



ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ  
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ  
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

## ΕΥΦΥΗ ΣΥΣΤΗΜΑΤΑ ΡΟΜΠΟΤ

**Βιβλιογραφική Εργασία:**

**«Εξελικτική Υπολογιστική και Γενετικοί Αλγόριθμοι»**

Σφυρής Γεώργιος	5161
Τζεμανάκη Αντωνία	5167
Τζιώνας Δημήτριος	5168

ΘΕΣΣΑΛΟΝΙΚΗ 2008

## Εισαγωγή

Διαμάχες αιώνων ανάμεσα στην θρησκεία και την επιστήμη έχουν θέμα την εξέλιξη του ανθρώπου και κατ' επέκταση όλων των έμβιων οργανισμών. Οι οπαδοί της εκ του μηδενός δημιουργίας του κόσμου από τον Θεό υποστηρίζουν επιπλέον ότι η θεωρία της εξέλιξης ως επιστημονική θεωρία δεν έχει τίποτα το πρακτικό να προσφέρει, καθώς δεν έχει καμία σχέση με την καθημερινή ζωή. Ωστόσο, κάτι τέτοιο, όπως αποδεικνύει η επιστήμη της βιολογίας, δεν ισχύει.

Φαινόμενα όπως η ανθεκτικότητα που αναπτύσσουν τα βακτήρια στα αντιβιοτικά ή η ανθεκτικότητα των καρκινικών κυττάρων στην χημειοθεραπεία, εξηγούνται ξεκάθαρα από την θεωρία της εξέλιξης και είναι άμεσες συνέπειες της μετάλλαξης και της φυσικής επιλογής. Κατανοώντας αυτές τις αρχές, δίνεται η δυνατότητα να αντιμετωπιστούν οι παραπάνω επιβλαβείς οργανισμοί.

Στις τελευταίες δεκαετίες, η συνεχής ανάπτυξη της τεχνολογίας κατάφερε να εκμεταλλευτεί τη θεωρία της εξέλιξης σε ένα τελείως διαφορετικό πεδίο. Έτσι λοιπόν, στην επιστήμη των υπολογιστών η **εξελικτική υπολογιστική** (evolutionary computation) είναι τομέας της τεχνητής νοημοσύνης (artificial intelligence - ειδικότερα computational intelligence) που περιλαμβάνει συνδυαστικά προβλήματα βελτιστοποίησης.

Η εξελικτική υπολογιστική χρησιμοποιεί επαναληπτική πρόοδο, όπως η ανάπτυξη σε πληθυσμό. Αυτός ο πληθυσμός στη συνέχεια επιλέγεται με μία καθοδηγούμενη - τυχαία αναζήτηση χρησιμοποιώντας παράλληλη επεξεργασία έτσι ώστε να επιτευχθεί ο επιθυμητός στόχος.

Επεξεργασίες τέτοιου είδους πολύ συχνά προκύπτουν από τους βιολογικούς μηχανισμούς της εξέλιξης και τη μελέτη των έμβιων οργανισμών. Έννοιες όπως ο πληθυσμός, η μετάλλαξη και η φυσική επιλογή (επιβίωση του πιο προσαρμοσμένου) για την εύρεση της ακριβέστερης λύσης του προβλήματος, δανείζονται από την ίδια τη βιολογία.

Οι Εξελικτικές τεχνικές περιλαμβάνουν κυρίως αλγορίθμους αναζήτησης που προσπαθούν να βρουν την βέλτιστη λύση του προβλήματος, οι σημαντικότεροι από τους οποίους είναι:

- Εξελικτικοί αλγόριθμοι, που περιλαμβάνουν **γενετικούς αλγορίθμους**, εξελικτικό προγραμματισμό, εξελικτικές στρατηγικές, γενετικό προγραμματισμό και συστήματα ταξινόμησης με δυνατότητα μάθησης.
- Αλγόριθμοι νοημοσύνης σμήνους (*swarm intelligence*), όπως αυτοί που προσομοιάζουν τη συμπεριφορά μίας κοινωνίας μυρμηγκιών (*ant colony optimization*) ή σμήνους σωματιδίων (*particle swarm optimization*).
- Σε μικρότερο βαθμό περιλαμβάνουν επίσης αυτό-οργάνωση (*self-organization*) και διαφορική εξέλιξη (*differential evolution*).

Στο παρόν κείμενο θα ασχοληθούμε με τους εξελικτικούς αλγορίθμους και ειδικότερα με τους γενετικούς αλγορίθμους.

## Ιστορική αναδρομή

Στη δεκαετία του '50, πριν ακόμα την χρήση των υπολογιστών σε ευρεία κλίμακα, προέκυψε η ιδέα της χρήσης των αρχών του Δαρβίνου για αυτόματη επίλυση προβλημάτων.

Μέχρι τη δεκαετία του '60, τρεις διαφορετικές προσεγγίσεις αυτής της ιδέας έχουν αναπτυχθεί σε τρία διαφορετικά πεδία.

Ο εξελικτικός προγραμματισμός (evolutionary programming) προτάθηκε από τον *Lawrence J. Fogel* στις Η.Π.Α., ενώ ο *John Henry Holland* ονόμασε τη μέθοδό του γενετικό αλγόριθμο (genetic algorithm). Στη Γερμανία οι *Ingo Rechenberg* και *Hans-Paul Schwefel* εισήγαγαν τις εξελικτικές στρατηγικές (evolution strategies).

Αυτές οι περιοχές αναπτύχθηκαν ξεχωριστά για περίπου 15 χρόνια. Από τις αρχές της δεκαετίας του '90 και έπειτα, θεωρούνται ξεχωριστές προσεγγίσεις (διάλεκτοι) μιας ενιαίας τεχνολογίας, αυτή του εξελικτικού προγραμματισμού (evolutionary programming).

Επιπλέον, στις αρχές της δεκαετίας του '90, μία τέταρτη προσέγγιση έκανε την εμφάνισή της, ο γενετικός προγραμματισμός (genetic programming) από τον *J. Koza*.

Αυτές οι ορολογίες ορίζουν τον τομέα της εξελικτικής υπολογιστικής (evolutionary computing), ενώ ο εξελικτικός προγραμματισμός (evolutionary programming), οι εξελικτικές στρατηγικές (evolution strategies), οι γενετικοί αλγόριθμοι (genetic algorithms) και ο γενετικός προγραμματισμός (genetic programming) θεωρούνται υποτομείς της.

## Γενετικοί Αλγόριθμοι

Γενικά, ο γενετικός αλγόριθμος (ΓΑ σε συντομογραφία) είναι μία προγραμματιστική τεχνική, που μιμείται την ίδια τη βιολογική εξέλιξη και που λειτουργεί σαν στρατηγική επίλυσης προβλημάτων.

Για ένα συγκεκριμένο προς επίλυση πρόβλημα, η είσοδος του γενετικού αλγορίθμου είναι ένα σετ (πληθυσμός) κωδικοποιημένων πιθανών λύσεων του προβλήματος (συχνά καλούνται χρωμοσώματα ή γονιδιώματα). Όλες οι λύσεις, ανεξάρτητα από τη φύση τους, αξιολογούνται σύμφωνα με ένα μέτρο, την **συνάρτηση ποιότητας (fitness function)**. Αυτή η συνάρτηση κάθε φορά πρέπει να ορίζεται ώστε να ταιριάζει στην κάθε περίπτωση, αφού εξαρτάται από τη φύση του προβλήματος.

Ο αλγόριθμος εξελίσσει τον πληθυσμό εφαρμόζοντας σ' αυτόν διάφορες γενετικές διαδικασίες επιλογής, διαδικασίες αναπαραγωγής, διαδικασίες μετάλλαξης.

Έτσι, δημιουργείται ένας νέος πληθυσμός που αντικαθιστά τον προηγούμενο. Η διαδικασία επαναλαμβάνεται μέχρι να «βρεθεί λύση».

Η είσοδος του ΓΑ μπορεί να είναι είτε κάποιες λύσεις οι οποίες είναι ήδη γνωστό ότι λειτουργούν, οπότε και ο σκοπός του γενετικού αλγορίθμου είναι να τις βελτιώσει, είτε (όπως και γίνεται συνήθως) κάποιες υποψήφιες λύσεις που έχουν δημιουργηθεί με τυχαίο τρόπο.

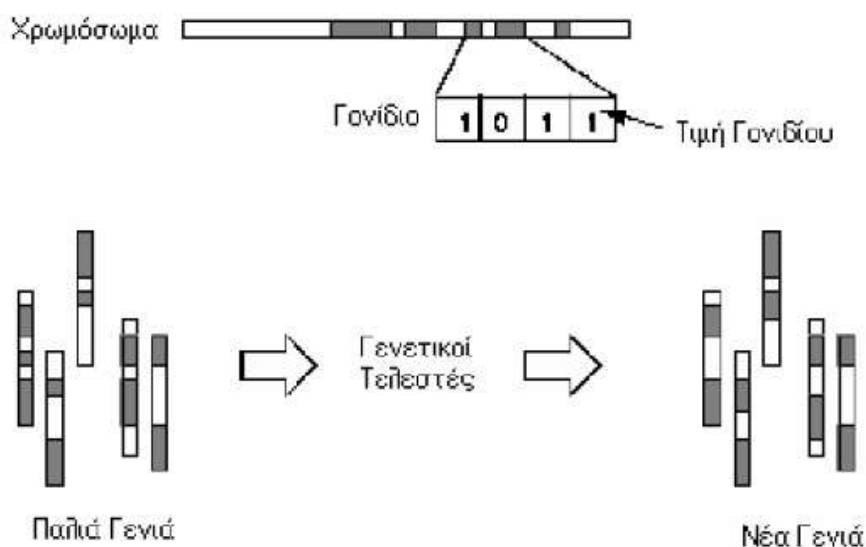
Σε ένα σετ λύσεων που δημιουργήθηκαν τυχαία, όπως είναι φυσικό οι περισσότερες δε θα δουλέψουν καθόλου, οπότε και θα διαγραφούν. Παρ' όλα αυτά, καθαρά για λόγους τύχης, μερικές είσοδοι μπορεί να δίνουν ένδειξη ότι ίσως δώσουν μία ικανοποιητική λύση στο πρόβλημα, έστω κι αν αυτή η ένδειξη είναι αδύναμη. Αυτές οι υποσχόμενες υποψήφιες λύσεις κρατούνται και αναπαράγονται.

Διάφορα αντίγραφα δημιουργούνται από αυτές, αλλά τα αντίγραφα αυτά δεν είναι ακριβή, αντίθετα εισάγονται σ' αυτά τυχαίες αλλαγές κατά τη διαδικασία της αντιγραφής, με τη βοήθεια των γενετικών τελεστών. Αυτοί οι ψηφιακοί απόγονοι μεταφέρονται στην επόμενη γενιά, οπότε και δημιουργείται εκ νέου ένα σετ υποψήφιων λύσεων, οι οποίες υποβάλλονται σε δεύτερο γύρο αξιολόγησης.

Οι υποψήφιες λύσεις που έγιναν χειρότερες ή που δε βελτιστοποιήθηκαν διαγράφονται. Αντίθετα, οι πιθανές λύσεις που βελτιώθηκαν χάρη στις τυχαίες διαφοροποιήσεις που εφαρμόστηκαν σ' αυτές και άρα είναι πιο αποτελεσματικές στην επίλυση του αρχικού προβλήματος, κρατούνται. Και πάλι, αυτές οι βελτιωμένες υποψήφιες λύσεις επιλέγονται και αντιγράφονται στην επόμενη γενιά αφού υποβληθούν για άλλη μία φορά σε τυχαίες αλλαγές. Η όλη διαδικασία συνεχίζεται όπως παραπάνω.

Το ζητούμενο είναι η μέση ικανότητα (fitness) του πληθυσμού να αυξάνεται με κάθε επανάληψη κι έτσι, μετά από εκατοντάδες ή χιλιάδες γενιές, να προκύψουν πολύ καλές λύσεις για το πρόβλημα.

Στο σχήμα δίνεται συνοπτικά η δομή του ΓΑ:



## Παραλληλισμός Ενοιών

Στην πραγματική Βιολογία, έχουμε Γενετική εφαρμοζόμενη σε πληθυσμούς χρωμοσωμάτων, που αντιπροσωπεύουν πληθυσμούς οργανισμών, οι οποίοι ανταγωνίζονται για να επιβιώσουν.

Στην τεχνητή Βιολογία, από την άλλη μεριά, έχουμε Γενετικούς τελεστές εφαρμοζόμενους σε πληθυσμούς συμβολοσειρών, που αντιπροσωπεύουν πληθυσμούς λύσεων, οι οποίες με τη σειρά τους επιλέγονται για να ικανοποιήσουν τους σκοπούς του χρήστη.

Πιο συγκεκριμένα, μία αντιστοίχιση της ορολογίας των γενετικών αλγορίθμων, φαίνεται παρακάτω:

Άτομο - Υποψήφια Λύση (χρωμόσωμα/ γονιδίωμα)

Επιδόσεις - Ποιότητα (fitness)

Περιβάλλον - Πρόβλημα

Φυσική επιλογή - Πιθανοτική επιλογή

Αναπαραγωγή - Ανασυνδυασμός

Γενιές ατόμων - Γενιές λύσεων

Βελτίωση είδους - Βελτιστοποίηση

## Μέθοδοι αναπαράστασης (κωδικοποίησης)

Πριν εφαρμοστεί ο γενετικός αλγόριθμος σε οποιοδήποτε πρόβλημα, είναι αναγκαία η εφαρμογή μίας μεθόδου που θα κωδικοποιήσει τις πιθανές λύσεις του προβλήματος σε μία μορφή ικανή να την αντιληφθεί ο υπολογιστής.

Η κωδικοποιημένη λύση ονομάζεται γενότυπος (genotype), ενώ η αποκωδικοποιημένη πληροφορία είναι ο φαινότυπος (phenotype).

➤ Μία κοινή προσέγγιση είναι να κωδικοποιηθούν οι λύσεις σαν **strings δυαδικών αριθμών**, δηλαδή ακολουθίες από '0' και '1', όπου η θέση του κάθε ψηφίου αντιπροσωπεύει την τιμή μίας συγκεκριμένης παραμέτρου της λύσης του προβλήματος.

➤ Μία άλλη παρόμοια προσέγγιση είναι η κωδικοποίηση των λύσεων σε **πίνακες ακεραίων ή δεκαδικών αριθμών**, όπου και πάλι η κάθε θέση του πίνακα αντιπροσωπεύει την τιμή μίας συγκεκριμένης παραμέτρου της λύσης του προβλήματος.

Αυτή η προσέγγιση δίνει μεγαλύτερη ακρίβεια και πολυπλοκότητα από αυτή της σχετικά περιορισμένης μεθόδου της δυαδικής κωδικοποίησης. Συχνά οι γενετικοί αλγόριθμοι που χρησιμοποιούνται για την εκπαίδευση νευρωνικών δικτύων χρησιμοποιούν αυτή τη μορφή κωδικοποίησης.

➤ Μία τρίτη προσέγγιση είναι η κωδικοποίηση των πιθανών λύσεων σε **strings χαρακτήρων**, όπου κάθε γράμμα αντιπροσωπεύει την τιμή μίας παραμέτρου της λύσης του προβλήματος.

Το προτέρημα των παραπάνω μεθόδων είναι ότι διευκολύνουν τον ορισμό τελεστών που προκαλούν τις τυχαίες αλλαγές στις υποψήφιας λύσεις του προβλήματος: μετατροπή του '0' σε '1' ή αντίστροφα, αφαίρεση/ πρόσθεση μίας τυχαίας ποσότητας από/ σε έναν αριθμού ή αλλαγή ενός γράμματος σε ένα άλλο.

➤ Ο *γενετικός προγραμματισμός*, διαθέτει μία ακόμα τεχνική αναπαράστασης, αυτή των **δένδρων**, όπου οι τυχαίες αλλαγές γίνονται με αλλαγή του τελεστή ή αλλοιώνοντας την τιμή ενός κόμβου του δέντρου ή αντικαθιστώντας ένα υποδέντρο με ένα άλλο.

Είναι σημαντικό να τονιστεί ότι οι εξελικτικοί αλγόριθμοι δε χρειάζονται να κωδικοποιήσουν τις υποψήφιας λύσεις σαν strings δεδομένων σταθερού μήκους. Μερικές φορές όντως συμβολίζονται με αυτόν τον τρόπο, αλλά κάποιες άλλες όχι.



## Μέθοδοι επιλογής

Η επιλογή αποσκοπεί στη βελτίωση της ποιότητας του πληθυσμού δίνοντας στα πιο ικανά άτομα μεγαλύτερη πιθανότητα να παραμείνουν σε επόμενες γενιές και εστιάζει την αναζήτηση σε υποσχόμενες περιοχές του χώρου λύσεων. Η πίεση που ασκείται από τον τελεστή επιλογής για το συνδυασμό χαρακτηριστικών δύο γονέων καλής ποιότητας, βασίζεται στην υπόθεση ότι αν τμήματα της αριθμητικής - γενετικής πληροφορίας, υπεύθυνα για την καλή ποιότητα των λύσεων συνδυαστούν, τότε οι πιθανότητες να προκύψουν απόγονοι καλύτερης ποιότητας αυξάνονται.

Υπάρχουν πολλές διαφορετικές τεχνικές με τις οποίες ο γενετικός αλγόριθμος διαλέγει τις υποψήφιες λύσεις που θα μεταβούν στην επόμενη γενιά. Παρακάτω ακολουθούν οι πιο συνηθισμένες, μερικές αμοιβαίως αποκλειόμενες, ενώ άλλες χρησιμοποιούνται συνδυαστικά.

### **Ελιτισμός (Elitist selection):**

Επιλέγονται οι πιο κατάλληλες (fit) τιμές (αυτές δηλαδή που δίνουν τα καλύτερα αποτελέσματα) σε κάθε γενιά. Οι περισσότεροι γενετικοί αλγόριθμοι δε λειτουργούν καθαρά ελιτιστικά, αντίθετα χρησιμοποιούν μία τροποποιημένη μορφή κατά την οποία επιλέγεται είτε η καλύτερη είτε μερικές από τις καλύτερες υποψήφιες λύσεις από κάθε γενιά ώστε να αντιγραφεί στην επόμενη, σε περίπτωση που δεν προκύψει κάτι καλύτερο.

### **Αναλογική επιλογή (Fitness-proportionate selection):**

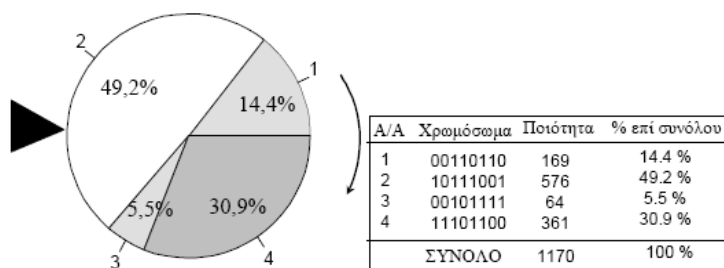
Περισσότερες προσαρμοσμένες (fit) τιμές είναι πιθανότερο να επιλεγούν, χωρίς όμως και να είναι σίγουρο. Απαιτεί δύο περάσματα ολόκληρου του πληθυσμού σε κάθε γενιά, κι έτσι μπορεί να γίνει χρονοβόρα. Ένα για να υπολογίσει τη μέση ικανότητα (fitness) του πληθυσμού και ένα για να υπολογίσει την αναμενόμενη ικανότητα του κάθε ατόμου.

### Ρουλέτα (Roulette-wheel selection):

Είναι μία μέθοδος ανάλογη με τη μέθοδο επιλογής που βασίζεται στην αναλογική. Εδώ, η πιθανότητα επιλογής ενός ατόμου (μίας υποψήφιας λύσης) είναι ανάλογη με το ποσό κατά το οποίο η τιμή της συνάρτησης αξιολόγησης για το συγκεκριμένο άτομο είναι μεγαλύτερη/ μικρότερη από την τιμή της για κάποιο ανταγωνιστικό.

Αυτό μπορεί να παρομοιαστεί με μία ρουλέτα, στην οποία κάθε άτομο παίρνει ένα κομμάτι του τροχού, αλλά οι καλύτερες λύσεις παίρνουν μεγαλύτερα κομμάτια από τις άλλες. Έτσι, η επιφάνεια της ρουλέτας διαιρείται σύμφωνα με τις τιμές της σχετικής ποιότητας των λύσεων, δηλαδή η γωνία κάθε τομέα της ρουλέτας γίνεται ανάλογα με τη σχετική ποιότητα. Αν οι σχετικές ποιότητες είναι ίσες τότε οι τομείς της ρουλέτας θα έχουν την ίδια επιφάνεια. Έπειτα, γυρίζει ο τροχός τόσες φορές όσους γονείς θέλουμε να επιλέξουμε, και κάθε φορά διαλέγεται αυτός που κατέχει το κομμάτι του τροχού που προκύπτει.

Ένα παράδειγμα της τεχνικής αυτής φαίνεται στο επόμενο σχήμα:



### SUS (stochastic universal sampling)

Με τη μέθοδο της ρουλέτας καταλήγουμε στον στατιστικά προσδοκώμενο αριθμό απογόνων για κάθε άτομο. Ωστόσο, καθώς οι πληθυσμοί που χρησιμοποιούνται στους ΓΑ είναι μικροί, στην πραγματικότητα ο αριθμός των απογόνων που αντιστοιχίζονται σε ένα άτομο απέχει συχνά από την αναμενόμενη τιμή του.

Αυτό συμβαίνει γιατί μία σχετικά απίθανη σειρά γύρων της ρουλέτας θα μπορούσε ίσως να αντιστοιχίσει όλους τους απογόνους στο χειρότερο (λιγότερο fit) άτομο του πληθυσμού (να τυχαίνει πάντα ο μικρότερος τομέας).

Μία τροποποιημένη μέθοδος που μειώνει αυτήν την διασπορά (τη διαφορά του πλήθους των πιθανών απογόνων από το αναμενόμενο) είναι η SUS, με την οποία αντί να γυρίζει η ρουλέτα  $N$  φορές για να πάρουμε ισάριθμους γονείς, γυρίζει μόνο μία φορά αλλά με  $N$  ξεχωριστούς δείκτες (βελόνες της ρουλέτας) με τους οποίους επιλέγονται οι γονείς.

### Παρατήρηση

Οι προαναφερθείσες τεχνικές στηρίζονται σε μία αυστηρή επιλογή σύμφωνα με την τιμή της συνάρτησης ποιότητας (κυρίως fitness-proportionate selection) του κάθε ατόμου. Έτσι, προκύπτει το ακόλουθο πρόβλημα, γνωστό σαν «**πρόωρη σύγκλιση**» (**premature convergence**) (το οποίο και δεν καταφέρνει να λύσει η SUS):

Όταν έχουμε από την αρχή της εκτέλεσης του ΓΑ μία «αυστηρή» επιλογή με βάση την ικανότητα (fitness) του κάθε ατόμου, έχουμε μία έμφαση στην εκμετάλλευση των πιο προσαρμοσμένων (fit) ατόμων, με αποτέλεσμα να μην «εξερευνούμε» άλλες περιοχές του πεδίου πιθανών λύσεων.

Στην αρχή οι συνέπειες δε είναι εμφανείς, καθώς εύκολα ξεχωρίζουν τα πιο ικανά άτομα (στην αρχή υπάρχει μεγάλη ποικιλία τιμών ποιότητας - fitness- για τα άτομα του πληθυσμού, και ξεχωρίζουν οι καλές από τις υπόλοιπες), στις επόμενες γενιές, όμως, τα άτομα που προκύπτουν έχουν παραπλήσιες τιμές ικανότητας με αποτέλεσμα να μην μπορεί να γίνει εύκολα επιλογή και να οδηγηθούμε σε τέλμα. Το πρόβλημα αυτό λύνεται προσπαθώντας να διατηρούμε μία ποικιλία κάθε φορά στις τιμές της συνάρτησης ποιότητας του πληθυσμού. Με λίγα λόγια, η εξέλιξη στηρίζεται στην ποικιλότητα των τιμών προσαρμογής των ατόμων του πληθυσμού.

### **Γενεαλογική αντικατάσταση (Generational replacement/selection):**

Οι απόγονοι των ατόμων που επιλέχθηκαν από κάθε γενιά αποτελούν εξ' ολοκλήρου την επόμενη γενιά, ενώ δεν κρατούνται άτομα μεταξύ των γενεών.

### **Κλιμάκωση (Scaling selection):**

Όσο η μέση ικανότητα (fitness) του πληθυσμού αυξάνεται, καθώς αυτό είναι σημάδι ότι η επιλογή γίνεται με ικανοποιητικό τρόπο, η συνάρτηση αξιολόγησης γίνεται όλο και πιο επιλεκτική - μεροληπτική. Αυτή η μέθοδος μπορεί να είναι ιδιαίτερα χρήσιμη για να γίνεται η καλύτερη επιλογή αργότερα, όταν όλα τα άτομα θα έχουν σχετικά υψηλή προσαρμογή (fitness) και μόνο μικρές διαφορές θα διακρίνουν το ένα ως πιο κατάλληλο από το άλλο.

### **Boltzman**

Συχνά είναι καλό κατά την επιλογή να μην υπάρχει σταθερή πίεση, αλλά μεταβλητή, με την έννοια, να επιτρέπεται μερικές φορές να περνούν στην επόμενη γενιά και λύσεις όχι τόσο ικανές (fit), για μεγαλύτερη ποικιλία στον πληθυσμό. Αργότερα, κατά την εκτέλεση του ΓΑ, μπορεί η πίεση να αυξηθεί, και τελικά να μεταφέρονται μόνο άτομα που δίνουν πολύ καλά αποτελέσματα.

Αυτό επιτυγχάνεται με τη μέθοδο επιλογής Boltzman, με τη βοήθεια ενός δείκτη "θερμοκρασίας". Μεγάλος δείκτης θερμοκρασίας σημαίνει πιο ελεύθερη επιλογή, ενώ μικρός δείκτης, πιο αυστηρή επιλογή. Ο δείκτης ξεκινάει με υψηλή τιμή (πιο ελεύθερη επιλογή) ώστε να έχουμε ποικιλία ατόμων στην αρχή, σε περίπτωση που μπορεί να προκύψει κάτι καλό από τα λιγότερο ικανά άτομα. Έπειτα, ο δείκτης χαμηλώνει (πιο μεροληπτική επιλογή), και με την αυστηρότερη επιλογή, τα άτομα που προέκυψαν με χαμηλή τιμή της συνάρτησης ποιότητας διαγράφονται.

### **Ιεραρχική επιλογή (Hierarchical selection):**

Σε αυτήν την τεχνική, τα άτομα περνούν από πολλαπλούς γύρους επιλογής σε κάθε γενιά. Οι αξιολογήσεις στα χαμηλά (αρχικά) επίπεδα είναι γρηγορότερες και λιγότερο διακριτικές, ενώ εκείνα τα άτομα που τα καταφέρνουν μέχρι και τα υψηλότερα επίπεδα, αξιολογούνται πιο αυστηρά.

Το κέρδος από αυτήν τη μέθοδο συνίσταται στο ότι μειώνει το συνολικό υπολογιστικό χρόνο, επειδή α) χρησιμοποιεί γρήγορη και λιγότερο επιλεκτική αξιολόγηση για να ξεχωρίσει την πλειοψηφία των υποσχόμενων υποψήφιων λύσεων έναντι των άχρηστων και β) υποβάλλει μόνο εκείνες που «επιβιώνουν» από τους αρχικούς γύρους σε πιο λεπτομερή και υπολογιστικά πιο χρονοβόρα αξιολόγηση.

### **Επιλογή μόνιμης κατάστασης (Steady-state selection):**

Οι απόγονοι των ατόμων που επιλέχθηκαν από κάθε γενιά επιστρέφουν στον προηγούμενο πληθυσμό και αντικαθιστούν κάποια από τα λιγότερο κατάλληλα άτομα της προηγούμενης γενιάς. Μόνο μερικά από τα άτομα κρατούνται μεταξύ των γενεών.

### **Κατάταξη (Rank selection):**

Σε κάθε άτομο στον πληθυσμό δίνεται ένας βαθμός (νούμερο κατάταξης) βασισμένο στην ικανότητα (fitness) του. Η επιλογή κάθε φορά βασίζεται σε αυτή την κατάταξη που προκύπτει, αντί στις απόλυτες διαφορές των τιμών της συνάρτησης αξιολόγησης.

Το πλεονέκτημα της μεθόδου είναι ότι μπορεί να αποτρέψει τα πολύ προσαρμοσμένα (very fit) άτομα απ' το να επικρατήσουν εις βάρος των λιγότερο προσαρμοσμένων σε πολύ αρχικό στάδιο. Σε περίπτωση που αυτό συνέβαινε, η γενετική ποικιλότητα του πληθυσμού θα μειωνόταν και προσπάθειες για εύρεση μίας ικανοποιητικής λύσης τελικά θα εμποδίζονταν.

Το μειονέκτημα της όμως, είναι ότι καθυστερώντας την πίεση επιλογής, ο ΓΑ είναι αρκετά αργός στην εύρεση ποιοτικών (fit) ατόμων, ενώ επιπλέον απαιτείται χρόνος για να ταξινομήσει ολόκληρο τον πληθυσμό σύμφωνα με το βαθμό του κάθε ατόμου. Είναι δηλαδή, χρονοβόρα διαδικασία.

### **Τουρνουά (Tournament selection):**

Η μέθοδος τουρνουά λειτουργεί πιο απλά από τη ρουλέτα. Μοιάζει με την μέθοδο κατάταξης όσον αφορά στην πίεση του τελεστή επιλογής, αλλά έχει το πλεονέκτημα να είναι υπολογιστικά πιο αποδοτική και λειτουργεί καλύτερα σε παράλληλη υλοποίηση.

Υποομάδες από άτομα επιλέγονται από τον ευρύτερο πληθυσμό, και τα μέλη της κάθε υποομάδας ανταγωνίζονται το ένα το άλλο. Μόνο ένα άτομο από κάθε υποομάδα επιλέγεται για αναπαραγωγή.

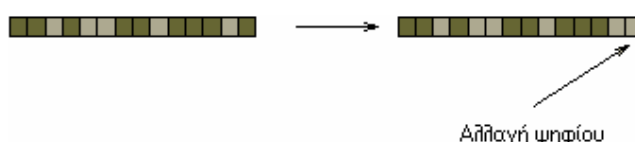
### Μέθοδοι αλλαγής

Όπως προαναφέρθηκε, αφού προκύψουν οι καλύτερες υποψήφιες λύσεις με την εφαρμογή μίας συνάρτησης επιλογής, θα πρέπει να μεταφερθούν στην επόμενη γενιά, αφού πρώτα αλλοιωθούν, ελπίζοντας ότι θα βελτιωθούν. Η βελτίωση επιτυγχάνεται κυρίως με τους εξής τρόπους:

➤ Ο πρώτος είναι η **μετάλλαξη (mutation)**, και είναι ίσως ο απλούστερος. Κατ' αντιστοιχία με τις αλλαγές που επιφέρει η μετάλλαξη στα γονίδια των οργανισμών στην πραγματική βιολογία, η μετάλλαξη στους γενετικούς αλγορίθμους, προκαλεί μικρές κάθε φορά αλλαγές στον κώδικα αναπαράστασης (γενότυπο) μιας πιθανής λύσης.

Για παράδειγμα, αν αρχικά έχει χρησιμοποιηθεί δυαδική κωδικοποίηση, κατά τη διαδικασία της μετάλλαξης, ένα ψηφίο (bit) της δυαδικής ακολουθίας που κωδικοποιεί μια πιθανή λύση επιλέγεται τυχαία σύμφωνα με την

πιθανότητα μετάλλαξης και υφίσταται διαταραχή δηλαδή αλλάζει η τιμή του. Αυτό έχει σαν αποτέλεσμα την αύξηση της ποικιλίας του πληθυσμού. Η πιθανότητα με την οποία εκτελείται ο τελεστής μετάλλαξης συνήθως επιλέγεται στο διάστημα [0.001, 0.01] (αναφέρεται σε όλα τα ψηφία του πληθυσμού). Στην επόμενη εικόνα φαίνεται η μετάλλαξη ενός ατόμου που έχει κωδικοποιηθεί δυαδικά:



Όταν το χρωμόσωμα κωδικοποιείται με πραγματικούς αριθμούς, τότε ο τελεστής της μετάλλαξης εφαρμόζεται διαφορετικά. Επιλέγεται το γονίδιο το οποίο θα υποστεί μετάλλαξη και μετατρέπεται σε πραγματική τιμή. Στην τιμή αυτή προστίθεται μια ποσότητα η οποία ακολουθεί συνήθως κανονική κατανομή.

➤ Η δεύτερη μέθοδος είναι η **διασταύρωση (crossover)**. Αποτελεί ουσιαστικά την καινοτομία και το κύριο μέσο για ποικιλομορφία στον γενετικό αλγόριθμο, ενώ η μετάλλαξη προστατεύει τον πληθυσμό από το να σταθεροποιηθεί και άρα, λειτουργεί περισσότερο παρασκηνιακά.

Όπως, λοιπόν, στον πραγματικό βιολογικό μηχανισμό, όπου κατά την αναπαραγωγή τα χρωμοσώματα των γονέων "ανασυνδυάζονται" προς δημιουργία του απογόνου, έτσι και στους ΓΑ παρατηρείται μία αναδιοργάνωση στους κώδικες αναπαραστάσης των γονέων, που καταλήγει στον απόγονο.

Ο τελεστής διασταύρωσης είναι υπεύθυνος για την εξερεύνηση του χώρου των πιθανών λύσεων. Μέσω αυτού, η αριθμητική πληροφορία (δηλαδή το «γενετικό υλικό») ανταλλάσσεται μεταξύ ενός ζεύγους τυχαία επιλεγμένων ατόμων. Με τη διασταύρωση επιλέγονται τμήματα των δύο γονέων και παράγεται ένας ή και δύο τεχνητοί απόγονοι.

Συνήθεις μορφές της διασταύρωσης είναι:

- Η διασταύρωση θέσης (Single-point crossover), όπου επιλέγεται τυχαία μία θέση στον κώδικα αναπαράστασης των γονέων, και ο ένας γονέας συνεισφέρει όλο το τμήμα του κώδικα του που βρίσκεται πριν από εκείνη τη θέση και ο άλλος γονέας όλο το τμήμα που βρίσκεται μετά. Ένα παράδειγμα φαίνεται στο παρακάτω σχήμα:

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

Η διασταύρωση αυτής της μορφής, όμως, υστερεί, διότι δε μπορεί να δώσει όλους τους πιθανούς συνδυασμούς. Για παράδειγμα, αν έχουμε τους γονείς  $11^{*****}1$  και  $****11^{**}$ , δεν είναι δυνατόν να προκύψει με αυτή την τεχνική ο απόγονος  $11^{**}11^{*}1$ .

- Η ομοιόμορφη διασταύρωση (uniform crossover), σύμφωνα με την οποία, στον κώδικα αναπαράστασης κάθε απογόνου, η τιμή σε κάθε θέση του κώδικα του απογόνου μπορεί να προέρχεται από το γενότυπο είτε του ενός είτε του άλλου γονέα, με ίση πιθανότητα.

Επιπλέον, μία ομοιόμορφη διασταύρωση μπορεί να είναι απλή, οπότε γίνεται σε ένα μόνο σημείο της ακολουθίας των δυαδικών ψηφίων ή πολλαπλή, οπότε γίνεται σε πολλά σημεία της ακολουθίας. Και στις δυο περιπτώσεις το σημείο ( ή τα σημεία) της διασταύρωσης επιλέγονται τυχαία. Μία απλή διασταύρωση φαίνεται στο επόμενο σχήμα:

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



Σαν αποτέλεσμα, ο τελεστής διασταύρωσης παράγει άτομα τα οποία ανήκουν σε μια κλειστή περιοχή του χώρου λύσεων με ακρότατα τις ελάχιστες και τις μέγιστες τιμές που εμφανίζονται στα διανύσματα των ατόμων για όλο τον πληθυσμό. Έτσι ο τελεστής αυτός συμβάλει στην εξερεύνηση μιας περιοχής του χώρου έρευνας και στην εύρεση καινούργιων ελαχίστων.

Η συχνότητα με την οποία εκτελείται ο τελεστής διασταύρωσης στον πληθυσμό των λύσεων εξαρτάται από μια προκαθορισμένη πιθανότητα διασταύρωσης. Οι τιμές της πιθανότητας διασταύρωσης στην πράξη επιλέγονται στο διάστημα  $[0.6, 0.95]$ , ενώ έχουν προταθεί και τεχνικές κατά τις οποίες η πιθανότητα διασταύρωσης προσαρμόζεται ανάλογα με την πορεία της αναζήτησης.

Με τη διασταύρωση έχουμε τη δυνατότητα να συνδυαστούν τα καλύτερα χαρακτηριστικά των δύο γονέων καταλήγοντας σε έναν απόγονο ο οποίος κληρονομεί τις ικανότητες τους και όσα έχουν ήδη «μάθει» και καμιά από τις «αδυναμίες» τους.

➤ Η τρίτη μέθοδος είναι η **μετανάστευση (migration)**. Η μέθοδος αυτή προσομοιάζει τη βιολογική μετανάστευση πληθυσμών από μία γεωγραφική περιοχή σε μία άλλη. Συγκεκριμένα, αυτό που γίνεται στο γενετικό αλγόριθμο είναι να χωρίζεται ο συνολικός πληθυσμός σε υπο-πληθυσμούς, να επιλέγονται τα καλύτερα άτομα από κάθε υπο-πληθυσμό και να αντικαθιστούν τα χειρότερα άτομα των γειτονικών υπο-πληθυσμών.

Η διαφορά είναι ότι τα άτομα αυτά αντιγράφονται στον επόμενο πληθυσμό και δεν αφαιρούνται από τον προηγούμενο. Η μετανάστευση συμβαίνει επαναληπτικά μία φορά μετά από ένα συγκεκριμένο αριθμό γενεών (π.χ. κάθε 20 γενεές).

➤ Η τέταρτη μέθοδος είναι η **αντιστροφή (inversion)**. Στην περίπτωση αυτή αντιστρέφεται η σειρά των στοιχείων σε ένα ολόκληρο τμήμα του κώδικα αναπαράστασης ενός ατόμου. Το τμήμα αυτό ορίζεται από δύο τυχαία σημεία πάνω στον κώδικα αναπαράστασης. Η μέθοδος αυτή, όμως, δεν έχει οδηγήσει σε καλά αποτελέσματα ως τώρα.

### Πλεονεκτήματα ΓΑ

Οι γενετικοί αλγόριθμοι, τελικά, αποδείχτηκαν να είναι εξαιρετικά ισχυροί και αποτελεσματικοί σαν μέθοδος επίλυσης προβλημάτων, επιδεικνύοντας την ισχύ των αρχών της εξέλιξης. Μερικά από τα πλεονεκτήματά τους είναι τα εξής:

- Δεν υπάρχουν περιορισμοί για το πρόβλημα και το είδος του χώρου λύσεων, είναι γενικά εφαρμόσιμοι.
- Μπορούν να επιλύουν δύσκολα προβλήματα γρήγορα και αξιόπιστα.
- Είναι μία μέθοδος που κάνει ταυτόχρονα εξερεύνηση του χώρου αναζήτησης και εκμετάλλευση της ήδη επεξεργασμένης πληροφορίας.
- Έχουν χαμηλό κόστος ανάπτυξης και εφαρμογής αλλά και είναι εύκολα επεκτάσιμοι και εξελίξιμοι.
- Επιδέχονται παράλληλη υλοποίηση, υπάρχει δηλαδή δυνατότητα παράλληλης εξερεύνησης πολλών υπο-περιοχών του χώρου λύσεων (Exploration) αλλά και διεξοδική εξερεύνηση περιοχών υψηλής ποιότητας (Exploitation).
- Εύκολα ενσωματώνουν και άλλες μεθόδους και μπορούν εύκολα να συνεργαστούν με τα υπάρχοντα μοντέλα και συστήματα.
- Μπορούν να δεχθούν λύσεις προτεινόμενες από τον χρήστη ενώ συγχρόνως παρέχουν πλήθος εναλλακτικών λύσεων.

- Δουλεύει με κωδικοποίηση του συνόλου των παραμέτρων και όχι με τις παραμέτρους.
- Ψάχνει μέσα από ένα σύνολο σημείων και όχι μόνο από ένα σημείο.
- Χρησιμοποιεί πιθανολογικούς κανόνες μετάβασης και όχι ντετερμινιστικούς.
- Κάνει χρήση της αντικειμενικής συνάρτησης και δεν έχει σχέση με παραγώγους.

### Εφαρμογές

Η εφαρμογή των γενετικών αλγορίθμων γίνεται συνήθως σε προβλήματα μη πραγματικού χρόνου (offline problems) καθώς ο περιορισμός χρόνου και η υπολογιστική πολυπλοκότητα δεν επιτρέπουν την εφαρμογή του σε πραγματικό χρόνο. Ωστόσο, έχουν βρεθεί όπως ήδη αναφέρθηκε και μέθοδοι παραλληλοποίησης της διαδικασίας για την εφαρμογή σε πραγματικό χρόνο (online).

Μερικά από τα πεδία στα οποία εφαρμόζονται συχνά οι γενετικοί αλγόριθμοι είναι τα εξής:

- Σχεδίαση μηχανών αεροπλάνων
- Αναγνώριση προσώπων
- Προγραμματισμός υπολογιστών
- Συστήματα αναγνώρισης εισβολών
- Σχεδίαση τηλεπικοινωνιακών δικτύων
- Βέλτιστη τοποθέτηση ενεργών στοιχείων για μείωση θορύβου
- Βελτιστοποίηση προβλημάτων με αριθμητικές παραμέτρους και συνδυαστικών προβλημάτων (Numerical and Combinatorial Optimization)
- Μοντελοποίηση και αναγνώριση συστημάτων (System modeling and identification)

- Χρονικός Προγραμματισμός και Έλεγχος Συστημάτων (Planning, Scheduling and Control)
- Σχεδίαση (Engineering Design)
- Εξόρυξη Δεδομένων (Data Mining)
- Machine Learning
- Artificial Life

### Επίλογος

Οι γενετικοί αλγόριθμοι έχουν χρησιμοποιηθεί σε μία μεγάλη ποικιλία επιστημονικών πεδίων επιλύοντας προβλήματα εξίσου ή πιο δύσκολα από αυτά που αντιμετωπίζονται από τον άνθρωπο. Πολλές φορές μάλιστα, οι λύσεις των γενετικών αλγορίθμων είναι πιο κομψές αλλά και περισσότερο αποτελεσματικές ή ακόμα και πιο περίπλοκες από αυτές που θα μπορούσε να δώσει ένας μηχανικός.

Είναι ένας κλάδος καινούριος αλλά πολλά υποσχόμενος λόγω των μεγάλων προτερημάτων του και της έντονης ποικιλομορφίας στον τρόπο εφαρμογής του.

### Βασική Βιβλιογραφία

1. Ηλεκτρονική εγκυκλοπαίδεια Wikipedia  
( [http://en.wikipedia.org/wiki/Genetic\\_algorithm](http://en.wikipedia.org/wiki/Genetic_algorithm)  
[http://en.wikipedia.org/wiki/Evolutionary\\_computation](http://en.wikipedia.org/wiki/Evolutionary_computation) )
2. Adam Marczyk - *Genetic Algorithm and Evolutionary Computation*
3. Ιάσων Βιττώριος - Διπλωματική εργασία (Α.Π.Θ., επιβλέπων Καθηγητής Βασίλειος Πετρίδης)
4. Hsiao-Lan Fang - *Genetic Algorithms in Timetabling and Scheduling*
5. Melanie Mitchell - *An introduction to Genetic Algorithms*

### Συμπληρωματική Βιβλιογραφία

1. Tetsuya Higuchi, Yong Liu, Xin Yao - *Evolvable Hardware*
2. Σωκράτης Κ. Κάτσικας - Σημειώσεις: Εξελικτικός προγραμματισμός: Νέες μέθοδοι βελτιστοποίησης
3. Σπυρίδων Α. Καζαρήλης - Σημειώσεις για Γενετικούς αλγόριθμους
4. Α. Γ. Σταφυλοπάτης - Σημειώσεις: Εξελικτικός υπολογισμός και Γενετικοί αλγόριθμοι