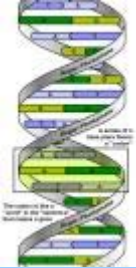


Γενετικοί Αλγόριθμοι

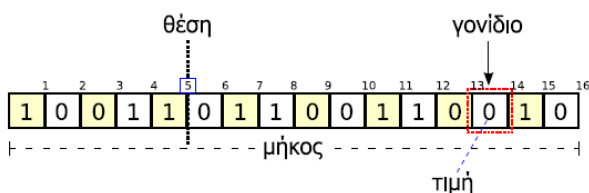


Φώτης Κόκκορας

ΑΤΕΙ Θεσσαλονίκης
Τμήμα Πληροφορικής

Αναπαράσταση Χρωσωμάτων

❖ Δυαδική Αναπαράσταση



❖ Αναπαράσταση Ακεραίων

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	4	5	3	3	7	7	9	16	12	10	14	10	12	9

❖ Αναπαράσταση Κινητής Υποδιαστολής

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
-1.2	5.4	8.1	3.3	1.2	2.1	1.4	-1.9	10	8.3	8.3	2.7	0.2	0.1	0.1	-0.2

❖ Αναπαράσταση Permutation (Μετάθεσης ή Συνδυασμού)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	2	4	5	3	8	6	7	9	16	12	15	14	10	11	13

- ❑ **κάθε γονίδιο εμφανίζεται μία φορά** (π.χ. πρόβλημα πλανόδιου πωλητή)

Αρχικοποίηση

- ❖ Εισάγονται αρχικές τιμές πληθυσμού, συμβατές με την αναπαράσταση που έχει επιλεγεί.
- ❖ Συνήθως είναι τυχαίες τιμές που παράγονται από μια γεννήτρια τυχαίων αριθμών.

Πληθυσμός

- ❖ Περιέχει υποψήφιες λύσεις.
- ❖ Συνήθως είναι μια δομή δεδομένων πίνακα (ή συνδεδεμένη λίστα).
 - ☐ Το μέγεθός της καθορίζεται από το μέγεθος του πληθυσμού με τον οποίο θέλουμε να δουλέψουμε.

Επιλογή Ατόμων για Αναπαραγωγή

- ❖ **Ζητούμενο:** Η επιλογή σχετικά καλών ατόμων με βάση την ποιότητα της λύσης που ορίζουν (δεν δημιουργούνται νέες λύσεις).
- ❖ **Ανάλογα με την ποιότητα των ατόμων.**

Αρχικά υπολογίζονται ποσοστά επιλογής για όλα τα άτομα και στη συνέχεια εφαρμόζεται κάποιος από τους ακόλουθους αλγόριθμους επιλογής:

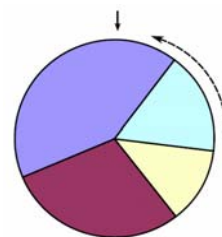
 - ☐ Cost Roulette Wheel
 - ☐ Stochastic Universal Sampling
- ❖ Rank Roulette Wheel
- ❖ Tournament Selection

Η Έννοια του Τροχού

- ❖ Δημιουργείται ένα τροχός (δίσκος) που χωρίζεται σε τομείς που αντιστοιχίζονται στα άτομα του πληθυσμού.
- ❖ Κάθε τομέας έχει μέγεθος ανάλογο με την ποιότητα του ατόμου.
- ❖ Έστω ότι θέλουμε να επιλέξουμε λ άτομα από πληθυσμό $P=(C_1, C_2, \dots, C_N)$ ατόμων, όπου τα άτομα είναι ταξινομημένα σε φθίνουσα σειρά ως προς την ποιότητα.
- ❖ Το μέγεθος των τομέων υπολογίζεται ως εξής:
 - ❑ Υπολογίζεται η συνολική τιμή ποιότητας του πληθυσμού: $F = \sum_{i=1}^{\lambda} f(C_i)$
 - όπου $i \in [1, \lambda]$ και f είναι η συνάρτηση ποιότητας
 - ❑ Υπολογίζεται η πιθανότητα επιλογής του ατόμου i : $P_i = \frac{f(C_i)}{F}$
 - ❑ Υπολογίζεται η αθροιστική πιθανότητα q_i για κάθε άτομο C_i : $q_i = \sum_{j=1}^{\lambda} P_j$
- ❖ **Παρατήρηση:** Η μέθοδος μεροληπτεί υπέρ των ατόμων με καλύτερη τιμή ποιότητας και μπορεί να οδηγήσει σε πρόωρη σύγκλιση του πληθυσμού σε τοπικό βέλτιστο του προβλήματος.

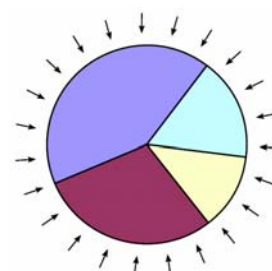
Cost Roulette Wheel

- ❖ Γυρίζουμε τον τροχό λ φορές (όσα και τα άτομα που θέλουμε να επιλέξουμε).
- ❖ Επιλέγουμε ένα άτομο τη φορά με την εξής διαδικασία:
 - ❑ Έστω r ένας τυχαίος αριθμός στο διάστημα $[0, 1]$.
 - αν $r \leq q_i$ τότε επιλέγεται το πρώτο άτομο
 - διαφορετικά επιλέγεται το i -στό άτομο για το οποίο ισχύει: $q_{i-1} < r \leq q_i$ με $i \in [2, \lambda]$



Stochastic Universal Sampling

- ❖ Αποτελεί προσομοίωση του Cost Roulette Wheel με τη διαφορά ότι υπάρχουν λ ομοιόμορφα κατανεμημένοι δείκτες γύρω από τον τροχό.
- ❖ Επιλέγουμε άτομα με την εξής διαδικασία:
 - ❑ Έστω r ένας τυχαίος αριθμός στο διάστημα $[0, 1/N]$.
 - Θυμίζουμε ότι N είναι το μέγεθος του αρχικού πληθυσμού από τον οποίο επιλέγουμε.
 - ❑ Για όσο ισχύει $r \leq q_i$ επιλέγεται το i -στό άτομο και ο αριθμός r αυξάνεται κατά $1/N$
 - ❑ Η διαδικασία επαναλαμβάνεται μέχρις ότου επιλεγούν τα λ ζητούμενα άτομα.



Αναπροσαρμογή Τιμών Ποιότητας

❖ Προβλήματα:

- ❑ Αν οι τιμές ποιότητας των ατόμων είναι παραπλήσιες, οι πιθανότητες που υπολογίζονται είναι σχεδόν ίδιες και η επιλογή γίνεται σχεδόν τυχαία!
- ❑ Οι μηχανισμοί με βάση την ποιότητα συμπεριφέρονται διαφορετικά στην ίδια συνάρτηση ποιότητας f , όταν αλλάζει η κλίμακα τιμών ποιότητας.

❖ Αντιμετώπιση: αναπροσαρμόζουμε με κάποιο τρόπο τις τιμές ποιότητας:

- ❑ **Windowing**: από κάθε τιμή ποιότητας αφαιρείται η τιμή ποιότητας του χειρότερου ατόμου
 - $f'(C_i) = f(C_i) - f_{min}$
- ❑ **Linear Scaling**: μεταβάλλουμε γραμμικά τις τιμές ποιότητας
 - $f'(C_i) = \alpha f(C_i) + \beta$ με τους συντελεστές α και β να ορίζονται από το χρήστη
- ❑ **Sigma Scaling**: Η τιμή ποιότητας μεταβάλλεται σύμφωνα με τη μέση τιμή \bar{f} και την τυπική απόκλιση σ του πληθυσμού:
 - $f'(C_i) = \max(f(C_i) + (\bar{f} - c \cdot \sigma), 0)$ c είναι μικρός ακέραιος (συνήθως στο διάστημα $[1, 5]$)
- ❑ **Power Law**: οι τιμές ποιότητας υψώνονται σε δύναμη k κοντά στο 1 (π.χ. 1.005)
 - $f'(C_i) = (f(C_i))^k$

Rank Roulette Wheel

❖ Παρόμοιος τελετής με τον Cost Roulette Wheel.

- ❑ Διαφέρει στον τρόπο υπολογισμού των πιθανοτήτων επιλογής (δηλαδή στο πρώτο στάδιο).

❖ Η πιθανότητα υπολογίζεται με βάση τη θέση (rank) που κατέχει το χρωμόσωμα στον ταξινομημένο ως προς την ποιότητα πληθυσμό.

- ❑ Ισχύει δηλαδή η σχέση δεξιά:

$$P_i = \frac{\lambda - i + 1}{\sum_{i=1}^{\lambda} i}$$

❖ Παράδειγμα:

- ❑ Έστω πληθυσμός 6 ατόμων ($\lambda=6$) σε ταξινομημένη σειρά ως προς την τιμή ποιότητας.
 - Η τιμή ποιότητας δεν μας ενδιαφέρει.
- ❑ Ο παρονομαστής της παραπάνω σχέσης είναι:
 - $\sum i = 1+2+3+4+5+6 = 21$
- ❑ Με εφαρμογή της παραπάνω σχέσης προκύπτουν οι πιθανότητες που φαίνονται στον πίνακα δεξιά.

i	P _i
1	28.6%
2	23.8%
3	19.0%
4	14.3%
5	9.5%
6	4.8%

Tournament Selection

❖ Πλεονεκτήματα:

- ☐ Δεν απαιτεί πληροφορίες από όλο τον πληθυσμό.
- ☐ Δεν απαιτεί ταξινόμηση, ως προς την ποιότητα, **όλου** του πληθυσμού.
 - Σε μεγάλους ή/και σε κατανεμημένους πληθυσμούς είναι υπολογιστικά ακριβή διαδικασία.

❖ Βήματα:

- ☐ Επιλέγονται τυχαία n άτομα ($n \geq 2$) από τον πληθυσμό.
- ☐ Από τα n άτομα, επιλέγεται αυτό με την καλύτερη τιμή ποιότητας.
- ☐ Η διαδικασία επαναλαμβάνεται μέχρι να συγκεντρωθούν τα λ άτομα που απαιτούνται.

❖ Έχοντας δημιουργήσει τον πληθυσμό που θα συμμετάσχει στην αναπαραγωγική διαδικασία, επιλέγουμε από αυτόν τυχαία ζευγάρια και εφαρμόζουμε τεχνικές ανασυνδυασμού (αναπαραγωγής).

- ☐περιγράφονται στη συνέχεια

Ανασυνδυασμός

❖ Δημιουργεί έναν ή περισσότερους απογόνους (συνήθως 2) συνδυάζοντας τα χαρακτηριστικά των γονέων.

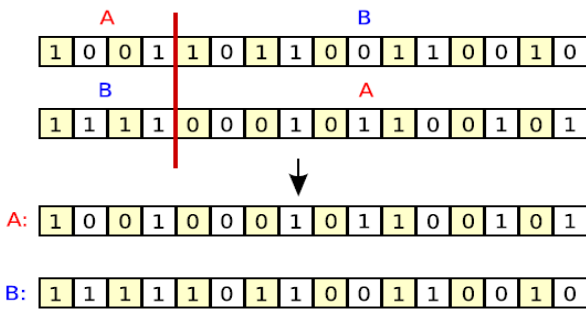
❖ Στα επόμενα, θεωρούμε ότι οι γονείς έχουν τη μορφή:

$$C_1 = (c_1^1, c_2^1, \dots, c_n^1) \text{ και } C_2 = (c_1^2, c_2^2, \dots, c_n^2)$$

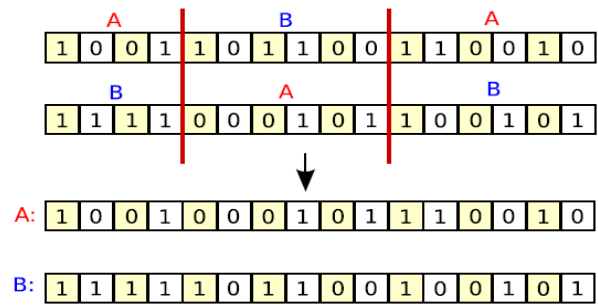
Τελεστές Ανασυνδυασμού (1/4)

Διαδική Αναπαράσταση

Single Point Crossover

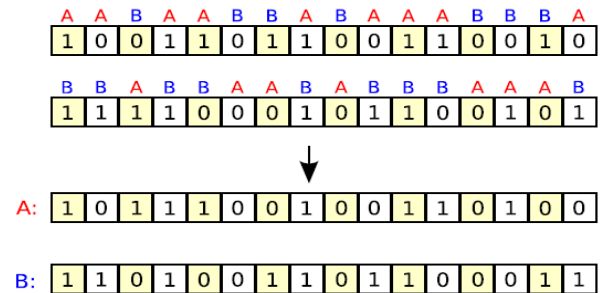


N Point Crossover



❖ Uniform

- ❑ Οι τιμές σε κάθε γονίδιο των απογόνων επιλέγονται τυχαία από τις τιμές των αντίστοιχων γονιδίων των γονέων.



Τελεστές Ανασυνδυασμού (2/4)

Αναπαράσταση Κινητής Υποδιαστολής

Flat

Δημιουργείται ένας απόγονος $H = (h_1, h_2, \dots, h_i, h_{i+1}, \dots, h_n)$, με h_i μία τυχαία τιμή από το διάστημα $[c_i^{min}, c_i^{max}]$, όπου $c_i^{min} = \min(c_i^1, c_i^2)$ και $c_i^{max} = \max(c_i^1, c_i^2)$.

Arithmetical

Δημιουργούνται δύο απόγονοι, $H_k = (h_1^k, h_2^k, \dots, h_i^k, h_{i+1}^k, \dots, h_n^k)$ με $h_i^1 = \lambda c_i^1 + (1 - \lambda)c_i^2$, όπου $k = 1, 2$, $h_i^2 = \lambda c_i^2 + (1 - \lambda)c_i^1$ και λ ένας τυχαίος αριθμός στο $[0, 1]$.

Extended Line

Δημιουργείται ένας μόνο απόγονος $H = (h_1, h_2, \dots, h_i, h_{i+1}, \dots, h_n)$, με $h_i = c_i^{min} + \alpha(c_i^{max} - c_i^{min})$ όπου α ένας τυχαίος αριθμός στο $[-0.25, 1.25]$, $c_i^{min} = \min(c_i^1, c_i^2)$ και $c_i^{max} = \max(c_i^1, c_i^2)$.

Average

Δημιουργείται ένας απόγονος $H = (h_1, h_2, \dots, h_i, h_{i+1}, \dots, h_n)$ με $h_i = \frac{c_i^1 + c_i^2}{2}$.

BLX- α

Δημιουργείται ένας απόγονος $H = (h_1, h_2, \dots, h_i, h_{i+1}, \dots, h_n)$ με h_i ένας τυχαίος αριθμός στο διάστημα $[c_i^{\min} - I\alpha, c_i^{\max} + I\alpha]$, όπου $c_i^{\min} = \min(c_i^1, c_i^2)$, $c_i^{\max} = \max(c_i^1, c_i^2)$, $I = c_i^{\max} - c_i^{\min}$ και α μία σταθερά. Παρατηρούμε ότι ο BLX- α με $\alpha = 0.0$ είναι ισοδύναμος με τον Flat.

Heuristic

Έστω ότι ο γονέας C_1 έχει καλύτερη τιμή ποιότητας από τον C_2 . Τότε δημιουργείται ένας απόγονος $H = (h_1, h_2, \dots, h_i, h_{i+1}, \dots, h_n)$ με $h_i = r(c_i^1 - c_i^2) + c_i^1$ όπου r ένας τυχαίος αριθμός στο διάστημα $[0, 1]$.

Τελεστές Ανασυνδυασμού (3/4)

Αναπαράσταση Ακεραίων

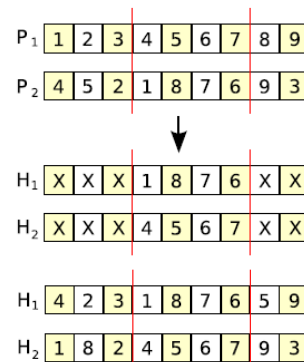
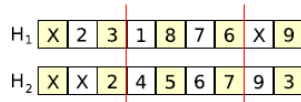
- ❖ Μπορούμε να χρησιμοποιήσουμε τελεστές παρόμοιους με αυτούς της δυαδικής αναπαράστασης.
- ❖ Οι τελεστές αναπαράστασης κινητής υποδιαστολής δεν εξυπηρετούν γιατί δεν παράγουν ακέραιες τιμές!

Τελεστές Ανασυνδυασμού (4/4)

Αναπαράσταση Permutation (Μετάθεσης)

❖ PMX (Partially Mapped Crossover)

- ❑ Επιλέγουμε ένα υποσύνολο των χρωμοσωμάτων διαλέγοντας τυχαία δύο σημεία και ανταλλάσσουμε τα εσωτερικά τμήματα. ➔
- ❑ Από τα υπόλοιπα γονίδια, κρατούμε αυτά που δεν εμφανίζονται ήδη στους απογόνους. ➡



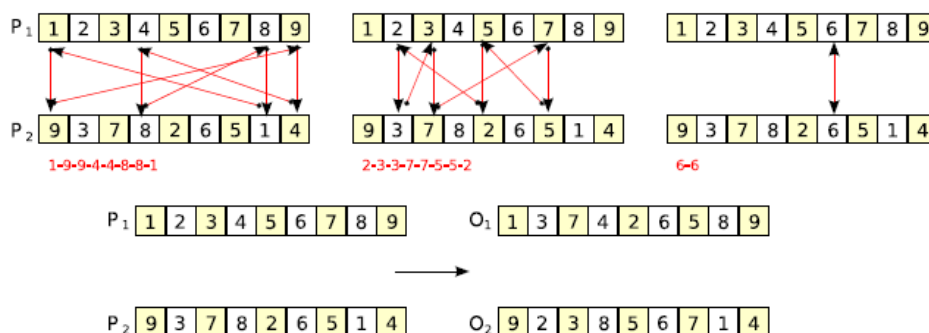
- ❑ Συμπληρώνουμε τα κενά με τα εναπομείναντα γονίδια των γονέων. ➔

❖ Παραλλαγή της PMX (εφαρμογή στους γονείς P1 και P2 προηγούμενου παραδείγματος)

- ❑ Επιλέγουμε ένα υποσύνολο των γονέων διαλέγοντας τυχαία δύο σημεία και κρατάμε τα εσωτερικά τμήματα. $H_1 = [X | X | X | 4 | 5 | 6 | 7 | X | X]$ και $H_2 = [X | X | X | 1 | 8 | 7 | 6 | X | X]$
- ❑ Για να συμπληρώσουμε το H_1
 - αναδιατάσσουμε τον γονέα P2 φέρνοντας το δεύτερο τμήμα εμπρός: **9 3** 4 5 2 1 8 7 6
 - αφαιρούμε τα γονίδια που υπάρχουν ήδη στον απόγονο H_1 : **9 3** 4 5 2 1 8 7 6 ➔ **9 3 2 1 8**
 - με τα γονίδια που έμειναν συμπληρώνουμε τον H_1 ως εξής: **2 1 8** | 4 5 6 7 | **9 3**
 - αντίστοιχα προκύπτει ο απόγονος H_2 : **3 4 5** | 1 8 7 6 | **9 2**

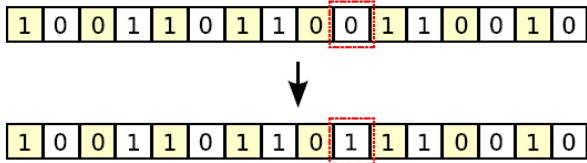
❖ Κυκλικός Ανασυνδυασμός (Cycle Recombination)

- ❑ Κανόνες "κύκλου":
 - 1) Εκκίνηση από πρώτο διαθέσιμο γονίδιο του P1.
 - 2) Κατακόρυφη μετάβαση από P1 σε P2.
 - 3) Μετάβαση από P2 στο ίδιο γονίδιο του P1.
 - 4) Επανάληψη από βήμα (2) μέχρι να "πέσουμε" στο γονίδιο εκκίνησης του βήματος (1).
- ❑ Επανάληψη του "κύκλου" μέχρι να χρησιμοποιηθούν όλα τα γονίδια του P1.
- ❑ Οι απόγονοι σχηματίζονται παίρνοντας γονίδια από τους γονείς, από αριστερά προς δεξιά, με τη σειρά που ορίζουν οι "κύκλοι":
 - τα γονίδια αυτούσια για "μονούς" κύκλους (1°, 3°, κ.ο.κ.)
 - τα γονίδια αντεστραμμένα για "ζυγούς" κύκλους (2°, 4°, κ.ο.κ.)



Μετάλλαξη (1/2)

❖ Δυναδικής Αναπαράστασης



❖ Αναπαράστασης Κινητής Υποδιαστολής

Uniform Random

Στην περίπτωση αυτή, το c'_i είναι ένας τυχαίος αριθμός από το διάστημα $[a_i, b_i]$, όπου a_i, b_i είναι αντίστοιχα το ελάχιστο και το μέγιστο όριο της τιμής του c_i .

Creep

Ο τελεστής αυτός λειτουργεί προσθέτοντας μία μικρή θετική ή αρνητική τιμή σε ένα τυχαίο γονίδιο, δηλαδή $c'_i = c_i + (b_i, a_i)r$, όπου r ένας τυχαίος αριθμός από το διάστημα $[-1, 1]$ και a_i, b_i το ελάχιστο και μέγιστο επιτρεπτό όριο του c_i γονιδίου.

Non-Uniform Random

Έστω t η τρέχουσα γενιά και g_{max} ο μέγιστος αριθμός επαναλήψεων. Τότε ο τελεστής αυτός έχει αποτέλεσμα:

$$c'_i = \begin{cases} c_i + \Delta(t, b_i - c_i) & \text{αν } r=0 \\ c_i - \Delta(t, c_i - a_i) & \text{αν } r=1 \end{cases}$$

όπου $\Delta(t, y) = y(1 - r^{(1 - \frac{t}{g_{max}})^b})$, με r ένας τυχαίος αριθμός ίσος με 0 ή 1 και b μία σταθερά που ορίζεται από τον χρήστη και δηλώνει το βαθμό εξάρτησης στο πλήθος των επαναλήψεων. Η συνάρτηση $\Delta(t, y)$ δίνει αποτέλεσμα στο διάστημα $[0, y]$.

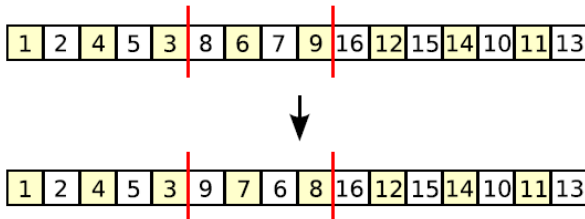
❖ Αναπαράσταση Ακεραίων

❑ Uniform Random και non-Uniform Random

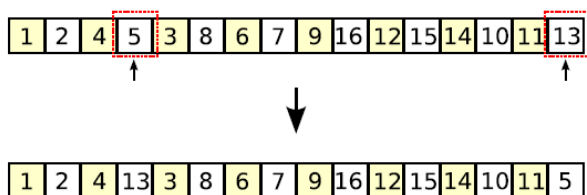
Μετάλλαξη (2/2)

❖ Αναπαράσταση Permutation (Μετάθεσης)

❑ Inversion



❑ Swap



Φυσική Επιλογή

(ή Τελεστής Αντικατάστασης)

- ❖ **Ζητούμενο:** να επιλέξουμε τα "καλύτερα" άτομα του τρέχοντος πληθυσμού (γενιάς) με τα οποία θα δημιουργήσουμε την επόμενη γενιά.
- ❖ Επιλέγουμε από:
 - ☐ μ άτομα της τρέχουσας γενιάς (όχι αυτά που επιλέχθηκαν για αναπαραγωγή!)
 - ☐ λ απογόνους (τα άτομα που προέκυψαν από την αναπαραγωγή)
- ❖ Επιπλέον μπορούμε να κρατούμε το καλύτερο "άτομο" (ελιτισμός – elitism).
 - ☐ Διατηρείται η καλύτερη λύση (στο σύνολο αρχικού πληθυσμού και παιδιών) μέχρι να βρεθεί καλύτερη.

Επιλογή ως προς την Ηλικία

- ❖ Κρατούνται οι απόγονοι και αντικαθιστούν αντίστοιχο πλήθος της τρέχουσας γενιάς.
1. Αντικατάσταση του χειρότερου. Τα λ χειρότερα άτομα του πληθυσμού της τρέχουσας γενιάς, επιλέγονται για να καταλάβουν τη θέση τους οι λ απόγονοι. Στη περίπτωση όπου οι απόγονοι είναι όσο όλος ο πληθυσμός, δηλαδή ισχύει $\lambda = \mu$, τότε αντικαθιστώνται όλα τα άτομα από τους απογόνους.
 2. Τυχαία αντικατάσταση. Η αντικατάσταση των ατόμων του πληθυσμού από τους απογόνους τους πραγματοποιείται με τυχαίο τρόπο. Δηλαδή επιλέγονται τυχαία τα λ άτομα που θα εγκαταλείψουν τον πληθυσμό και την θέση τους καταλαμβάνουν οι απόγονοι.



Επιλογή ως προς την Ποιότητα

❖ Κριτήριο Επιλογής είναι η τιμή ποιότητας.

1. **Tournament.** Από το σύνολο των ατόμων (απόγονοι και τρέχουσα γενιά) επιλέγεται ένα μικρό υποσύνολο k ατόμων και το άτομο με την χειρότερη τιμή ποιότητας απομακρύνεται. Η διαδικασία αυτή επαναλαμβάνεται έως ότου παραμείνουν μ άτομα.
2. **GENITOR.** Από το σύνολο όλων των ατόμων ($\lambda + \mu$) απομακρύνονται τα λ χειρότερα άτομα. Παρόλα αυτά η μέθοδος αυτή μπορεί να οδηγήσει τον ΓΑ σε πρόωρη σύγκλιση, καθώς τον κατευθύνει μόνο προς τα υπάρχοντα ποιοτικότερα άτομα.



Συνθήκη Τερματισμού ΓΑ

Ο ΓΑ εκτελείται επαναληπτικά έως ότου ικανοποιηθεί η συνθήκη ή ικανοποιηθούν οι συνθήκες τερματισμού που έχουν οριστεί. Συνήθως η συνθήκη τερματισμού ορίζεται από το πρόβλημα, ωστόσο οι περισσότερο συνηθισμένες είναι η εξής:

1. Μέγιστο πλήθος επαναλήψεων.
2. Μέγιστος χρόνος εκτέλεσης του ΓΑ.
3. Η μη βελτίωση της ποιότητας του καλύτερου ατόμου για ένα προκαθορισμένο πλήθος γενεών.
4. Η εύρεση της βέλτιστης ή μίας αποδεκτής λύσης.