναπροσαρμόζει τις δύο πιθανότητες ( $P_{cross}$  και  $P_{mut}$ ) με τον εξής τρόπο:
  - a. Αν το ποσοστό των ατόμων που απέχει ποιοτικά  $1\%$  από τον καλύτερο του πληθυσμού είναι μικρότερο από ένα ποσοστό  $P_{or1}$

- (πχ: 10%) και ταυτόχρονα η πιθανότητα διασταύρωσης  $P_{\text{cross}}$  είναι μικρότερη από το 90%, τότε η πιθανότητα διασταύρωσης,  $P_{\text{cross}}$ , αυξάνει κατά ένα μικρό ποσοστό  $P_{\text{rate}}$  (πχ: 10%), ενώ αντίστοιχα η πιθανότητα μετάλλαξης  $P_{\text{mut}}$  μειώνεται κατά το ίδιο ποσοστό  $P_{\text{rate}}$ .
- b. Αντίθετα, αν το ποσοστό που βρίσκεται κοντά στο 1% του καλύτερου, είναι μεγαλύτερο από ένα άλλο ποσοστό  $P_{\text{op2}}$  (πχ: 50%), τότε σημαίνει ότι ο αλγόριθμος αρχίζει να συγκλίνει επικίνδυνα, οπότε ελέγχουμε ξανά την πιθανότητα  $P_{\text{cross}}$  και αν είναι μεγαλύτερη από πχ: 10%, τότε μειώνουμε την πιθανότητα κατά  $P_{\text{rate}}$ , ενώ αυξάνουμε αντίστοιχα την πιθανότητα του τελεστή μετάλλαξης  $P_{\text{mut}}$  κατά την ίδια ποσότητα.
3. **Έλεγχος Προόδου του Βέλτιστου.** Η λειτουργία αυτής της μεθόδου δεν κάνει χρήση των ποσοστών σύγκλισης του πληθυσμού, κοντά στον καλύτερο κάθε γενιάς, αλλά ελέγχει το πόσο βελτιώθηκε η μέγιστη τιμή της ποιότητας. Αν η βελτίωση υπολείπεται ενός συγκεκριμένου ποσοστού και ταυτόχρονα η πιθανότητα μετάλλαξης είναι μικρότερη από το 90%, τότε η πιθανότητα  $P_{\text{mut}}$  αυξάνει κατά  $P_{\text{rate}}$ , ενώ αντίστοιχα η πιθανότητα  $P_{\text{cross}}$  ελαττώνεται κατά την ίδια ποσότητα. Σε αυτόν τον αλγόριθμο πρέπει να δώσουμε ιδιαίτερη προσοχή γιατί πρέπει να «ζυγίσουμε» την προσδοκόμενη βελτίωση, του μεγίστου της ποιότητας, με βάση το χρονικό διάστημα που θα ελέγχουμε τα στοιχεία και θα αναπροσαρμόζουμε τις πιθανότητες.
  4. **Εναλλαγή Πιθανοτήτων.** Αυτή η μέθοδος δεν χρησιμοποιεί στατιστικά δεδομένα, ούτε εκτελεί κανέναν πολύπλοκο έλεγχο. Για την ακρίβεια αυτή η μέθοδος εναλλάσσει τις πιθανότητες  $P_{\text{cross}}$  και  $P_{\text{mut}}$  κάθε φορά που εφαρμόζεται.
  5. **Τυχαία Επιλογή μεθόδου.** Με αυτή την επιλογή, δεν εφαρμόζεται κάποια νέα μεθοδολογία, αλλά επιλέγεται με ίση πιθανότητα μία από τις προαναφερθείσες μεθόδους αναπροσαρμογής των πιθανοτήτων.

Μελετώντας καλύτερα το θέμα της αναπροσαρμογής των πιθανοτήτων στους ΓΑ μπορούμε να συμπεράνουμε ότι πρόκειται για ένα ολόκληρο πεδίο έρευνας, το οποίο τα τελευταία χρόνια μελετάται όλο και περισσότερο από διάφορους επιστήμονες. Οι μεθοδολογίες που έχουν αναπτυχθεί προσπαθούν να προσαρμόσουν διάφορες έννοιες «ανταμοιβής» των αποδοτικών τελεστών ενσωματώνοντας τις πιθανότητες μέσα στο χρωμόσωμα δίνοντάς τους έτσι τη δυνατότητα να συνεξελίσσονται (co-evolve) με τα υπόλοιπα γονίδια του χρωμοσώματος. Περισσότερες πληροφορίες σχετικά με τις μεθόδους αναπροσαρμογής των πιθανοτήτων στους ΓΑ μπορούμε να βρούμε στο [5], όπου γίνεται μια έρευνα πάνω σε αυτό το θέμα.

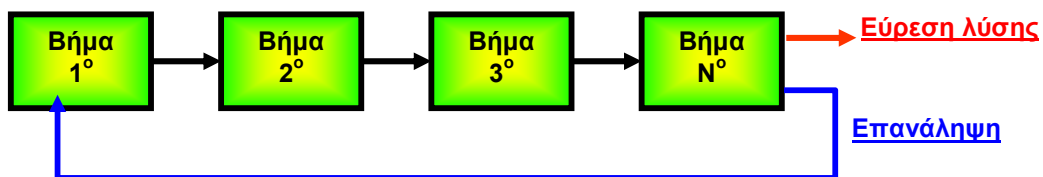
## 5.6 Σενάρια εκτέλεσης Εξελικτικού Αλγορίθμου

Η αρχική θεώρηση του αλγορίθμου περιελάμβανε την υλοποίηση ενός Σειριακού Γενετικού Αλγορίθμου (ΣΓΑ), στον οποίο θα προσαμόζονταν όσο το δυνατόν περισσότεροι γενετικοί τελεστές, για να δωθεί η δυνατότητα της πραγμα-

τοποίησης αρκετών διαφορετικών πειραμάτων. Όσο περισσότερο όμως γινότανε αντιληπτή η παράλληλη φύση των Εξελικτικών Αλγορίθμων τόσο μεγάλωνε και η επιθυμία της υλοποίησης ενός συστήματος που θα εκμεταλλεύεται καλύτερα τις δυνατότητες των ΕΑ στην αναζήτηση της βέλτιστης λύσης. Με αφετηρία λοιπόν αυτήν την επιθυμία ο θεωρούμενος αρχικά ΣΓΑ μετατράπηκε, σχετικά εύκολα, σε ΠΓΑ (Παράλληλο Γενετικό Αλγόριθμο). Φυσικά η έννοια της παραλληλοποίησης αφορούσε αποκλειστικά την προσέγγιση του διαχωρισμού (decomposition approach) και όχι την εφαρμογή του ΣΓΑ σε παράλληλο Η/Υ. Η προσέγγιση του διαχωρισμού περιγράφεται στην ενότητα (2.5.2) και στην παρούσα υλοποίηση του αλγορίθμου εφαρμόστηκε το μοντέλο των νησίδων (island model) ή αλλιώς το μοντέλο του ΠΓΑ χαμηλής ανάλυσης (coarse-grained PGA). Παρ'όλαυτά η δυνατότητα εκτέλεσης ενός και μόνο ΣΓΑ εξακολούθησε να ιφίσταται κανονικά δίνοντας έτσι στον αλγόριθμο μία πλειάδα «σεναρίων» εκτέλεσης. Αυτά τα διαφορετικά «σενάρια» εκτέλεσης του ΕΑ περιγράφονται στις αμέσως επόμενες ενότητες.

### 5.6.1 Ένας πληθυσμός – Μία παραμετροποίηση

Η πρώτη και απλούστερη εκτέλεση του ΓΑ αφορά την υλοποίηση ενός ΣΓΑ. Σε αυτό το «σενάριο» έχουμε την δημιουργία ενός πληθυσμού με ενιαία παραμετροποίηση.



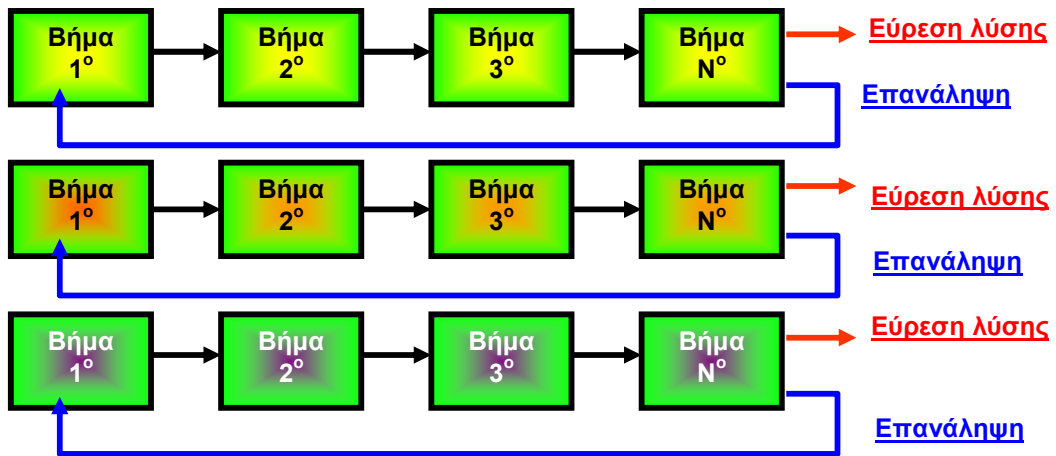
Σχήμα 5.8. Α σενάριο εκτέλεσης ΕΑ.

Όπως βλέπουμε και στο παραπάνω σχήμα τα βήματα του αλγορίθμου διαδέχονται το ένα το άλλο και στο τελικό βήμα (Βήμα N<sup>ο</sup>) είτε έχουμε βρεί λύση, οπότε και σταματάμε την εξέλιξη, είτε συνεχίζουμε πάλι από το πρώτο βήμα.

### 5.6.2 ‘N’ πληθυσμοί – ‘N’ παραμετροποιήσεις

Σε αυτήν την υποκατηγορία έχουμε τη δυνατότητα εφαρμογής τριών διαφορετικών «σεναρίων» εκτέλεσης. Αυτά είναι τα εξής:

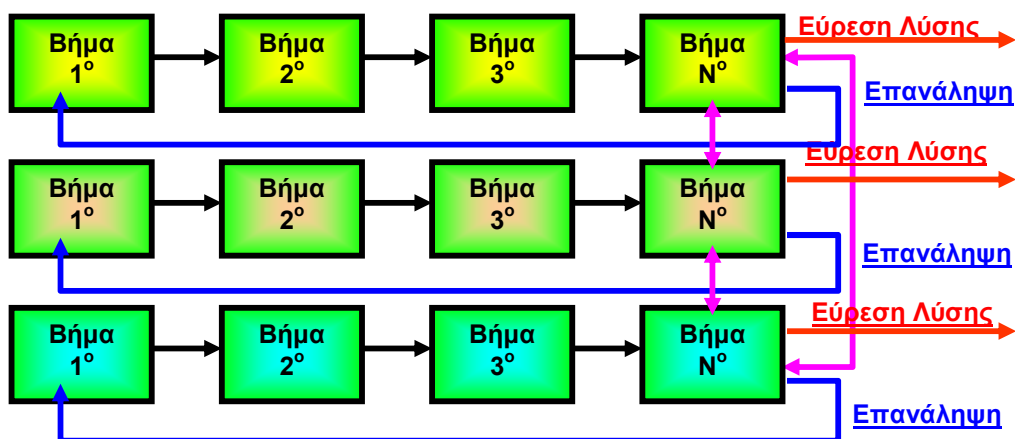
1. **Παράλληλη εξέλιξη πληθυσμών, χωρίς επικοινωνία.** Αυτή η επιλογή εκτέλεσης αποτελεί, στην ουσία, μία γενίκευση του προηγούμενου σεναρίου (5.6.1). Δηλαδή έχουμε την ταυτόχρονη εξέλιξη ‘N’ διαφορετικών πληθυσμών με ξεχωριστές παραμετροποιήσεις, χωρίς όμως να έχουν καμία επικοινωνία μεταξύ τους.



Σχήμα 5.9. Β σενάριο εκτέλεσης ΕΑ.

Όπως βλέπουμε και στο παραπάνω σχήμα κάθε πληθυσμός (χρωματίζεται με διαφορετικό χρώμα στο κέντρο των τετραγώνων) εξελίσσεται ταυτόχρονα με τους υπόλοιπους χωρίς όμως να έχουνε καμία επικοινωνία μεταξύ τους. Ο τερματισμός ενός πληθυσμού (δηλαδή η εύρεση λύσης) δεν διακόπτει την εξέλιξη των υπολοίπων. Κάθε πληθυσμός μπορεί να έχει τα δικά του κριτήρια τερματισμού.

2. **Παράλληλη εξέλιξη πληθυσμών, με επικοινωνία.** Αυτό το «σενάριο» αποτελεί την ουσιαστική υλοποίηση του μοντέλου του διαχωρισμού. Με αυτή την επιλογή έχουμε κάθε πληθυσμό, ο οποίος δύναται να έχει εντελώς ξεχωριστή παραμετροποίηση από του υπολοίπους, να αποτελεί μία ξεχωριστή νησίδα. Περιοδικά οι πληθυσμοί ανταλλάσσουν μεταξύ τους άτομα (δηλαδή γενετικό υλικό) και κατόπιν συνεχίζουν πάλι την εξέλιξη τους.



Σχήμα 5.10. Γ σενάριο εκτέλεσης ΕΑ.

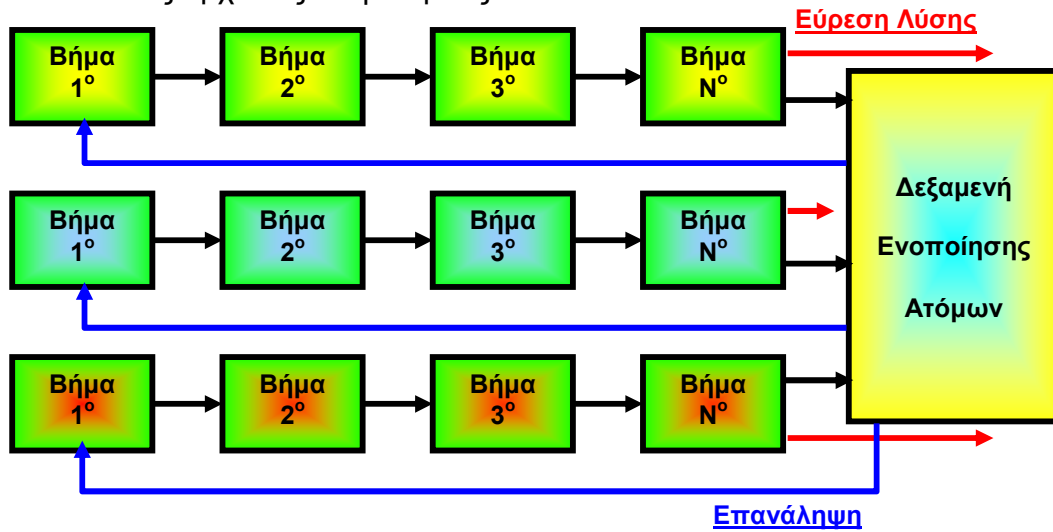
Όπως βλέπουμε και στο παραπάνω σχήμα (5.10), με πορφυρό χρώμα εμφανίζονται τα βέλη που δείχνουν τις πιθανές μετακινήσεις ατόμων μεταξύ των παράλληλα εξελισσόμενων πληθυσμών. Κοινές παράμετροι για όλους τους πληθυσμούς αποτελούν: **α)** η μέθοδος μετακίνησης και **β)** το χρονικό διάστημα που θα μεσολαβεί μεταξύ δύο διαδοχικών μετακινήσεων, πχ: 100 γενιές (γνωστό και ως «διάστημα απομόνωσης» των πληθυσμών).

Στις μεθόδους μετακίνησης ο χρήστης μπορεί να επιλέξει μία από τις τέσσερις παρεχόμενες μεθόδους: **1)** Όλοι σε όλους, **2)** Σε τυχαίο προορισμό, **3)** Μόνο στο διπλανό (Κυκλική Μετακίνηση) και **4)** Τυχαία επιλογή μεθόδου. Στην πρώτη επιλογή (1), που αποτελεί και την προεπιλεγμένη μέθοδο, κάθε πληθυσμός (νησίδα) στέλνει και δέχεται άτομα (χρωμοσώματα) από και προς όλους τους πληθυσμούς που βρίσκονται σε εξέλιξη (βλέπε σχήμα 2.6.α, ενότητα 2.5.2). Δηλαδή αν κάποιος πληθυσμός τερματίσει την εξελικτική του διαδικασία, επειδή ικανοποιήθηκε κάποιο κριτήριο τερματισμού, δε συμμετέχει πλέον στις μετακινήσεις χρωμοσωμάτων. Στη δεύτερη επιλογή (2) κάθε πληθυσμός επιλέγει με τυχαίο τρόπο έναν άλλο πληθυσμό και ανταλλάσει με αυτόν χρωμοσώματα. Στην τρίτη επιλογή (3) οι πληθυσμοί δημιουργούν έναν νοητό κύκλο και κάθε πληθυσμός στέλνει άτομα σε αυτόν που βρίσκεται στη δεξιά του πλευρά και δέχεται άτομα από αυτόν που βρίσκεται στην αριστερή του πλευρά (βλέπε σχήμα 2.6.β, ενότητα 2.5.2). Η τέταρτη και τελευταία επιλογή (4) επιλέγει με τυχαίο τρόπο μία από τις τρεις προηγούμενες, κάθε φορά που πρόκειται να γίνει μετακίνηση ατόμων. Η επιλογή των μεθόδων γίνεται ισοπίθανα, με αποτέλεσμα στο τέλος της διαδικασίας όλες οι μέθοδοι να έχουν εφαρμοστεί περίπου ίσες φορές.

Από την άλλη μεριά κάθε πληθυσμός είναι ελεύθερος να επιλέγει με ξεχωριστό τρόπο τόσο το πλήθος, όσο και την ποιότητα των ατόμων που θα δέχεται και θα αποστέλει. Και στις δύο περιπτώσεις παρέχονται αρκετοί από τους τελεστές επιλογής, που ήδη προαναφέρθηκαν, για την επιλογή τόσο των ατόμων που θα αποδημίσουν (emigrants), όσο και αυτών που θα εποικίσουν (immigrants). Και σε αυτό το σενάριο εξέλιξης η ικανοποίηση ενός κριτηρίου τερματισμού από έναν πληθυσμό (δηλαδή η εύρεση λύσης) δεν τερματίζει και την συνολική εξέλιξη των υπολοίπων πληθυσμών, απλώς σταματάει τη δική του εξέλιξη και φυσικά δεν συμμετέχει πλέον στις ανταλλαγές των ατόμων. Θα πρέπει να σημειώσουμε ότι η ανταλλαγή ατόμων μεταξύ των πληθυσμών συνεχίζεται μέχρις ότου μείνει μόνο ένας, εν εξέλιξη, πληθυσμός.

- 3. Παράλληλη εξέλιξη πληθυσμών, με ενοποίηση στο τελικό βήμα.** Αυτή η επιλογή εξέλιξης, αποτελεί μία «παραποίηση» του προηγούμενου σεναρίου εξέλιξης. Και εδώ έχουμε ένα σύνολο πληθυσμών οι οποίοι εξελίσσονται παράλληλα, αλλά αντί για οργανωμένα και περιοδική ανταλλαγή γενετικού υλικού έχουμε τυχαία και συνεχή ανταλλαγή ατόμων. Δηλαδή στο τέλος κάθε γενιάς τα άτομα από όλους τους πληθυσμούς συλ-

λέγονται σε μία «δεξαμενή» και κατόπιν ανακατανέμονται με τυχαίο τρόπο πίσω στους αρχικούς πληθυσμούς.



Σχήμα 5.11. Ένα σενάριο εκτέλεσης ΕΑ.

Και στο παραπάνω σχήμα τα βέλη με κόκκινο χρώμα υποδεικνύουν την εύρεση λύσης σε κάθε πληθυσμό, με μπλέ χρώμα τις επιστροφές των ατόμων πίσω στους πληθυσμούς για την εκκίνηση της επόμενης γενιάς και τα βέλη με μαύρο χρώμα υποδεικνύουν τις μεταβάσεις του αλγορίθμου από το ένα βήμα στο άλλο.

Θα πρέπει ακόμα να αναφέρουμε ότι οι παράλληλα εξελισσόμενοι πληθυσμοί δεν είναι υποχρεωτικό να είναι ισάριθμοι. Επίσης επειδή οι πληθυσμοί εξελίσσονται σε σειριακό Η/Υ η συνολική ταχύτητα του αλγορίθμου θα εξαρτάται πάντα από την πιο «αργή» παραμετροποίηση. Ακόμα, στην περίπτωση όπου έχουμε ανταλλαγή γενετικού υλικού, με οργανωμένο τρόπο (ΈΓ σενάριο εκτέλεσης – σχήμα 5.10), ενδέχεται να έχουμε αυξομειώσεις των ατόμων των πληθυσμών, αν η παραμετροποίηση των τελεστών μετακίνησης δε γίνει με σωστό τρόπο. Τέλος κάθε πληθυσμός θα αξιολογείται από την ίδια συνάρτηση αξιολόγησης (evaluation function), αλλά τόσο το πλήθος των περιορισμών που θα ικανοποιεί, όσο και οι τιμές της βαρύτητας κάθε περιορισμού μπορούν να διαφέρουν από πληθυσμό σε πληθυσμό.

## 5.7 Προσομοίωση Πολέμου μεταξύ πληθυσμών

Η προσομοίωση της πολεμικής κατάστασης εφαρμόζεται πειραματικά και προσπαθεί να προσομοιώσει την επιβολή ενός ισχυρότερου πληθυσμού σε άλλους πιο αδύναμους πληθυσμούς. Με την έννοια του ισχυρού εννοούμε αυτόν τον πληθυσμό που είναι καλύτερα προσαρμοσμένος στις συνθήκες του περιβάλλοντός του, το οποίο εκφράζεται άμεσα με την τιμή που αποδίδει η συνάρτηση αξιολόγησης (evaluation function) σε κάθε χρωμόσωμα. Τα είδη προσομοίωσης πολέμου που παρέχει ο συγκεκριμένος αλγόριθμος είναι τρία : α)



Εμφύλιος, β) Πρόκληση και γ) Παγκόσμιος πόλεμος. Η προσπάθεια που έγινε για την υλοποίηση αυτού του τμήματος του αλγορίθμου επικεντρώθηκε στην όσο το δυνατόν λιγότερη παραμετροποίηση, από πλευράς χρήστη, χωρίς όμως να χαθούν πολλές δυνατότητες της προσομοίωσης.

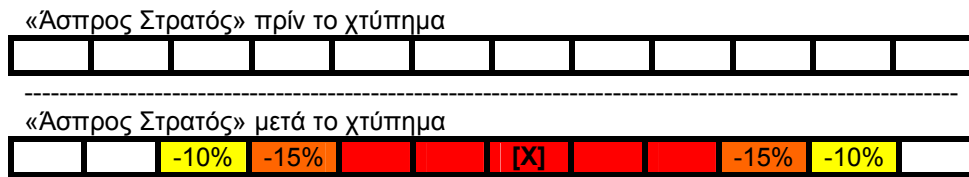
Παρακάτω περιγράφονται αναλυτικά και τα τρία είδη πολέμου:

1. **Εμφύλιος Πόλεμος (Civil War)**. Στον εμφύλιο πόλεμο κάθε πληθυσμός χωρίζεται σε δύο, ισάριθμα, αντίπαλα στρατόπεδα (το «Άσπρο» και το «Μαύρο»)<sup>4</sup>. Η επιλογή των ατόμων, κάθε στρατοπέδου, γίνεται τυχαία όπως επίσης και η επιλογή των στρατιωτών κάθε αντιμαχόμενης παράταξης. Το ποσοστό των ατόμων κάθε παράταξης που θα αποτελούν τους στρατιώτες είναι το ίδιο και στα δύο στρατόπεδα και ορίζεται από αυτόν που θα κάνει την παραμετροποίηση του ΕΑ. Έτσι, για παράδειγμα, αν κάθε στρατόπεδο έχει πενήντα (50) άτομα και το ποσοστό είναι 17%, τότε από κάθε πλευρά θα έχουμε  $5 \cdot 17 / 100 = 9$  στρατιώτες. Η μάχη διεξάγεται σε δύο φάσεις, που είναι σχεδόν πανομοιότυπες και στα άλλα δύο είδη πολέμου.

a. **«Βομβαρδισμοί»**. Η φάση του «βομβαρδισμού» είναι αυτή η οποία προηγείται πάντα και εφαρμόζεται κυρίως για λόγους δικαιοσύνης, σε περιπτώσεις όπου οι δύο αντίπαλοι στρατοί έχουν μεγάλη αριθμητική διαφορά. Έτσι για κάθε δέκα στρατιώτες η αντίπαλη πλευρά δικαιούται μία βόμβα. Κάθε «βόμβα» προκαθορίστηκε να επιδρά σε εννιά (9) άτομα του αντίπαλου στρατού. Για να καταλάβουμε καλύτερα τον τρόπο λειτουργίας αυτής της φάσης του πολέμου θα δούμε το παρακάτω παράδειγμα, μαζί με το σχήμα 5.12. Αν θεωρήσουμε ότι τα δύο αντίπαλα στρατόπεδα έχουν το μεν «Άσπρο» 12 στρατιώτες και το δε «Μαύρο» 7, τότε σύμφωνα με τον παραπάνω κανόνα το «Μαύρο» στρατόπεδο δικαιούται μία βόμβα ( $12 / 10 = 1$ ), ενώ το «Άσπρο» καμία ( $7 / 10 = 0$ ). Σημειώνεται ότι για την εύρεση του πλήθους των βομβών εφαρμόζεται η ακέραια διαίρεση του πλήθους των στρατιωτών με τον αριθμό δέκα (10). Έπειτα θεωρούμε τους στρατιώτες κάθε στρατεύματος παρατεταγμένους εκατέρωθεν σε μία ευθεία γραμμή. Η τοποθέτηση των στρατιωτών σε κάθε πλευρά γίνεται με τυχαίο τρόπο και δεν παίζει κανένα ρόλο η ποιότητα του χρωμοσώματος. Επιλέγεται, επίσης τυχαία, το επίκεντρο του χτυπήματος και τότε η θέση που χτυπήθηκε σύν / πλύν δύο θέσεις δεξιά και αριστερά του επικέντρου διαγράφονται οριστικά από τη μάχη, ενώ η ποιότητα σύν / πλύν τεσσάρων θέσεων χάνει ένα ποσοστό από την ποιότητά του θεωρώντας ότι «τραυματίστηκαν» από το χτύπημα.

---

<sup>4</sup> Οι ονομασίες «Άσπρο» και «Μαύρο» προέρχονται από την επιρροή του παιχνιδιού στρατηγικής < Σκάκι >.



Σχήμα 5.12. Παράδειγμα εφαρμογής βομβαρδισμού.

Όπως διακρίνουμε και στο παραπάνω σχήμα με το σύμβολο «[X]» σημαδεύουμε το επίκεντρο του χτυπήματος. Με κόκκινο χρώμα τσεκάρονται τα χρωμοσώματα που θα διαγραφούν οριστικά, με πορτοκαλί χρώμα αυτά που έχασαν το 15% της ποιότητάς τους, λόγω του χτυπήματος, και με κίτρινο χρώμα αυτοί που έχασαν το 10% από την ποιότητά τους. Οπότε το πλήθος του «Άσπρου» στρατού που θα συμμετέχει στην επόμενη φάση θα είναι επτά άτομα ( $12 - 5 = 7$ ). Θα πρέπει να επισημάνουμε πως το πλήθος των βομβών από κάθε πλευρά αποφασίζεται μία φορά στην αρχή. Αν μετά το πέρας των «βομβαρδισμών» προκύψει κάποιος πληθυσμός του οποίου το πλήθος είναι πάνω από δέκα άτομα, τότε η μάχη συνεχίζεται κανονικά στη δεύτερη φάση και δεν γίνονται εκ νέου βομβαρδισμοί.

- b. **«Μάχη Σώμα με Σώμα»** Η δεύτερη φάση της μάχης περιλαμβάνει «μάχη σώμα με σώμα». Επιλέγεται ένας στρατιώτης από κάθε στρατό και αυτός με τη μεγαλύτερη τιμή στην ποιότητα κερδίζει. Ο ηττημένος διαγράφεται οριστικά, ενώ ο νικητής συνεχίζει κανονικά αλλά «λαβωμένος». Δηλαδή χάνει από την ποιότητά του το σύνολο της ποιότητας του αντιπάλου που κέρδισε. Έτσι αν ο νικητής είχε αρχικά τιμή στην ποιότητά του ίση με 0.9 και ο ηττημένος ίση με 0.5, τότε ο νικητής θα συνεχίσει τον πόλεμο με ποιότητα 0.4. Στην περίπτωση που οι δύο αντίπαλοι έχουν ακριβώς την ίδια ποιότητα, τότε ο νικητής επιλέγεται στην τύχη, με ίση πιθανότητα και για τους δύο, ενώ η ποιότητα του νικητή ορίζεται μόλις στο δέκα τοις εκατό της αρχικής του ποιότητας. Δηλαδή αν είχε ποιότητα ίση με 0.75, θα συνέχισι τη μάχη με ποιότητα 0.075. Η παραπάνω διαδικασία επαναλαμβάνεται μέχρις ότου το ένα από τα δύο στρατεύματα εξοντωθεί ολοκληρωτικά. Τότε έχουμε και το τέλος του πολέμου.

Μετά το πέρας του πολέμου ακολουθεί η φάση κατά την οποία συλλέγονται πάλι όλα τα άτομα που απέμειναν και από τις δύο παρατάξεις, στρατιώτες και μη, και συνεχίζουν κανονικά την εξέλιξη.

2. **Πρόκληση (Challenge War)**. Σε αυτό το είδος του πολέμου αρχικά συλλέγονται όλοι οι πληθυσμοί που θα συμμετέχουν στον πόλεμο και παρατάσσονται σε μία «νοητή» σειρά με τυχαίο τρόπο. Κατόπιν κάθε πληθυσμός έχει το δικαίωμα να προκαλέσει μία και μόνο μία φορά σε πόλεμο έναν άλλο πληθυσμό. Μεταξύ αυτών των δύο πληθυσμών

διεξάγεται κανονικά πόλεμος όπως περιγράφηκε και παραπάνω, με την εξής όμως διαφορά. Η επιλογή των στρατιωτών κάθε πληθυσμού δεν γίνεται απαραίτητα με τυχαίο τρόπο, αλλά δίνεται η δυνατότητα στους εύρωστους πληθυσμούς να επιλέξουν τους στρατιώτες τους από την «ελίτ» τους αν και μόνο αν το ποσοστό των ατόμων του πληθυσμού που βρίσκεται κοντά στο 1% επί του καλύτερου ατόμου, είναι μεγαλύτερο από το ποσοστό επί του συνολικού πληθυσμού που θα αποτελούν τους στρατιώτες. Δηλαδή αν το ποσοστό του πληθυσμού που πρέπει να «στρατευτεί» είναι 15% και η «ελίτ» του πληθυσμού είναι 25%, τότε οι στρατιώτες θα επιλεγούν αποκλειστικά από την «ελίτ». Σε αντίθετη περίπτωση θα επιλεγούν με τυχαίο τρόπο. Ο νικητής του πολέμου θα ενσωματώσει στον δικό του πληθυσμό όλα τα άτομα που απέμειναν από τον ηττημένο. Η παραπάνω διαδικασία επαναλαμβάνεται μέχρις ότου όλοι οι πληθυσμοί προκαλέσουν από μία φορά κάποιον άλλο πληθυσμό, ή αν μείνει μόνο ένας πληθυσμός. Στο τέλος οι εναπομείναντες πληθυσμοί συνεχίζουν κανονικά την εξέλιξή τους.

- 3. Παγκόσμιος (World War).** Αυτό το είδος του πολέμου παρουσιάζει και το μεγαλύτερο ενδιαφέρον αν και ενδείκνυται μόνο για πειράματα που περιλαμβάνουν πολλούς συνεξελίσσαντες πληθυσμούς. Σε αυτή την περίπτωση συλλέγονται όλοι οι πληθυσμοί που θα συμμετέχουν στον πόλεμο και χωρίζονται σε δύο αντίπαλα στρατόπεδα με τυχαίο τρόπο. Το πλήθος των πληθυσμών από κάθε πλευρά είναι περίπου το ίδιο (πχ: 3 – 3, 4 – 5 , κτλ) χωρίς όμως να λαμβάνεται υπόψιν το μέγεθος κάθε πληθυσμού. Κατόπιν με βάση το ίδιο ποσοστό κάθε πληθυσμός συνεισφέρει στα κοινά στρατεύματα με τον τρόπο που περιγράφηκε λίγο νωρίτερα στο '2' (*Πρόκληση*). Δηλαδή και εδώ δίνουμε τη δυνατότητα στους εύρωστους πληθυσμούς να χρησιμοποιήσουν στρατεύματα από την «ελίτ» τους. Αφού μαζευτούν τα στρατεύματα από κάθε πλευρά τότε διαξάγεται κανονικά η μάχη σε δύο φάσεις, όπως περιγράφηκε στο '1' (*Εμφύλιος Πόλεμος*). Και εδώ ο πόλεμος τελειώνει όταν ο ένας από τους δύο στρατούς («Ασπρος» ή «Μαύρος») χάσει όλα τα στρατεύματά του. Το χαρακτηριστικό αυτού του πολέμου είναι ότι με το πέρας των μαχών οι ηττημένοι πληθυσμοί συλλέγονται όλοι μαζί σε εννιαίο χώρο και κατόπιν μοιράζονται στους νικητές πληθυσμούς, σαν λάφυρα, με βάση το ποσοστό που είχε συνεισφέρει κάθε πληθυσμός στα κοινά στρατεύματα. Δηλαδή αν το σύνολο των ηττημένων είναι πχ: 100 άτομα και οι νικητές είναι τρεις, τότε δεν θα μοιραστούν ισόποσα τα εκατό άτομα (δηλαδή από 33 ο κάθε ένας), αλλά ανάλογα με το πόσα στρατεύματα είχαν συνεισφέρει στην αρχή του πολέμου. Δηλαδή αν έχουμε τρεις νικητές (πχ: Α, Β, Γ) και τα ποσοστά επί του συνολικού στρατού ήταν πχ: 30% από τον 'Α', 50% από τον 'Β' και 20% από τον 'Γ', τότε ο 'Α' πληθυσμός θα πάρει το 30% των ηττημένων, ο 'Β' το 50% και ο 'Γ' το υπόλοιπο 20%. Κατόπιν οι νικητές του πολέμου συνεχίζουν κανονικά της εξελικτικής τους διαδικασία έχοντας στο δυναμικό τους και τα επιπλέον άτομα που πήραν σαν «λάφυρα».

Οι παράμετροι που μπορούμε να ορίσουμε σε αυτές τις προσομοιώσεις είναι **α)** το χρονικό διάστημα που θα μεσολαβεί μεταξύ δύο διαδοχικών πολεμικών συγκρούσεων, **β)** το είδος του πολέμου που θα διεξαχθεί (με προεπιλεγμένο είδος τον «Εμφύλιο»), **γ)** το ποσοστό επί του πληθυσμού που θα αποτελούν τους στρατιώτες κάθε πληθυσμού (με αρχική τιμή το 15%) και **δ)** την πιθανότητα να συμμετάσχει ένας πληθυσμός στον πόλεμο (με προεπιλεγμένη τιμή το 50%). Οι τρεις πρώτες παράμετροι είναι κοινές για όλους τους συμμετέχοντες πληθυσμούς, ενώ μόνο η τελευταία μπορεί να διαφέρει από πληθυσμό σε πληθυσμό.

## 5.8 Προσομοίωση Ίωσης πληθυσμού

Η προσομοίωση της «ίωσης» ενός πληθυσμού, αφορά κυρίως πολυπληθείς πληθυσμούς οι οποίοι εξελίσσονται για αρκετά μεγάλο χρονικό διάστημα. Η ιδέα αυτής της προσομοίωσης προήρθε από την παρατήρηση ότι, συνήθως, στα φυσικά περιβάλλοντα μετά από μεγάλους και καταστροφικούς πολέμους σειρά παίρνουν οι ασθένειες που επιδρούν ακόμα πιο αρνητικά στα μέλη του περιβάλλοντος. Η ασθένιση των πληθυσμών πραγματοποιείται περιοδικά, συνήθως μετά από μεγάλα χρονικά διαστήματα, και μπορεί να συμβεί με βάση μία πιθανότητα η οποία θα καθορίσει αν τελικά ο πληθυσμός θα νοσήσει ή όχι.

Η επίδραση του φαινομένου της ίωσης, σε έναν πληθυσμό, επιφέρει καταστροφικές, για αυτόν, συνέπειες καθώς δε γίνεται καμία διάκριση στα άτομα που θα νοσήσουν, με αποτέλεσμα να υπάρχει η πιθανότητα να πληγούν άτομα υψηλής ποιότητας. Η πραγματοποίηση της ίωσης μπορεί να συμβεί με δύο τρόπους και πλήττει ένα ποσοστό του πληθυσμού. Οι τρόποι ίωσης του πληθυσμού είναι δύο:

1. **Όλοι οι νοσήσαντες θα έχουν την ίδια ασθένεια.** Με τον όρο ασθένεια εννοούμε μία «μολυσματική», αλλά επιτρεπτή, τιμή γονιδίου η οποία θα αντικαταστήσει κάθε τιμή σε ένα ποσοστό των γονιδίων του χρωμοσώματος. Δηλαδή επιλέγεται, τυχαία, μία επιτρεπτή τιμή και τοποθετείται σε ένα μικρό ποσοστό των γονιδίων. Τα γονίδια που θα αλλάξουν τιμή επιλέγονται επίσης τυχαία.

Υγιές χρωμόσωμα :



Μολυσματική (αλλά επιτρεπτή) τιμή :



Μολυσμένο χρωμόσωμα :



Σχήμα 5.13. Μόλυνση χρωμοσώματος σε ποσοστό 40%.

2. **Κάθε άτομο θα έχει ξεχωριστή ασθένεια.** Αυτή η επιλογή ίωσης είναι παρόμοια με την προηγούμενη με την διαφορά ότι κάθε φορά που επιλέγεται το άτομο που θα νοσήσει παράγεται μία νέα μολυσματική τιμή.

Οι παράμετροι που μπορούμε να ορίσουμε σε αυτήν την προσομοίωση είναι: **α)** το χρονικό διάστημα που, περιοδικά, θα νοσούν οι πληθυσμοί, **β)** η πιθανότητα να ιωθεί κάθε πληθυσμός (με αρχική τιμή το 20%), **γ)** ο τύπος ίωσης κάθε πληθυσμού ( με προεπιλεγμένο τύπο αυτόν όπου όλοι θα έχουν τον ίδιο ιό), **δ)** το ποσοστό του πληθυσμού που θα πληγεί από την ασθένεια (με προεπιλεγμένη τιμή το 10%) και **ε)** το ποσοστό κάθε χρωμοσώματος που θα μολύνει ο «ιός» (με αρχική τιμή το 20%). Από τις παραπάνω παραμέτρους οι δύο πρώτες (**α** και **β**) είναι κοινές για όλους τους συνεξελισσόμενους πληθυσμούς, ενώ οι υπόλοιπες τρεις μπορούν να διαφέρουν από πληθυσμό σε πληθυσμό.

## 5.9 Κριτήρια τερματισμού εξέλιξης

Τα κριτήρια τερματισμού της εξελικτικής διαδικασίας είναι πέντε και στην περίπτωση των παράλληλα εξελισσόμενων πληθυσμών τα τρία από αυτά μπορούν να παραμετροποιηθούν διαφορετικά σε κάθε πληθυσμό, έτσι ώστε κάθε πληθυσμός να τερματίζει με τον δικό του τρόπο. Τα κριτήρια τερματισμού παρατίθενται και περιγράφονται παρακάτω:

1. **Εύρεση Βέλτιστης Λύσης.** Το κριτήριο αυτό αφορά την εύρεση ενός προγράμματος εξετάσεων το οποίο ικανοποιεί πλήρως όλους τους περιορισμούς ικανοποίησης και η τιμή της αξιολόγησής του είναι η μονάδα (1). Η εύρεση της βέλτιστης λύσης σε μία εξελικτική διαδικασία εξαρτάται από πάρα πολλά πράγματα όπως το αρχικό μέγεθος του πληθυσμού, το πλήθος και η αυστηρότητα των περιορισμών ικανοποίησης, η παραμετροποίηση των γενετικών τελεστών, το μέγιστο χρονικό διάστημα της αναζήτησης (μέγιστο πλήθος γενεών εξέλιξης), και άλλα. Θα πρέπει επίσης να αναφέρουμε ότι ανάλογα με τα δεδομένα, εισαγωγής, του αλγορίθμου η βέλτιστη λύση είναι πολύ πιθανό να μην υπάρχει. Δηλαδή στο χώρο αναζήτησης του προβλήματος να μην υπάρχει βέλτιστο με τιμή τη μονάδα, αλλά το ολικό βέλτιστο να είναι κατά πολύ μικρότερο. Αυτό το φαινόμενο σε πολλές περιπτώσεις είναι πιθανό να το γνωρίζουμε εκ των προτέρων, ενώ σε άλλες περιπτώσεις μπορεί να μην είναι φανερό.
2. **Μέγιστο πλήθος γενεών εξέλιξης.** Αυτό το κριτήριο αφορά το μέγιστο χρονικό διάστημα που επιτρέπουμε στον αλγόριθμο να αναζητά την καλύτερη δυνατή λύση του προβλήματος. Σαν προεπιλεγμένη τιμή τοποθετήθηκε η χίλια (1000).
3. **Μέγιστο πλήθος γενεών, με σταθερή τιμή της ποιότητας.** Επειδή ο χώρος αναζήτησης σε προβλήματα συνδυαστικής βελτιστοποίησης, όπως είναι και το πρόβλημα του χρονοπρογραμματισμού των εξετάσεων, είναι τεράστιος και δεν υπάρχει δυνατότητα με τη σημερινή τεχνολογία να αναζητηθούν όλες οι πιθανές λύσεις του χώρου, σε λογικό χρονικό διάστημα, οι ΕΑ κατευθύνουν την αναζήτησή τους σε ποιοτικά καλύτερες περιοχές του χώρου με τη βοήθεια τόσο των γενετικών τελεστών, όσο και των τελεστών επιλογής. Παρ'όλ'αυτά όμως αν φανταστούμε το χώρο

αναζήτησης σε τρεις διαστάσεις (3Δ) να περιλαμβάνει βουνά, πεδιάδες και χαράδρες είναι πολύ πιθανό ο ΕΑ να αναζητάει τη λύση σε ένα επίπεδο μέρος χωρίς να μπορεί να προσδιορίσει με βεβαιότητα κάποια περιοχή όπου η ποιότητα είναι καλύτερη. Σε αυτές τις περιπτώσεις ο αλγόριθμος παλινδρομεί («κολλάει») και δεν μπορεί να ξεφύγει από αυτές τις περιοχές του χώρου. Οπότε αν τον αφήσουμε σε μια τέτοια κατάσταση ο ΕΑ θα παραμείνει αενάως στο ίδιο μέρος. Για να αποφύγουμε λοιπόν αυτές τις δυσάρεστες καταστάσεις ορίζουμε ένα άνω όριο στο χρονικό διάστημα που μπορεί ο ΕΑ να παραμείνει στην ίδια, ποιοτικά, θέση. Η προεπιλεγμένη τιμή αυτού του κριτηρίου είναι συνήθως το ήμισυ του μέγιστου πλήθους γενεών εξέλιξης, δηλαδή πεντακόσια (500).

4. **Μέγιστο πλήθος αξιολογήσεων.** Αυτό το κριτήριο τερματισμού επιτρέπει σε έναν πληθυσμό να αξιολογήσει ένα συγκεκριμένο πλήθος ατόμων. Συνήθως η τιμή αυτή είναι ακέραιο πολλαπλάσιο του μεγέθους του πληθυσμού, δηλαδή αν ένας πληθυσμός αριθμεί πενήντα (50) άτομα τότε μία πιθανή τιμή σε αυτό το κριτήριο θα ήταν χίλιες πεντακόσιες (1500) αξιολογήσεις. Δηλαδή ο πληθυσμός θα αξιολογείται για τριακόσιες (300) γενεές. Η προεπιλεγμένη τιμή τοποθετείται δυναμικά και υπολογίζεται κάθε φορά από το γινόμενο του μεγέθους του πληθυσμού επί του μέγιστου πλήθους γενεών εξέλιξης.
5. **Ελάχιστη απόσταση από τη βέλτιστη λύση.** Στην ουσία το κριτήριο αυτό αποτελεί το «ελάχιστο σφάλμα». Δηλαδή είναι μία προσέγγιση του πόσο κοντά στη βέλτιστη λύση θεωρούμε ότι μία λύση μας ικανοποιεί. Έτσι, για παράδειγμα, αν ορίσουμε αυτήν την τιμή στο 0.05 (5%), τότε θεωρούμε ότι οι λύσεις κοντά στο 0.95 (95%) είναι αρκετά ικανοποιητικές και επομένως με την εύρεση μιάς τέτοιας λύσης σταματάμε την περαιτέρω εξέλιξη του πληθυσμού. Η προεπιλεγμένη τιμή αυτού του κριτηρίου είναι το μηδέν (0.0).

# ΚΕΦΑΛΑΙΟ 6

## Αποτελέσματα

Στο παρόν κεφάλαιο παρουσιάζονται μερικές δοκιμαστικές εφαρμογές του ΕΑ που έγιναν, σε πειραματικά δεδομένα, για να αποδειχτεί και στην πράξη τόσο η χρησιμότητά του, όσο και η αποτελεσματικότητά του στην εύρεση ποιοτικών προγραμμάτων εξεταστικής περιόδου. Στις ενότητες που ακολουθούν παρουσιάζονται τα πειράματα που διεξήχθησαν, καθώς και τα αποτελέσματα που προέκυψαν από αυτές τις δοκιμές. Τόσο οι παραμετροποιήσεις των ΕΑ που εφαρμόστηκαν, όσο και τα στατιστικά δεδομένα που συνελέχθησαν θα παρουσιαστούν με συντομία, καθώς τα πλήρη δεδομένα θα παραδωθούν σε συμπαγή δίσκο (CD) που θα συνοδεύει την πτυχιακή εργασία.

### 6.1 Χαρακτηριστικά πειραμάτων

Τα πειράματα που διεξήχθησαν είχαν σκοπό, όχι τόσο να βρουν ένα πλήρως ικανοποιητικό πρόγραμμα εξετάσεων, όσο να επιβεβαιώσουν ότι ο αλγόριθμος που αναπτύχθηκε ήταν πραγματικά αποτελεσματικός και επιβεβαίωσε τις προσδοκίες του γράφοντος για σωστή λειτουργία και ικανοποιητικά αποτελέσματα.

Όλα τα πειράματα έκαναν χρήση του ίδιου συνόλου δεδομένων, που το αποτελούσαν μία δοκιμαστική Βάση Δεδομένων (Β.Δ). Σε αυτήν τη Β.Δ περιέχονταν εκατόν πενήντα τέσσερις (154) φοιτητές, πενήντα δύο (52) μαθήματα (που διδάσκονται στο τμήμα Πληροφορικής), εννέα (9) αίθουσες διδασκαλίας εκ των οποίων οι τέσσερις (4), με συνολική χωρητικότητα διακοσίων εξήντα (260) καθισμάτων, αφορούσαν αίθουσες «θεωρίας» και πέντε (5), με συνολική χωρητικότητα εκατόν εξήντα (160) καθισμάτων, αφορούσαν αίθουσες «εργαστηρίων». Από τις προαναφερθείσες αίθουσες έγινε χρήση μόνο των τεσσάρων αιθουσών «θεωρίας» καθώς το πρόγραμμα εφαρμόστηκε για την παραγωγή προγραμμάτων εξετάσεων θεωρητικών μαθημάτων. Επίσης στη Β.Δ είχαν τοποθετηθεί σαράντα (40) χρονοθυρίδες, ίσης χρονικής διάρκειας (2h:30'), που

απλώνονταν ισομερώς σε δέκα (10) ημέρες εξεταστικής περιόδου. Δηλαδή είχαμε τέσσερις χρονοθυρίδες για κάθε μέρα της εξεταστικής περιόδου. Τέλος στη Β.Δ είχαν δημιουργηθεί οχτακόσιες πενήντα (850) δηλώσεις μαθημάτων, των φοιτητών.

Πραγματοποιήθηκαν συνολικά έξι (6) πειράματα, κάθε ένα εκ των οποίων εκτελέστηκε πέντε φορές, έτσι ώστε να παρουσιαστούν οι μέσοι όροι τόσο των χρόνων εκτέλεσης, όσο και των τελικών τιμών στα αποτελέσματά τους. Τα έξι αυτά πειράματα μπορούμε να πούμε ότι χωρίζονται σε τρεις δυάδες πειραμάτων με παρόμοιες παραμετροποιήσεις, έτσι ώστε να δούμε κάποιες διαφορές και να συγκρίνουμε τα αποτελέσματα που επιφέρουν κάποιες δυνατότητες του ΕΑ που αναπτύχθηκε. Οι τρεις δυάδες των πειραμάτων παρουσιάζονται με λεπτομέρειες παρακάτω:

1. Η πρώτη δυάδα πειραμάτων ήταν σειριακές υλοποιήσεις (δηλαδή με έναν εξελισσόμενο πληθυσμό κάθε φορά) και περιελάμβαναν πληθυσμούς των πενήντα (50) ατόμων. Το πρώτο από αυτά (**Exp01**)<sup>5</sup>, που αποτελεί και την προεπιλεγμένη παραμετροποίηση του αλγορίθμου, εξέλιξε έναν πληθυσμό των πενήντα ατόμων για χίλιες (1000) γενιές. Ο τελεστής Ανασυνδυασμού που επιλέχτηκε ήταν ο “UniformCrossover” και εφαρμοζόταν με πιθανότητα  $P_{cross} = 0.7$  ή 70%. Ο τελεστής Μετάλλαξης ήταν ο “IntMutator” και εφαρμοζόταν με πιθανότητα  $P_{mut} = 0.3$  ή 30%. Ο τελεστής Επιλογής ήταν ο “TournamentSelector” που διεξήγαγε τουρνουά των δύο (2) ατόμων. Η παραμετροποίηση περιελάμβανε τη χρήση «ελιτισμού» ενός ατόμου, ενώ στην επόμενη γενιά επιβίωναν μόνο οι απόγονοι. Το δεύτερο πείραμα (**Exp02**) είχε την ίδια παραμετροποίηση με το πρώτο, με τη διαφορά ότι, σύν τοις άλλοις, εφάρμοζε και αναπροσαρμογή στις πιθανότητες εφαρμογής των τελεστών Ανασυνδυασμού και Μετάλλαξης ( $P_{cross}$  και  $P_{mut}$ ) κάθε πενήντα (50) γενιές.
2. Η δεύτερη δυάδα πειραμάτων ήταν επίσης σειριακές υλοποιήσεις και περιελάμβαναν πληθυσμούς των εκατό (100) ατόμων το κάθε ένα. Το τρίτο πείραμα (**Exp03**) εξέλιξε τα άτομα για τρεις χιλιάδες (3000) γενιές και για τελεστή Ανασυνδυασμού εφάρμοζε τον “UniformCrossover” με πιθανότητα εφαρμογής  $P_{cross} = 0.7$  ή 70%, ενώ για τελεστή Μετάλλαξης τον “IntMutator” με πιθανότητα εφαρμογής  $P_{mut} = 0.3$  ή 30%. Σαν τελεστής Επιλογής, είχε επιλεγεί ο “RouletteWheelSelector”, ενώ ο αλγόριθμος εφάρμοζε και «ελιτισμό» ενός ατόμου. Τέλος, την επόμενη γενιά την αποτελούσαν τα άτομα με την καλύτερη ποιότητα. Το τέταρτο πείραμα (**Exp04**) ήταν όμοιο με το τρίτο, με την προσθήκη της αναπροσαρμογής των πιθανοτήτων ( $P_{cross}$  και  $P_{mut}$ ) κάθε εκατό (100) γενιές.
3. Η τρίτη δυάδα πειραμάτων παρουσίαζε και το μεγαλύτερο ενδιαφέρον καθώς γινόταν σύγκριση ενός σειριακού ΕΑ με έναν παράλληλο ΕΑ. Ο σειριακός ΕΑ (**Exp05**) περιελάμβανε έναν πληθυσμό των διακοσίων (200) ατόμων που τα εξέλιξε για τέσσερις χιλιάδες (4000) γενιές. Οι τελεστές Ανασυνδυασμού και Μετάλλαξης, καθώς και οι πιθανότητες εφαρμογής τους ήταν όμοιοι με αυτούς των προηγούμενων πειραμάτων, δηλαδή

---

<sup>5</sup> “Exp01” είναι η κωδική ονομασία του πρώτου πειράματος. Ανάλογα ονοματίζονται και τα υπόλοιπα πειράματα, (πχ: Exp02, Exp03, ..., κτλ).



“UniformCrossover” με  $P_{\text{cross}} = 0.7$  και “IntMutator” με  $P_{\text{mut}} = 0.3$ . Για τον τελεστή Επιλογής εφαρμόστηκαν τρεις διαφορετικοί τελεστές, στις συνολικά πέντε εκτελέσεις του πειράματος. Αυτοί ήταν **α)** “Ranking-Selector” στις δύο πρώτες εκτελέσεις, **β)** “BestNSelector” στις επόμενες δύο και **γ)** “GreedyOverSelector” στην τελευταία εκτέλεση. Επίσης εφαρμόζε «ελιτισμό» ενός ατόμου, ενώ την επόμενη γενιά την αποτελούσαν πάντα τα καλύτερα άτομα του συνολικού πληθυσμού (γονείς και απόγονοι). Σε όλες τις εκτελέσεις είχαμε την εφαρμογή της αναπροσαρμογής των πιθανοτήτων ( $P_{\text{cross}}$  και  $P_{\text{mut}}$ ) κάθε εκατό (100) γενιές. Το έκτο και τελευταίο πείραμα (**Exp06**) περιελάμβανε τέσσερις πληθυσμούς των πενήντα (50) ατόμων που συνεξελίσσονταν παράλληλα για τέσσερις χιλιάδες (4000) γενιές. Οι τέσσερις αυτοί πληθυσμοί είχαν την ίδια παραμετροποίηση με τη διαφορά ότι εφαρμόζανε διαφορετικούς τελεστές Επιλογής, για να διερευνά ο κάθε ένας τον χώρο αναζήτησης με διαφορετικό τρόπο. Οι δύο πρώτοι (υπο)πληθυσμοί εφαρμόζαν τον τελεστή Επιλογής “TournamentSelector”, ο τρίτος τον “RouletteWheelSelector” και ο τέταρτος τον “RankingSelector”. Οι τελεστές Ανασυνδυασμού και Μετάλλαξης καθώς και οι πιθανότητες εφαρμογής αυτών, ήταν όμοιες με τα προηγούμενα πειράματα. Επίσης η εφαρμογή «ελιτισμού» ενός ατόμου, καθώς και η αναπροσαρμογή των πιθανοτήτων ( $P_{\text{cross}}$  και  $P_{\text{mut}}$ ) κάθε εκατό (100) γενιές, ήταν σταθερή παράμετρος σε όλους τους πληθυσμούς. Τέλος οι (υπο)πληθυσμοί επικοινωνούσαν μεταξύ τους, ανταλλάσσοντας γενετικό υλικό (χρωμοσώματα) κάθε εκατό (100) γενιές. Η μέθοδος ανταλλαγής ήταν η προεπιλεγμένη, δηλαδή «Όλοι σε όλους» και κάθε πληθυσμός έστελνε προς όλους τους υπόλοιπους αντίγραφα των δύο καλύτερων ατόμων του. Έπειτα κάθε πληθυσμός αντικαθιστούσε τα δύο χειρότερα άτομα του δικού του πληθυσμού με τα δύο καλύτερα από αυτά που είχε συλλέξει στη φάση της ανταλλαγής των ατόμων. Την επόμενη γενιά, σε κάθε (υπο)πληθυσμό, την αποτελούσαν τα άτομα με την καλύτερη ποιότητα.

Θα πρέπει σε αυτό το σημείο να αναφέρουμε ότι όλα τα παραπάνω πειράματα ικανοποιούσαν το ίδιο σύνολο περιορισμών με τη διαφορά ότι από δυάδα σε δυάδα αύξανε ο βαθμός δυσκολίας για μερικούς περιορισμούς. Δηλαδή σε όλα τα πειράματα οι περιορισμοί που συμμετείχαν στη διαμόρφωση των προγραμμάτων εξετάσεων ήταν οι «1», «4», «5», «6», «7» (βλέπε ενότητα 5.2.1 – «Περιορισμοί Ικανοποίησης»). Οι περιορισμοί «2», «3» δε συμμετείχαν καθόλου στη διαμόρφωση των προγραμμάτων καθώς αποτελούν πίο εξειδικευμένες περιπτώσεις. Οι ποινή που αποδίδονταν στους περιορισμούς «1» και «7» ήταν η μονάδα (1.0), δηλαδή η πίο μεγάλη, και ήταν κοινή σε όλα τα πειράματα. Από την άλλη μεριά οι ποινές που αποδίδονταν στους περιορισμούς «4», «5» και «6» είχαν τιμές (0.8) στην πρώτη δυάδα πειραμάτων (**Exp01** και **Exp02**), (0.9) στη δεύτερη δυάδα των πειραμάτων (**Exp03** και **Exp04**) και (1.0) στην τρίτη δυάδα των πειραμάτων (**Exp05** και **Exp06**). Βλέπουμε λοιπόν ότι υπήρχε μία κλιμάκωση της δυσκολίας όσο άλλαζαν και οι παραμετροποιήσεις του EA με

κορύφωση τα τελευταία δύο πειράματα όπου ο βαθμός δυσκολίας ήταν ο μέγιστος.

## 6.2 Παρουσίαση αποτελεσμάτων

Όλα τα πειράματα που παρουσιάστηκαν στην προηγούμενη ενότητα πραγματοποιήθηκαν στον ίδιο ηλεκτρονικό υπολογιστή (Pentium3 450Mhz, 192MB R.A.M), κάνοντας χρήση των ίδιων δεδομένων, όπως προαναφέρθηκε. Τα αποτελέσματα που παρουσιάζονται εδώ είναι ενδεικτικά και αναφέρονται κυρίως στους χρόνους εκτέλεσης<sup>6</sup> των πειραμάτων και στη μέγιστη τιμή της ποιότητας που επιτεύχθηκε σε κάθε πειραματική εκτέλεση. Το σύνολο των στατιστικών δεδομένων έχει αποθηκευτεί και θα διατεθεί σε συνοδευτικό υλικό μαζί με το κείμενο της πτυχιακής εργασίας.

Εκτέλεση	Χρόνος	Μέγιστη Ποιότητα	Μ.Ο Χρόνου	Μ.Ο Μεγ.Ποιότητας
1 <sup>η</sup>	00:02:00:594	0.539	<b>00:02:00:284</b>	<b>0.574</b>
2 <sup>η</sup>	00:01:59:761	0.585		
3 <sup>η</sup>	00:01:58:969	0.573		
4 <sup>η</sup>	00:02:02:305	0.594		
5 <sup>η</sup>	00:01:59:791	0.581		

Σχήμα 6.1. Συνοπτικά αποτελέσματα πειράματος "Exp01".

Εκτέλεση	Χρόνος	Μέγιστη Ποιότητα	Μ.Ο Χρόνου	Μ.Ο Μεγ.Ποιότητας
1 <sup>η</sup>	00:01:42:916	0.583	<b>00:01:40:764</b>	<b>0.585</b>
2 <sup>η</sup>	00:01:36:038	0.616		
3 <sup>η</sup>	00:01:42:097	0.584		
4 <sup>η</sup>	00:01:42:056	0.577		
5 <sup>η</sup>	00:01:40:714	0.564		

Σχήμα 6.2. Συνοπτικά αποτελέσματα πειράματος "Exp02".

Εκτέλεση	Χρόνος	Μέγιστη Ποιότητα	Μ.Ο Χρόνου	Μ.Ο Μεγ.Ποιότητας
1 <sup>η</sup>	00:12:25:602	0.583	<b>00:11:12:658</b>	<b>0.578</b>
2 <sup>η</sup>	00:10:44:756	0.560		
3 <sup>η</sup>	00:11:03:333	0.584		
4 <sup>η</sup>	00:10:54:791	0.589		
5 <sup>η</sup>	00:10:54:810	0.576		

Σχήμα 6.3. Συνοπτικά αποτελέσματα πειράματος "Exp03".

<sup>6</sup> Η κωδικοποίηση της χρονικής διάρκειας είναι η εξής: ΩΩ:ΛΛ:ΔΔ:ΧΔ

Εκτέλεση	Χρόνος	Μέγιστη Ποιότητα	Μ.Ο Χρόνου	Μ.Ο Μεγ.Ποιότητας
1 <sup>η</sup>	00:12:13:284	0.593	<b>00:12:06:548</b>	<b>0.576</b>
2 <sup>η</sup>	00:11:53:895	0.582		
3 <sup>η</sup>	00:11:40:676	0.608		
4 <sup>η</sup>	00:13:09:295	0.556		
5 <sup>η</sup>	00:11:35:589	0.539		

Σχήμα 6.4. Συνοπτικά αποτελέσματα πειράματος “Εχρ04”.

Εκτέλεση	Χρόνος	Μέγιστη Ποιότητα	Μ.Ο Χρόνου	Μ.Ο Μεγ.Ποιότητας
1 <sup>η</sup>	00:29:11:997	0.567	<b>00:28:35:137</b>	<b>0.570</b>
2 <sup>η</sup>	00:28:05:942	0.572		
3 <sup>η</sup>	00:27:59:523	0.570		
4 <sup>η</sup>	00:27:35:110	0.580		
5 <sup>η</sup>	00:30:03:112	0.561		

Σχήμα 6.5. Συνοπτικά αποτελέσματα πειράματος “Εχρ05”.

Εκτέλεση	Χρόνος	Μέγιστη Ποιότητα	Μ.Ο Χρόνου	Μ.Ο Μεγ.Ποιότητας
1 <sup>η</sup>	00:29:46:998	0.575	<b>00:23:15:784</b>	<b>0.577</b>
2 <sup>η</sup>	00:23:44:338	0.582		
3 <sup>η</sup>	00:19:56:810	0.587		
4 <sup>η</sup>	00:21:45:647	0.554		
5 <sup>η</sup>	00:21:05:127	0.585		

Σχήμα 6.6. Συνοπτικά αποτελέσματα πειράματος “Εχρ06”.

Οι παραπάνω πίνακες αποτελεσμάτων χρωματίζονται όμοια ανάλογα με τη δυάδα των πειραμάτων. Θα πρέπει να τονίσουμε ότι οι χρόνοι που παρουσιάζονται είναι πραγματικοί και έχουν ακριβεία χιλιοστού του δευτερολέπτου. Από την άλλη μεριά όπου παρατηρείται το φαινόμενο της μεγάλης απόκλισης από τη μία εκτέλεση στην άλλη δεν έχει να κάνει τόσο με τη συγκεκριμένη εφαρμογή, αλλά περισσότερο με το χρόνο εκτέλεσης που του δώθηκε από το Λειτουργικό Σύστημα (ΛΣ). Δηλαδή για να γίνει η διεξαγωγή των πειραμάτων όσο γίνεται πιό γρήγορα τερματίστηκαν όλες οι περιττές εφαρμογές που «έτρεχαν» στον Η/Υ και «έτρεχε» μόνο η διεξαγωγή των πειραμάτων. Παρ’όλα αυτά εμφανίστηκαν, σχετικά, μεγάλες χρονικές αποκλίσεις όχι μόνο ανάμεσα στις διαδοχικές εκτελέσεις του ίδιου πειράματος, αλλά ανάμεσα και σε διαφορετικές παραμετροποιήσεις που κανείς θα περίμενε να είναι πιό χρονοβόρες εξαιτίας της εφαρμογής περισσότερων λειτουργιών, όπως είναι αυτή της αναπροσαρμογής των πιθανοτήτων.

Συγκρίνοντας τα αποτελέσματα των δύο πρώτων πινάκων (σχήμα 6.1 και 6.2) που αφορούν την πρώτη δυάδα των πειραμάτων, βλέπουμε ότι η εφαρμογή της αναπροσαρμογής των πιθανοτήτων στο δεύτερο πείραμα (**Exp02**) έφερε καλύτερα ποιοτικά αποτελέσματα και μάλιστα σε μικρότερο χρονικό διάστημα, κάτι που δικαιολογήθηκε από τα ανωτέρω. Αν δούμε τώρα τα αποτελέσματα της δεύτερης δυάδας, που εμφανίζονται στα σχήματα 6.3 και 6.4 θα παρατηρήσουμε ότι ενώ το τέταρτο πείραμα (**Exp04**) είναι σαφώς πιο αργό, κάτι που δικαιολογείται από την εφαρμογή της αναπροσαρμογής των πιθανοτήτων, δεν εμφανίζει καλύτερα αποτελέσματα. Αυτό ίσως να οφείλεται στο γεγονός ότι το σύνολο των γενεών εξέλιξης δεν ήταν πολύ μεγάλο (μόλις 3000) για τον πληθυσμό των εκατό (100) ατόμων που εξελίσσονταν. Τέλος αν δούμε και την τρίτη δυάδα πειραμάτων (σχήματα 6.5 και 6.6) θα παρατηρήσουμε ότι η εφαρμογή του παράλληλου ΕΑ (**Exp06**) όχι μόνο είχε καλύτερα αποτελέσματα από αυτά του **Exp05**, αλλά επιτεύχθηκαν και σε μικρότερο χρονικό διάστημα.

Συνολικά, αν κρίνουμε τις δυάδες μεταξύ τους θα δούμε ότι παρόλο που ο Μ.Ο της ποιότητας των πειραμάτων είναι περίπου ο ίδιος (περίπου 0.575) δε θα πρέπει να ξεχνάμε ότι από δυάδα σε δυάδα ο βαθμός δυσκολίας αύξανε, καθώς αύξανε και η ποινή που αποδίδεται στους περιορισμούς. Αυτό σημαίνει ότι το 0.575 του 6<sup>ου</sup> πειράματος είναι σαφώς καλύτερο από το αντίστοιχο του 1<sup>ου</sup>.

Θα πρέπει επίσης να επισημάνουμε ότι κατά τη διάρκεια της εκτέλεσης του ΕΑ ένα αρκετά μεγάλο σύνολο στατιστικών δεδομένων συλλέγεται για κάθε πληθυσμό, για κάθε γενιά. Το σύνολο αυτό περιλαμβάνει δεκαεφτά (17) κατηγορίες δεδομένων. Αυτές είναι:

1. Μέγιστη Ποιότητα πληθυσμού
2. Ελάχιστη Ποιότητα πληθυσμού
3. Μέσος Όρος ποιότητας (Μ.Ο)
4. Μέση Τιμή ποιότητας (Μ.Τ)
5. Ποσοστό πληθυσμού πάνω από το Μ.Ο
6. Ποσοστό πληθυσμού πάνω από τη Μ.Τ
7. Ποσοστό (επί του πληθυσμού) των Μεταλλαγμένων ατόμων
8. Ποσοστό (επί του πληθυσμού) των Διασταυρωμένων ατόμων
9. Πιθανότητα Μετάλλαξης ( $P_{mut}$ )
10. Πιθανότητα Διασταύρωσης ( $P_{cross}$ )
11. Ποσοστό (επί του πληθυσμού) σύγκλισης κοντά στο 1% του καλύτερου ατόμου του πληθυσμού
12. Ποσοστό (επί του πληθυσμού) σύγκλισης κοντά στο 10% του καλύτερου ατόμου του πληθυσμού
13. Ποσοστό (επί του πληθυσμού) σύγκλισης κοντά στο 20% του καλύτερου ατόμου του πληθυσμού
14. Ποσοστό (επί του πληθυσμού) σύγκλισης κοντά στο 30% του καλύτερου ατόμου του πληθυσμού
15. Ποσοστό (επί του πληθυσμού) σύγκλισης κοντά στο 40% του καλύτερου ατόμου του πληθυσμού
16. Ποσοστό (επί του πληθυσμού) σύγκλισης κοντά στο 50% του καλύτερου ατόμου του πληθυσμού
17. Υπόλοιπο ποσοστό πληθυσμού

Όλα αυτά τα στοιχεία είναι διαθέσιμα και για τα παραπάνω πειράματα, αλλά για λόγους απλότητας δεν παρουσιάζονται.

Τέλος, όσον αφορά τα αποτελέσματα από τα παραπάνω πειράματα και βασιζόμενοι σε περαιτέρω ανάλυση που έγινε κατά τη διάρκεια των πειραμάτων, σε επίπεδο χρωμοσώματος, μπορούμε να πούμε ότι στο 85% των εκτελέσεων οι περιορισμοί «1» και «7» που αφορούν ο πρώτος τη χωρητικότητα των εξεταζόμενων θέσεων και ο δεύτερος τις συγκρούσεις πρώτου βαθμού, μεταξύ των μαθημάτων κάθε φοιτητή, ικανοποιήθηκαν πλήρως στο μέγιστο βαθμό. Αυτό οφείλεται κυρίως στο γεγονός ότι το δείγμα των εκατόν πενήντα τεσσάρων (154) φοιτητών πάνω στο οποίο δοκιμάστηκε ο αλγόριθμος είναι αρκετά μικρό σε σχέση με τα πραγματικά δεδομένα του τμήματος Πληροφορικής, όπου συνήθως συναντάμε πλήθος άνω των οκτακοσίων (800) φοιτητών, σε κάθε εξεταστική περίοδο. Από την άλλη μεριά ο περιορισμός «4», που αφορά την ικανοποίηση των απαιτητικών μαθημάτων ικανοποιήθηκε σε ποσοστό 90% στις περισσότερες εκτελέσεις, ενώ οι περιορισμοί «5» και «6» ήταν αυτοί στους οποίους οφείλεται η μέτρια τιμή της ποιότητας των αποτελεσμάτων. Ο περιορισμός «5» αφορά τις συγκρούσεις δεύτερου, τρίτου, κτλ βαθμού. Αυτός ο περιορισμός είναι εκ των πραγμάτων πολύ δύσκολο να ικανοποιηθεί στο μέγιστο βαθμό καθώς η ικανοποίηση αυτού του περιορισμού για έναν φοιτητή ή ένα υποσύνολο των φοιτητών, μπορεί να προκαλεί την μή-ικανοποίησή του σε ένα άλλο υποσύνολο των φοιτητών. Σύν τοις άλλοις οι δηλώσεις που υπήρχαν στη δοκιμαστική ΒΔ είχαν δημιουργηθεί με τέτοιο τρόπο ώστε το μεγαλύτερο τμήμα των φοιτητών να έχει μαθήματα από όλα τα εξάμηνα σπουδών. Κάτι τέτοιο, φυσικά, στην πράξη δεν ισχύει καθώς οι φοιτητές συνήθως επιλέγουν το μεγαλύτερο μέρος των μαθημάτων τους από το τρέχον εξάμηνο στο οποίο βρίσκονται. Τέλος ο περιορισμός «6» που προβλέπει να μην εξετάζονται μαθήματα του ίδιου εξαμήνου την ίδια μέρα, ήταν εξ αρχής αδύνατο να ικανοποιηθεί πλήρως καθώς από τις δηλώσεις προέκυπτε ότι από το έβδομο (7<sup>ο</sup>) εξάμηνο σπουδών είχαν δηλωθεί δεκαπέντε (15) μαθήματα. Οπότε μέσα στη χρονική διάρκεια των δέκα ημερών που διαρκούσε συνολικά η εξεταστική περίοδος θα είχαμε πέντε μαθήματα του 7<sup>ου</sup> εξαμήνου να συνυπάρχουν με άλλα του ίδιου εξαμήνου. Η καλύτερη περίπτωση βρέθηκε όταν αυτά τοποθετήθηκαν ανά δυάδες και η χειρότερη όταν σε μία μέρα εξετάσεων τοποθετήθηκαν έξι (6) μαθήματα του ίδιου εξαμήνου.

Εν κατακλείδι, τα αποτελέσματα που προέκυψαν από τα παραπάνω πειράματα ήταν κάτι παραπάνω από ενθαρρυντικά καθώς ο σκοπός, όπως αναφέρθηκε και στην αρχή του κεφαλαίου, δεν ήταν η εύρεση ενός βέλτιστου προγράμματος εξετάσεων, αλλά η επιβεβαίωση ότι ο εξελικτικός αλγόριθμος που αναπτύχθηκε δουλεύει και μάλιστα πολύ ικανοποιητικά.

# ΚΕΦΑΛΑΙΟ 7

## Συμπεράσματα - Περαιτέρω Έρευνα

Η παρούσα εργασία απέδειξε, ότι με τον αλγόριθμο που ανέπτυξε έφερε θετικά αποτελέσματα στη δημιουργία προγραμμάτων εξετάσεων, στη δοκιμαστική βάση δεδομένων στην οποία εφαρμόστηκε. Σε όλα τα πειράματα που περιγράφηκαν στο προηγούμενο κεφάλαιο δεν εφαρμόστηκε η βελτιστοποίηση του πίνακα των δηλώσεων, καθώς το δείγμα ήταν σχετικά μικρό και τα αποτελέσματα αυτής της διαδικασίας δε θα ήταν αρκετά εμφανή. Άλλωστε η συγκεκριμένη ενέργεια αφορά την ταχύτερη επεξεργασία των δεδομένων και όχι την καλύτερη ποιότητα των λύσεων. Επίσης σε όλα τα προηγούμενα παραδείγματα η εφαρμογή της διαδικασίας αναπροσαρμογής των πιθανοτήτων έφερε καλύτερα, ποιοτικά, αποτελέσματα απ'ότι όταν δεν εφαρμόστηκε, ενώ θετική κρίνεται και η συνεισφορά της εφαρμογής του «ελιτισμού».

Ακόμα, η χρήση Παράλληλων Εξελικτικών Αλγορίθμων σε σύγκριση με αυτή των Σειριακών ΕΑ (βλέπε πειράματα **Exp05** και **Exp06**) έφερε καλύτερα, ποιοτικά, αποτελέσματα χωρίς όμως να επιτυγχάνει και ουσιαστική βελτίωση στη χρονική διάρκεια της εξέλιξης, για προβλήματα ίδιου μεγέθους, μίας και η εφαρμογή των ΕΑ έγινε σε σειριακό Η/Υ.

Παρ'όλ'αυτά η πραγματική δύναμη και αξία του ΕΑ που αναπτύχθηκε θα φανεί στην πράξη μόνο εάν εφαρμοστεί σε πραγματικά δεδομένα.

Μερικές κατευθύνσεις που πιθανών να χρίζουν περαιτέρω διερεύνησης είναι οι εξής:

- **Υλοποίηση των ΕΑ σε παράλληλο Η/Υ.** Ο αλγόριθμος που αναπτύχθηκε στην παρούσα εργασία δίνει τη δυνατότητα στο χρήστη της εφαρμογής να υλοποιήσει και να πειραματιστεί με μια πληθώρα διαφορετικών ΕΑ. Οι διαφορές μεταξύ των ΕΑ αγγίζουν σχεδόν όλα τα βασικά στοιχεία τους. Έτσι μπορούμε να υλοποιήσουμε από έναν απλό (σειριακό) ΕΑ με σταθερές πιθανότητες ανασυνδυασμού και μεταλλάξης, χωρίς την εφαρμογή ελιτισμού στον οποίο θα επιβιώνουν μόνο οι απόγονοι κάθε γενιάς (canonical), μέχρι έναν παράλληλο ΕΑ ο οποίος θα χρησιμοποιεί μία

πληθώρα γενετικών τελεστών ανασυνδυασμού και μετάλλαξης, θα εφαρμόζει αναπροσαρμογή στις πιθανότητες εφαρμογής των γενετικών τελεστών και θα προσομοιώνει πολύπλοκες διαδικασίες μετανάστευσης, πολέμου, κ.α. Φυσικά όλα τα παραπάνω πραγματοποιούνται σε σειριακό Η/Υ, πράγμα που περιορίζει τη μέγιστη δυνατή επιτάχυνση σε μία άκρως παράλληλη διαδικασία, όπως είναι αυτή της εξέλιξης. Για την επίτευξη λοιπόν καλύτερων αποτελεσμάτων καλό θα ήταν να γίνει μία προσπάθεια υλοποίησης του ΕΑ σε παράλληλο Η/Υ. Με αυτόν τον τρόπο θα μπορούσαμε όχι μόνο να διερευνήσουμε μεγαλύτερους χώρους αναζήτησης του προβλήματός, αλλά να πάρουμε τα αποτελέσματα σε μικρότερο χρονικό διάστημα.

- **Έμμεση Αναπαράσταση.** Όπως αναφέραμε σε προηγούμενη ενότητα (κεφάλαιο 3<sup>ο</sup>), η αναπαράσταση μίας πιθανής λύσης του προβλήματος σε μορφή χρωμοσώματος είναι ουσιαστικής σημασίας, γιατί η αναπαράσταση παίζει πολύ σημαντικό ρόλο σε ολόκληρη τη λειτουργία του ΕΑ. Στην παρούσα εργασία, όπως αναφέραμε και στο τέταρτο κεφάλαιο, εξετάστηκε μόνο η περίπτωση της άμεσης αναπαράστασης (direct representation). Παρόλο που σε παρόμοια προβλήματα χρονοπρογραμματισμού αυτή είναι η πιο συνηθισμένη επιλογή, καλό θα ήταν να γίνει μία προσπάθεια υλοποίησης έμμεσης αναπαράστασης του χρωμοσώματος (indirect representation). Με αυτόν τον τρόπο θα είχαμε ένα μέτρο σύγκρισης μεταξύ δύο διαφορετικών αναπαραστάσεων όχι της ίδιας, αλλά διαφορετικής κατηγορίας. Ακόμα θα μπορούσε να δοθεί έμφαση στη δημιουργία μίας αναπαράστασης η οποία να εμφανίζει σε μικρότερο βαθμό το πρόβλημα της επίστασης (epistasis – gene interaction).
- **Εφαρμογή Ευριστικών Αλγορίθμων Αναζήτησης.** Σε όλες τις πιθανές υλοποιήσεις των ΕΑ που μπορούν να εφαρμοστούν, με τον αλγόριθμο που αναπτύχθηκε στην παρούσα εργασία, η μοναδική αξιοποίηση παλαιότερης γνώσης πάνω στο πρόβλημα που επιλύει είναι αυτή της αρχικοποίησης του αρχικού πληθυσμού με παλαιότερα χρωμοσώματα από προγενεστερες εξελικτικές διαδικασίες. Καμία άλλη χρήση βοηθητικής γνώσης δεν μπορεί να εφαρμοστεί για τη δημιουργία των προγραμμάτων εξετάσεων. Αυτό, φυσικά, έγινε σκόπιμα γιατί θέλαμε ο αλγόριθμος να είναι όσο το δυνατόν πιο ανεξάρτητος. Έχει όμως αποδειχτεί στην πράξη ότι η χρήση των ΕΑ μαζί με άλλες μεθόδους τοπικής αναζήτησης μπορεί να αποφέρει καλύτερα αποτελέσματα. Αυτές οι μέθοδοι τοπικής αναζήτησης κάνουν χρήση κάποιας ιδιαίτερης γνώσης που υπάρχει για το πρόβλημα και συνήθως σχετίζονται στενά με αυτό. Αυτό μπορεί να αφαιρεί, σε μικρό ή μεγάλο βαθμό, τη γενικότητα και την ανεξαρτησία του αλγορίθμου αλλά σε περιπτώσεις που κάτι τέτοιο δεν μας ενδιαφέρει είναι αποδεκτή η χρήση τέτοιων ευριστικών μεθόδων.
- **Κατευθυνόμενη Μετάλλαξη.** Η χρήση κατευθυνόμενης μετάλλαξης (directed mutation), έχει δείξει θετικά αποτελέσματα σε πολλές εφαρμογές

της στους ΕΑ. Παρ'όλ'αυτά η ένδειξη αυτή έχει προκύψει από εμπειρικές παρά από θεωρητικές μελέτες. Στον αλγόριθμο που αναπτύχθηκε, στην παρούσα εργασία, γίνεται χρήση μόνο μεταβλητής μετάλλαξης (*variable mutation*), κατά την οποία η πιθανότητα με την οποία εφαρμόζεται η μετάλλαξη αυξομειώνεται ανάλογα με κάποια στατιστικά στοιχεία, που αφορούν τη σύγκλιση του πληθυσμού που εξελίσσεται, έτσι ώστε να διατηρείται η ανομοιότητα (*diversity*) ανάμεσα στα άτομα του πληθυσμού. Περαιτέρω μελέτη, λοιπόν, θα μπορούσε να γίνει τόσο σε θεωρητικό, όσο και σε πειραματικό επίπεδο για τα ωφέλη που θα μπορούσαμε να έχουμε από την εφαρμογή ενός είδους κατευθυνόμενης μετάλλαξης στη λειτουργία του Εξελικτικού Αλγορίθμου (ΕΑ).



## Βιβλιογραφία

- [1] Αδαμίδης Παναγιώτης, «Συνεργαζόμενοι Πληθυσμοί – Βελτίωση της απόδοσης Παράλληλων Γενετικών Αλγορίθμων, χρησιμοποιώντας Συνεργαζόμενους Πληθυσμούς με Διαφορετικές Συμπεριφορές Εξέλιξης.», Διδακτορική Διατριβή, Αριστοτέλιο Πανεπιστήμιο Θεσσαλονίκης -Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Ηλεκτρονικών Υπολογιστών, Θεσσαλονίκη 1996.
- [2] Βλαχάβας Ιωάννης και Κεφαλάς Πέτρος και Βασιλειάδης Νικόλαος και Ρεφανίδης Ιωάννης και Κόκκορας Φώτιος και Σακελλαρίου Ηλίας, «Τεχνητή Νοημοσύνη» (κεφ: 2 – [σελ:13 .. 72], 5 – [σελ: 182 .. 193]), εκδόσεις Γαρταγάνη, Θεσσαλονίκη 2002.
- [3] Adamidis Panagiotis, “Parallel Evolutionary Algorithms: A Review”, Dept. of Applied Informatics, University of Macedonia 1998.
- [4] Martin Pelikan and David E.Goldberg, “Genetic Algorithms, Clustering, and the Breaking of Symmetry”, Illinois Genetic Algorithms Laboratory, IlliGAL Report No.2000013 March 2000.
- [5] Andrew Laurence Tuson, “Adapting Operator Probabilities In Genetic Algorithms”, Ph.D, University Of Edinburgh – Dept. of Artificial Intelligence, 1995.
- [6] Helio J.C. Barbosa and Asla Medeiros e Sa, “On Adaptive Operator Probabilities in Real Coded Genetic Algorithms”, Universidade Federal de Rio de Janeiro – Instituto de Matematica, Laboratorio Nacional de Computacao Cientifica.
- [7] Hugo Terashima Marin and Manuel Valenzuela-Rendon and Peter Ross, “Evolution of Constraint Satisfaction Strategies in Examination Timetabling”, ITESM-Centro de Inteligencia Artificial Mexico, University of Edinburgh-Division of Informatics, February 2000.
- [8] Darrell Whitley, “A Genetic Algorithm Tutorial”, Colorado State University – Dept. of Computer Science.
- [9] EvoStim: The EVONET Working Group on Evolutionary Scheduling and Timetabling, “The State of The Art in Evolutionary Approaches to Timetabling and Scheduling”.
- [10] E.K. Burke and S.Petrovic and R.Qu, “Case based Heuristic Selection for Examination Timetabling”, School of Computer Science and Information Technology, Jubilee Campus, University of Nottingham.
- [11] Hugo Terashima Marin, “A Comparison of GA-based Methods Graph Coloring Methods for solving the Timetabling Problem”, September 1994.
- [12] Peter Ross and Dave Corne and Hsiao –Lan Fang, “Improving Evolutionary Timetabling with Delta Evaluation and Directed Mutation”, University of Edinburgh – Dept. of Artificial Intelligence, 1994.
- [13] Sushil J.Louis, “Predicting time to convergence for GAs”, 1997.
- [14] Edmund Kieran Burke and Sanja Petrovic, “Recent Research Directions in Automated Timetabling”, School of Computer Science and Information

- Technology, University of Nottingham, Accepted for publication in European Journal of Operational Research – EJOR, 2002.
- [15] Edmund Burke and David Elliman and Rupert Weare, “Specialised Recombinative Operators for Timetabling Problems”, Department of Computer Science, University of Nottingham, UK.
  - [16] E.K.Burke and J.P.Newall and R.F.Weare, “A Memetic Algorithm for University Exam Timetabling”, Department of Computer Science – University of Nottingham, UK.
  - [17] E.K.Burke and D.G.Elliman and R.Weare, “Extensions To A University Exam Timetabling System”, University of Nottingham, UK.
  - [18] E.K.Burke and D.G.Elliman and R.Weare, “Automated Scheduling Of University Exams”, University of Nottingham, UK.
  - [19] Tony Wong and Pascal Côté and Paul Gely, “Final Exam Timetabling: A Practical Approach”, Department Of Automated Production Engineering, Dean of resource management.
  - [20] Edmund Burke and David Elliman and Rupert Weare, “A Genetic Algorithm Based University Timetabling System”, University of Nottingham, UK, 1994.
  - [21] M.Dimopoulou and P.Miliotis, “Implementation Of A University Course And Examination Timetabling System”, Athens University of Economics and Business, January 2000.
  - [22] Edmund Burke and Dave Elliman and Peter Ford and Rupert Weare, “Examination Timetabling in British Universities – A Survey”, Department of Computer Science – University of Nottingham, UK.
  - [23] Bryant A.Julstrom, “Adaptive Operator Probabilities in a Genetic Algorithm that Applies Three Operators”, St.Cloud State University – Dept. of Computer Science.
  - [24] Peter Ross and Emma Hart and Dave Corne, “Some Observation about GA-based Exam Timetabling”, University of Edinburgh – Dept.of AI, University of Reading – Dept.of Computer Science.
  - [25] Bernd Bullheimer, “An Examination Scheduling Model to Maximize Students' Study Time”, Department of Management Science – University of Vienna, Austria.
  - [26] Liam T.G Merlot and Natashia Boland and Barry D. Hughes and Peter J.Stuckey, “A Hybrid Algorithm for the Examination Timetabling Problem”, Department of Mathematics and Statistics - University of Melbourne Australia, Department of Computer Science and Software Engineering - University of Melbourne Australia.
  - [27] Roman V. Belavkin, “Genetic Algorithms – Lecture 8”, Middlesex University.
  - [28] E.K. Burke and J.P. Newall and R.F. Weare, “A Simple Heuristically Guided Search for the Timetable Problem”, Department of Computer Science – University of Nottingham.
  - [29] Mauro Annunziato and Stefano Pizzuti, “Adaptive Parameterization of Evolutionary Algorithms Driven by Reproduction and Competition”, ENEA – Agency for New technologies, Energy and Environment – Rome Italy.

- [30] Wilhelm Erben, "A Grouping Genetic Algorithm for Graph Colouring and Exam Timetabling", Department of Computer Science, FH Konstanz – University of Applied Sciences – Germany.
- [31] Philippe David, "A Constraint – Based Approach for Examination Timetabling Using Local Repair Techniques", Ecole des Mines de Nantes – France.
- [32] Sachi Nandan Mukherjee, "Timetabling using Cellular Genetic Algorithms with Adaptive Mutation Operators", Department of Computer Science, University of York, 2001.
- [33] Goldberg D. E. "Genetic Algorithms in Search, Optimization & Machine Learning". Reading, MA: Addison-Wesley, 1989.
- [34] Edmund Burke and David Elliman and Rupert Weare, "A Hybrid Genetic Algorithm for Highly Constraint Timetabling", University of Nottingham.
- [35] Dave Corve and Peter Ross and Hsiao-Lan Fung, "Fast Practical Evolutionary Timetabling", DAI Research Paper No (708), 1994.
- [36] Herbert Schild, "Οδηγός της Java2", απόδοση: Δημοσθένης Κ.Φιστουρής, εκδόσεις: Μ.Γκιούρδας.
- [37] Laura Lemay and Rogers Cadenhead, "Πλήρες Εγχειρίδιο της Java2 platform", απόδοση: Ιωάννης Β.Σαμαράς, εκδόσεις: Μ.Γκιούρδας.

# ΠΑΡΑΡΤΗΜΑ Α:

## ΘΕΩΡΗΜΑ ΣΧΗΜΑΤΩΝ (SCHEMA THEOREM)

Το **θεώρημα των προτύπων σχημάτων** (*schema theorem*) προτάθηκε αρχικά από τον Holland (1975) για να εξηγήσει πως λειτουργούν οι Γενετικοί Αλγόριθμοι (ΓΑ) μέσω της διάδοσης σε ολόκληρο τον πληθυσμό όμοιων τμημάτων γενετικού κώδικα. Ένα **πρότυπο σχήμα** (*scheme*) είναι ένα τμήμα γενετικού κώδικα ιδιαίτερης ομοιότητας το οποίο περιγράφει ένα υποσύνολο των συμβολοσειρών με ομοιότητες σε συγκεκριμένες θέσεις. Χωρίς να περιορίζουμε την γενίκευση, μπορούμε να υποθέσουμε ότι χρησιμοποιούμε αναπαράσταση με δυαδικές συμβολοσειρές δηλ. με την χρήση του αλφάβητου  $\{0, 1\}$ . Για τη δημιουργία σχημάτων εισάγεται στο αλφάβητο και το **αδιάφορο** (*don't care*) σύμβολο "\*". Έτσι έχουμε το αλφάβητο  $\{0, 1, *\}$ . Οποιαδήποτε συμβολοσειρά χρησιμοποιείται με την χρήση αυτού του αλφάβητου αποτελεί ένα πρότυπο σχήμα. Ένα πρότυπο σχήμα ταιριάζει σε κάποια συμβολοσειρά εάν κάθε 1 αντιστοιχεί σε 1, κάθε 0 αντιστοιχεί σε 0 και κάθε \* αντιστοιχεί είτε σε 1, είτε σε 0 σε όλες τις αντίστοιχες θέσεις. Τα πρότυπα σχήματα μας επιτρέπουν να περιγράψουμε με μεγαλύτερη ακρίβεια τις ομοιότητες μεταξύ διαφορετικών συμβολοσειρών.

Σαν παράδειγμα ας σκεφθούμε τα σχήματα και τις σειρές μήκους 5. Το σχήμα **\*0000** ταιριάζει σε δύο σειρές, στις  $\{10000, 00000\}$ . Ένα άλλο παράδειγμα είναι το σχήμα **\*111\*** που περιγράφει το τετραμελές υποσύνολο  $\{01110, 01111, 11110, 11111\}$ . Τελευταίο και πιο περιεκτικό παράδειγμα, το σχήμα **0\*1\*\*** που ταιριάζει σε οποιαδήποτε από τις οκτώ σειρές μήκους 5, οι οποίες ξεκινάνε με ένα 0 και έχουν έναν 1 στην τρίτη θέση. Όπως αντιλαμβανόμαστε η ιδέα ενός σχήματος μας δίνει ένα ισχυρό και στέρεο τρόπο για να μιλάμε για τις

ομοιότητες ανάμεσα σε σειρές πεπερασμένου μήκους, με βάση ένα αλφάβητο. Θα έπρεπε να δώσουμε έμφαση στο ότι το \* είναι μόνο ένα μετασύμβολο. Ο ΓΑ δεν το επεξεργάζεται ρητά. Απλώς, παίζει το ρόλο ενός σημειογραφικού μηχανισμού που επιτρέπει την περιγραφή όλων των πιθανών ομοιοτήτων μεταξύ σειρών ενός συγκεκριμένου μήκους και αλφαβήτου.

Το μέτρημα του συνολικού αριθμού των πιθανών σχημάτων, είναι μία διαφωτιστική άσκηση. Στο προηγούμενο παράδειγμα, για  $l=5$ , σημειώνουμε ότι υπάρχουν  $3 \times 3 \times 3 \times 3 \times 3 = 3^5 = 243$  διαφορετικά περιγράμματα ομοιομορφίας, επειδή κάθε μία από τις 5 θέσεις μπορεί να είναι 0,1 ή \*. Γενικά, για αλφάβητα βαθμού  $k$  (ο αριθμός των χαρακτήρων του αλφαβήτου), υπάρχουν  $(k + 1)^l$  σχήματα. Εκ πρώτης όψεως, φαίνεται ότι τα σχήματα κάνουν την αναζήτηση πιο δύσκολη. Για ένα αλφάβητο με  $k$  στοιχεία υπάρχουν μόνο  $k^l$  διαφορετικές σειρές μήκους  $l$ . Γιατί να ασχοληθούμε με τα  $(k + 1)^l$  σχήματα και να διευρύνουμε το χώρο ενδιαφέροντος; Θέτοντας το θέμα διαφορετικά, το παράδειγμα μήκους 5 έχει τώρα μόνο  $2^5 = 32$  διαφορετικές εναλλακτικές σειρές. Γιατί να κάνουμε τα πράγματα πιο δύσκολα λαμβάνοντας υπόψη  $3^5 = 243$  σχήματα. Πόσες πληροφορίες αποκτούμε μελετώντας τις ομοιότητες; Η απάντηση στο ερώτημα αυτό σχετίζεται με τον αριθμό των μοναδικών σχημάτων που περιλαμβάνονται στον πληθυσμό. Για να μετρηθεί αυτή η ποσότητα με ακρίβεια, απαιτούνται γνώσεις των σειρών μέσα σε κάποιο πληθυσμό. Μπορούμε να θέσουμε ένα όριο στον αριθμό των σχημάτων μέσα σε ένα συγκεκριμένο πληθυσμό, μετρώντας πρώτα τον αριθμό των σχημάτων που περιέχονται σε μία μοναδική σειρά, και μετά υπολογίζοντας ένα άνω όριο πάνω στο συνολικό αριθμό των σχημάτων μέσα στον πληθυσμό.

Για να γίνει αυτό πρακτικά αντιληπτό, ας εξετάσουμε, για παράδειγμα, τη σειρά 11111, η οποία είναι μέλος  $2^5$  σχημάτων μιας και κάθε θέση μπορεί να πάρει την πραγματική της τιμή, ή ένα αδιάφορο σύμβολο. Γενικότερα, κάθε σειρά περιέχει  $2^l$  σχήματα. Αποτέλεσμα αυτού είναι ένας πληθυσμός μεγέθους  $n$  να περιλαμβάνει μεταξύ  $2^l$  και  $n \cdot 2^l$  σχήματα, ανάλογα με την ποικιλία του. Αυτό το γεγονός πιστοποιεί την προηγούμενη διαίσθησή μας. Το αρχικό κίνητρο για να μελετήσουμε τις σημαντικές ομοιομορφίες, ήταν η λήψη περισσότερων πληροφοριών που θα μας βοηθούσαν να καθοδηγήσουμε την αναζήτησή μας. Το επίμαχο θέμα της απαρίθμησης αποδεικνύει ότι, ακόμα και σε μέτριου μεγέθους πληθυσμούς, περιέχεται ένας πλούτος πληροφοριών σχετικά με σημαντικές ομοιότητες. Θα εξετάσουμε τον τρόπο με τον οποίο οι ΓΑ εκμεταλλεύονται αποδοτικά αυτές τις πληροφορίες. Σε αυτό το κρίσιμο σημείο, φαίνεται να είναι αναγκαία κάποια παράλληλη επεξεργασία, αν πρόκειται να κάνουμε χρήση όλων αυτών των πληροφοριών με μία χρονική μέθοδο.

Από τα  $2^l$  έως  $n \cdot 2^l$  σχήματα, πόσα είναι πραγματικά επεξεργάσιμα από το ΓΑ, κατά ένα χρήσιμο τρόπο; Το ερώτημα απαντάται αν εξετάσουμε το αποτέλεσμα

της αναπαραγωγής, της διασταύρωσης και της μεταλλαγής στην ανάπτυξη ή στην παρακμή σημαντικών σχημάτων από γενεά σε γενεά. Το αποτέλεσμα της αναπαραγωγής σε κάποιο σχήμα είναι εύκολο να προσδιοριστεί. Εφόσον οι καλύτερα προσαρμοσμένες σειρές έχουν μεγαλύτερες πιθανότητες επιλογής, έχουμε έναν κατά μέσο όρο αυξημένο αριθμό δειγμάτων στα παρατηρημένα υποδείγματα της καλύτερης ομοιομορφίας. Πάντως, η αναπαραγωγή από μόνη της δεν είναι αρκετή για τη διερεύνηση νέων στοιχείων στο χώρο αναζήτησης. Τότε τί συμβαίνει σε κάποιο σχήμα όταν εισάγεται η έννοια της διασταύρωσης; Όταν η διασταύρωση δεν τέμνει ένα σχήμα, τότε δεν επιφέρει καμία αλλαγή. Αν όμως η τομή γίνει μέσα στο σχήμα τότε μπορεί να το αποδιοργανώσει. Σκεφθείτε για παράδειγμα, τα δύο σχήματα **1\*\*\*0** και **\*\*11\***. Το πρώτο είναι πιθανό να αποδιοργανωθεί από τη διασταύρωση, ενώ το δεύτερο είναι σχετικά απίθανο να καταστραφεί. Αποτέλεσμα αυτού είναι, τα σχήματα μικρού μήκους να μένουν ως έχουν μετά τη διασταύρωση, και να αναπαράγονται σε ένα καλό ρυθμό δειγμάτων από τον τελεστή της αναπαραγωγής. Η μεταλλαγή σε φυσιολογικούς, χαμηλούς ρυθμούς, δε διασπά ένα συγκεκριμένο σχήμα πολύ συχνά και εμείς μένουμε σε ένα εντυπωσιακό συμπέρασμα. Τα σχήματα υψηλής ποιότητας και μικρού μήκους (τα οποία ονομάζουμε *δομικά σύνολα* - building blocks) είναι πολλαπλασιασμένα γενεά προς γενεά δίνοντας εκθετικά αυξανόμενα δείγματα στα παρατηρημένα βέλτιστα. Όλο αυτό συμβαίνει παράλληλα, χωρίς κάποια ιδιαίτερη τήρηση βιβλίων ή κάποια ιδιαίτερη μνήμη ξέχωρα από εκείνη του πληθυσμού μας.

Για να εξετάσουμε τον τρόπο διάδοσης των προτύπων σχημάτων από γενιά σε γενιά, πρέπει να εισάγουμε δύο ακόμη έννοιες. Η **τάξη** (*order*) ενός πρότυπου σχήματος  $H$ , συμβολιζόμενη ως  $o(H)$ , είναι ο αριθμός των σταθερών θέσεων στο πρότυπο σχήμα, δηλ. ο αριθμός των συμβόλων 1 και 0, ή αλλιώς ο συνολικός αριθμός των συμβόλων μετά την αφαίρεση του αριθμού των αστερίσκων (\*). Το **μήκος ορισμού** (*defining length*) ενός πρότυπου σχήματος  $H$ , συμβολιζόμενο ως  $\delta(H)$ , είναι η απόσταση μεταξύ της πρώτης και της τελευταίας σταθερής θέσης στο πρότυπο σχήμα. Η ανάλυση του θέματος διευκολύνεται εάν θεωρήσουμε ότι μία γενιά ενός ΓΑ αποτελείται από δύο βήματα: το πρώτο είναι η επιλογή και το δεύτερο η **διασταύρωση** (*crossover*), η οποία είναι ο ανασυνδυασμός στους ΓΑ και η μετάλλαξη.

επιλογή

διασταύρωση, μετάλλαξη

Τρέχον πληθυσμός  $\Rightarrow$  Ενδιάμεσος πληθυσμός  $\Rightarrow$  Νέος πληθυσμός

Το θεώρημα των προτύπων σχημάτων για ένα απλό ΓΑ με διασταύρωση (ανασυνδυασμό ενός σημείου), μετάλλαξη αντιστροφής δυαδικού ψηφίου και επιλογή με την μέθοδο της ρουλέτας δίνεται από τον παρακάτω τύπο:

$$m(H, t+1) \geq m(H, t) \frac{f(H)}{\bar{f}} \left( 1 - p_c \frac{\delta(H)}{l-1} \left( 1 - \frac{1}{n} m(H, t) \frac{f(H)}{\bar{f}} \right) \right) (1 - p_m)^{o(H)}$$

όπου  $p_c$  είναι η πιθανότητα ανασυνδυασμού και  $p_m$  η πιθανότητα μετάλλαξης. Ο τύπος αυτός δείχνει ότι ευνοούνται πρότυπα σχήματα τα οποία έχουν μικρό μήκος ορισμού, χαμηλή τάξη και ποιότητα άνω της μέσης. Τα σχήματα αυτά αυξάνονται συνεχώς σε επόμενες γενιές. Ενώ αυτή είναι μία πολύ χρήσιμη ιδιότητα των ΓΑ, η πρακτική χρήση της είναι πολύ περιορισμένη.

# ΠΑΡΑΡΤΗΜΑ Β:

## Exam Timetabling Glossary

### [Contents]

- 1) Terms
- 2) Variable Definitions
- 3) Constraints

#### 1) Terms

**Clashes:** an occurrence of a student having 2 exams scheduled in the same session.

**Exam (examination):** a meeting of students for the purpose of completing a test.

**Room:** a place where students will sit their exam.

**Scheduled:** an exam that has been allocated to a session.

**Session:** a designated time and duration during which exams may be held (also called time or time slot).

**Student – exam:** a person sitting an exam. This is to differentiate from a student (who sits multiple exams). The total number of student-exams is much larger than the total number of students, as each student is counted for each exam they sit (which may be more than one).

**Timetable:** a feasible allocation of exams to sessions and rooms.

**Unscheduled:** an exam that has not been allocated to a session.

#### 2) Variable Definitions

$E = \{1...n\}$  denotes the given set of exams.

$T = \{1...v\}$  denotes the given set of exam sessions.

$R = \{1...r\}$  denotes the given set of exam rooms.

$s_i$  is the number of students sitting exam  $i$  for all  $i$  in  $E$ .



**A** is a set of exam restrictions. An element  $(i,m)$  of **A** represents the fact that exam  $i$  cannot be scheduled in session  $m$ .

**B** is a set of room restrictions. An element  $(k,m)$  of **A** represents the fact that room  $k$  is not available in session  $m$ .

**C<sub>k</sub>** represents the capacity of room  $k$ , for all  $k$  in **R**.

**H<sub>m</sub>** represents the total capacity of session  $m$  for all  $m$  in **T**. **H** will vary depending on **B**.

**D<sub>ij</sub>** is the number of students enrolled in both exams  $i$  and  $j$  for all  $i,j$  in **E**.

**x** is a timetable with all exams allocated to sessions.

**x<sub>i</sub>** is an integer variable indicating the session exam  $i$  is scheduled in.

**y<sub>i</sub>** is an integer variable indicating the room exam  $i$  is scheduled in.

**neq** = not equal to.

**SUM {i in E}** = sum over all  $i$  in **E**.

**%** = Remainder after integer division. For example,  $7\%3 = 1$ .

### 3) Constraints

**Capacity:** The total number of exams scheduled in a room,  $k$ , in a session,  $m$ , must be less than the capacity of the room.

**SUM {i in E} [(x<sub>i</sub> == m, y<sub>i</sub> == k) s<sub>i</sub>] <= C<sub>k</sub>**, for all  $k$  in **R**, for all  $m$  in **T**.

**Clashing:** A student cannot have two exams scheduled in the same session.

**x<sub>i</sub> neq x<sub>j</sub>**, for all exams  $i, j$  in **E**, **i neq j**, and **D<sub>ij</sub> > 0**.

**Exam Availability:** For all exam/session  $(i, m)$  elements of **A**, exam  $i$  cannot be scheduled in session  $m$ .

**x<sub>i</sub> neq m**, for all elements  $(i,m)$  of **A**.

**Room Availability:** No exam may be scheduled in a room when it is not available.

**For x<sub>i</sub> = m: y<sub>i</sub> neq k**, for all elements  $(k, m)$  of **B**.

If room allocation is not considered, this constraint manifests as **H<sub>m</sub>** not being constant for all  $m$ .

**Total Capacity:** The total number of exams scheduled in a session must be less than the total capacity of all rooms available in that session.

**SUM {i in E} [(x<sub>i</sub> == m) s<sub>i</sub>] <= H<sub>m</sub>**, for all  $m$  in **T**.