



Πολυτεχνείο
Κρήτης

Σχολή Μηχανικών Παραγωγής και Διοίκησης

*Εισαγωγή στην Τεχνητή
Νοημοσύνη*

Γενετικοί Αλγόριθμοι

Τσαφαράκης Στέλιος

Μεθευρετικοί Αλγόριθμοι πληθυσμού λύσεων

- Λειτουργούν με πληθυσμό λύσεων (population based algorithms).
- Χρήση πληροφορίας από ένα πληθυσμό λύσεων κατά τη διάρκεια της αναζήτησης.
- Σε κάθε επανάληψη οι λύσεις που αποτελούν τον πληθυσμό αλληλεπιδρούν/συνδυάζονται μεταξύ τους.
- Στόχος η δημιουργία καλύτερων λύσεων σε κάθε επανάληψη.
- Αποτελεσματικότερη αντιμετώπιση της εξάρτησης από τις αρχικές συνθήκες:
 - Πολλαπλές λύσεις => αναζήτηση από διαφορετικά σημεία του χώρου των λύσεων.
- Οι διάφοροι αλγόριθμοι διαφέρουν ως προς τον τρόπο αλληλεπίδρασης και εκμετάλλευσης της πληροφορίας των λύσεων του πληθυσμού

Εξελικτικοί Αλγόριθμοι

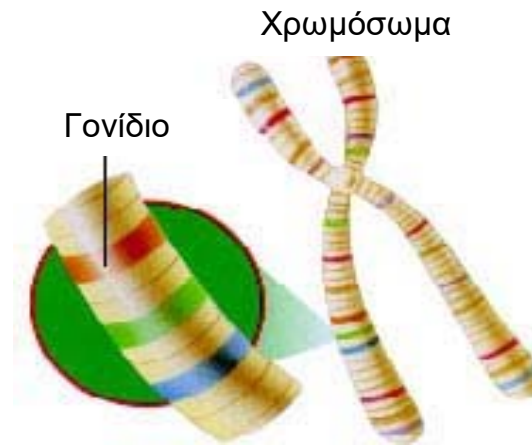
- Η αναζήτηση βασίζεται στη μίμηση της βιολογικής διαδικασίας της εξέλιξης στη φύση (Δαρβίνος).
- Διαφορετικοί πληθυσμοί συνδυάζονται μεταξύ τους και εξελίσσονται.
- Οι αλγόριθμοι βασίζονται σε διαδικασίες της εξέλιξης όπως φυσική επιλογή, αρχή της επιβίωσης του ικανότερου (survival of the fittest) και αναπαραγωγή.
- Χρησιμοποιούν ένα τελεστή διασταύρωσης πάνω σε 2 ή περισσότερους γονείς
- Στους απογόνους που δημιουργούνται εφαρμόζουν ένα τελεστή μετάλλαξης ο οποίος συνήθως μετακινεί τη λύση από ένα σημείο σε κάποιο άλλο με τη χρήση κάποιου βήματος που στηρίζεται σε κάποια κατανομή πιθανότητας.
- Αποτελούν στοχαστικές επαναληπτικές διαδικασίες με σταθερό μέγεθος πληθυσμού σε κάθε γενιά (επανάληψη).

Γενετικοί αλγόριθμοι

- Ο βασικός μηχανισμός των Γενετικών Αλγορίθμων (ΓΑ) είναι εμπνευσμένος από τη Δαρβινική θεωρία της εξέλιξης
- Οι οργανισμοί που δε μπορούν να επιβιώσουν στο περιβάλλον τους πεθαίνουν, ενώ οι υπόλοιποι πολλαπλασιάζονται μέσω της αναπαραγωγής.
- Οι απόγονοι παρουσιάζουν μικρές διαφοροποιήσεις από τους προγόνους τους, ενώ συνήθως υπερισχύουν αυτοί που συγκεντρώνουν τα καλύτερα χαρακτηριστικά.
- Σποραδικά συμβαίνουν τυχαίες μεταλλάξεις, από τις οποίες οι περισσότερες οδηγούν τα μεταλλαγμένα άτομα στο θάνατο, αν και είναι πιθανό να οδηγήσουν στη δημιουργία νέων "καλύτερων" οργανισμών.
- Αν το περιβάλλον μεταβάλλεται με αργούς ρυθμούς, τα διάφορα είδη μπορούν να εξελίσσονται σταδιακά ώστε να προσαρμόζονται σε αυτό.

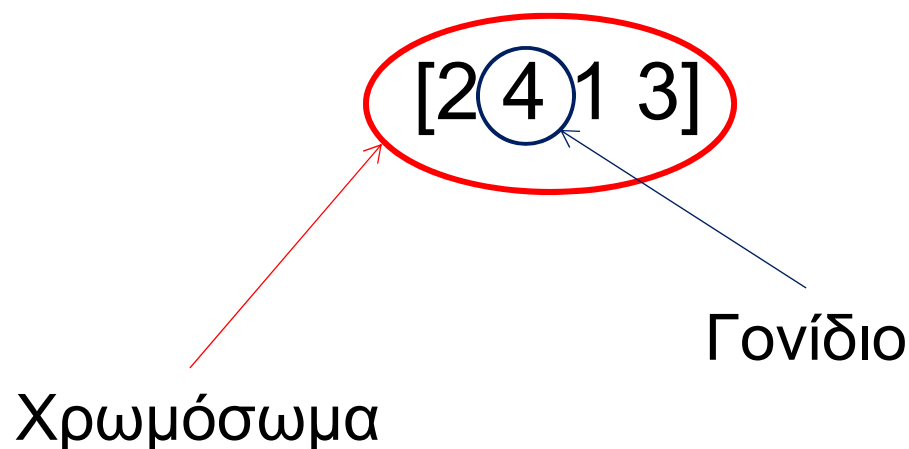
Αναπαράσταση Υποψήφιων Λύσεων (1/2)

- Στους βιολογικούς οργανισμούς, ένα χρωμόσωμα είναι ένα μεγάλο μόριο (ακολουθία) DNA και περιέχει έναν αριθμό γονιδίων.
- Στο πραγματικό DNA το αλφάβητο έχει μήκος τέσσερα και αποτελείται από τα γράμματα A, G, T και C που αντιστοιχούν στα τέσσερα διαφορετικά νουκλεοτίδια (βάσεις) που το συνθέτουν (Adenine, Guanine, Thymine και Cytosine).
- Στον αλγόριθμο το χρωμόσωμα μοντελοποιείται με ένα διάγραμμα των μεταβλητών του προβλήματος.



Αναπαράσταση Υποψήφιων Λύσεων (2/2)

- Μια υποψήφια λύση του προβλήματος αναπαρίσταται από ένα *χρωμόσωμα* (chromosome)
- Το χρωμόσωμα αποτελείται από *γονίδια* (genes – μεταβλητές του προβλήματος) τα οποία λαμβάνουν διάφορες τιμές (alleles).
- Π.χ. σε ένα πρόβλημα με 4 βασίλισσες μια υποψήφια λύση μπορεί να αναπαρασταθεί από ένα χρωμόσωμα όπου κάθε γόνος αναπαριστά τη θέση μιας βασίλισσας και παίρνει τιμές 1, 2, 3, ή 4 ανάλογα με τη θέση της βασίλισσας στην αντίστοιχη στήλη



4		♔		
3			♔	
2	♔			
1		♔		

Γενετικοί αλγόριθμοι - Λειτουργία

- Οι ΓΑ είναι αλγόριθμοι πληθυσμού λύσεων (population based) που λειτουργούν με επαναλήψεις (iterations)
- Σε κάθε επανάληψη τα χρωμοσώματα του πληθυσμού (υποψήφιες λύσεις) του προβλήματος συνδυάζουν το γενετικό τους υλικό (γονίδια) με στόχο να παράξουν καταλληλότερα χρωμοσώματα (καλύτερες λύσεις για το πρόβλημα).
- Η *καταλληλότητα* (fitness) ενός χρωμοσώματος αξιολογείται με βάση μια *συνάρτηση καταλληλότητας* (fitness function - αντικειμενική συνάρτηση του προβλήματος)
- Τα χρωμοσώματα συνδυάζουν το γενετικό τους υλικό με τη χρήση του τελεστή της *διαστάυρωσης* (crossover)
- Το γενετικό υλικό ορισμένων χρωμοσωμάτων υπόκειται τυχαίες μεταβολές με τη χρήση του τελεστή της *μετάλλαξης* (mutation)
- Μετά την εφαρμογή των 2 τελεστών επιλέγονται τα χρωμοσώματα που θα επιβιώσουν στην επόμενη γενιά (επανάληψη του αλγορίθμου)
- Με την παραπάνω διαδικασία ο αλγόριθμος αναζητά καλύτερες λύσεις σε κάθε επανάληψη

Τελεστές Γενετικού Αλγορίθμου

- Διασταύρωση (crossover): 2 χρωμοσώματα (γονείς-parents) συνδυάζουν το γενετικό τους υλικό ανταλλάσσοντας τμήματα και παράγουν 2 νέα χρωμοσώματα (απόγονοι-offspring).
- Μετάλλαξη (mutation): Κάθε γόνος ενός χρωμοσώματος μπορεί να αλλάξει τιμή με μια πιθανότητα p_m .
- Επιλογή (selection): Μετά την ολοκλήρωση της αναπαραγωγής (διασταύρωση & μετάλλαξη) επιλέγονται τα χρωμοσώματα (συνήθως) με τη μεγαλύτερη καταλληλότητα (survival of the fittest) τα οποία αποτελούν τον πληθυσμό που επιβιώνει στην επόμενη γενιά.

Διαδικασία Επιλογής Γονέων (1/2)

- Απόδοση πιθανοτήτων επιλογής προς αναπαραγωγή στα μέλη ενός πληθυσμού υποψηφίων λύσεων:
 - Κάποιοι γονείς με υψηλή καταλληλότητα ενδέχεται να επιλεγούν προς αναπαραγωγή περισσότερες από 1 φορές
 - Κάποιοι γονείς με χαμηλή καταλληλότητα ενδέχεται να μην επιλεγούν καθόλου.
- Κατά τη διαδικασία επιλογής, αρχικά οι υποψήφιες λύσεις αντιγράφονται σε μια δεξαμενή ζευγαρώματος (*mating pool*).
 - Σε αυτήν αντιγράφονται μέλη του αρχικού πληθυσμού, με πιθανότητα ανάλογη της καταλληλότητάς τους.
 - Για την επιλογή των χρωμοσωμάτων που θα αντιγραφούν στη δεξαμενή ζευγαρώματος χρησιμοποιούνται αρκετές τεχνικές.

Διαδικασία Επιλογής Γονέων (2/2)

Τεχνική της ρουλέτας

1. Παράγεται το άθροισμα S όλων των τιμών αξιολόγησης των υποψηφίων λύσεων.
2. Επιλέγεται ένας τυχαίος αριθμός n , από το 0 μέχρι το S , χρησιμοποιώντας συνάρτηση ομοιόμορφης κατανομής για τη δημιουργία των τυχαίων αριθμών.
3. Επαναληπτικά εξετάζεται κάθε υποψήφια λύση και η τιμή του προστίθεται σε έναν καταχωρητή K .
4. Αν η τιμή του K γίνει μεγαλύτερη ή ίση του n , η λύση επιλέγεται και ο K μηδενίζεται. Στην αντίθετη περίπτωση εκτελείται πάλι το 3.
5. Αν δεν έχει επιλεγεί ικανοποιητικός αριθμός υποψηφίων λύσεων εκτελείται το 2, αλλιώς ο αλγόριθμος τερματίζει.

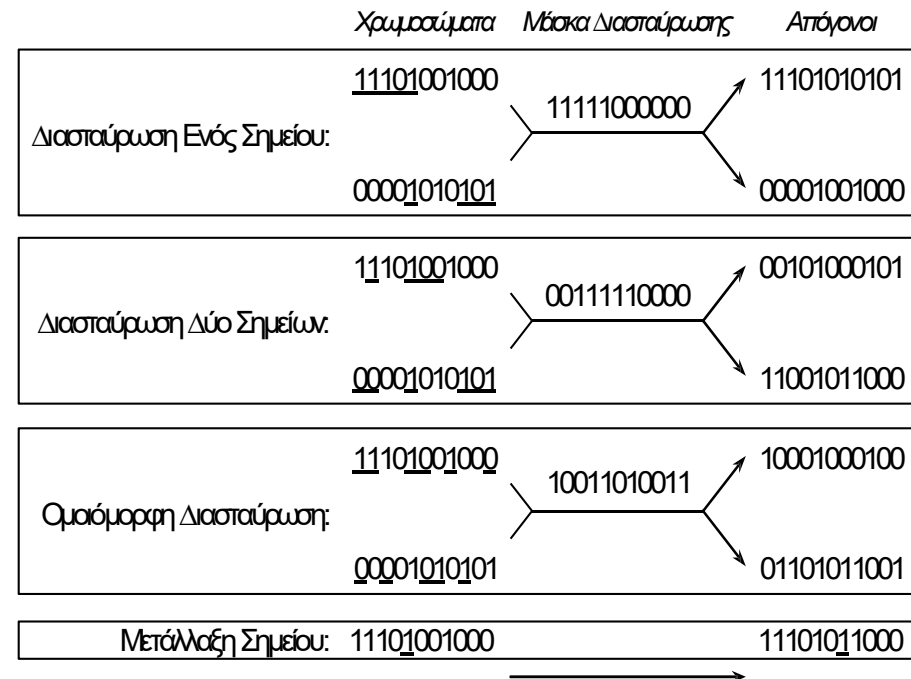
Άλλη τεχνική: επιλογή αναλογικής καταλληλότητας (fitness proportionate selection)

- Πιθανότητα επιλογής χρωμοσώματος x_i :

$$P(x_i) = \frac{\text{Καταλληλότητα}(x_i)}{\sum_{j=1}^N \text{Καταλληλότητα}(x_j)}$$

Αναπαραγωγή

- Αναπαραγωγή είναι η διαδικασία δημιουργίας απογόνων.
- Σε αυτή εμπλέκονται ένα σύνολο από παραλλαγές των 2 βασικών τελεστών οι οποίοι αντιστοιχούν σε διαδικασίες της βιολογικής εξέλιξης.
- Οι πιο συνηθισμένοι τελεστές είναι:
 - Διασταύρωση ενός σημείου (single- point crossover)
 - Διασταύρωση δύο σημείων (two-point crossover)
 - Ομοιόμορφη διασταύρωση (uniform crossover)
 - Μετάλλαξη σημείου (point mutation)



Ακέραιη αναπαράσταση

Διασταύρωση

- Ενός σημείου (one point):

P1 [50 44 30 08 23 | 82 13 04 09]

P2 [04 21 05 49 12 65 84 10 32]

O1 [50 44 30 08 23 65 84 10 32]

O2 [04 21 05 49 12 82 13 04 09]

- Δύο σημείων:

P1 [50 44 30 | 08 23 82 13 | 04 09]

P2 [04 21 05 49 12 65 84 10 32]

O1 [50 44 30 49 12 65 84 04 09]

O2 [04 21 05 08 23 82 13 10 32]

- Ομοιόμορφη (uniform)

Μετάλλαξη

[50 44 30 08 23 82 13 04 09]

[50 44 91 08 23 82 13 04 75]

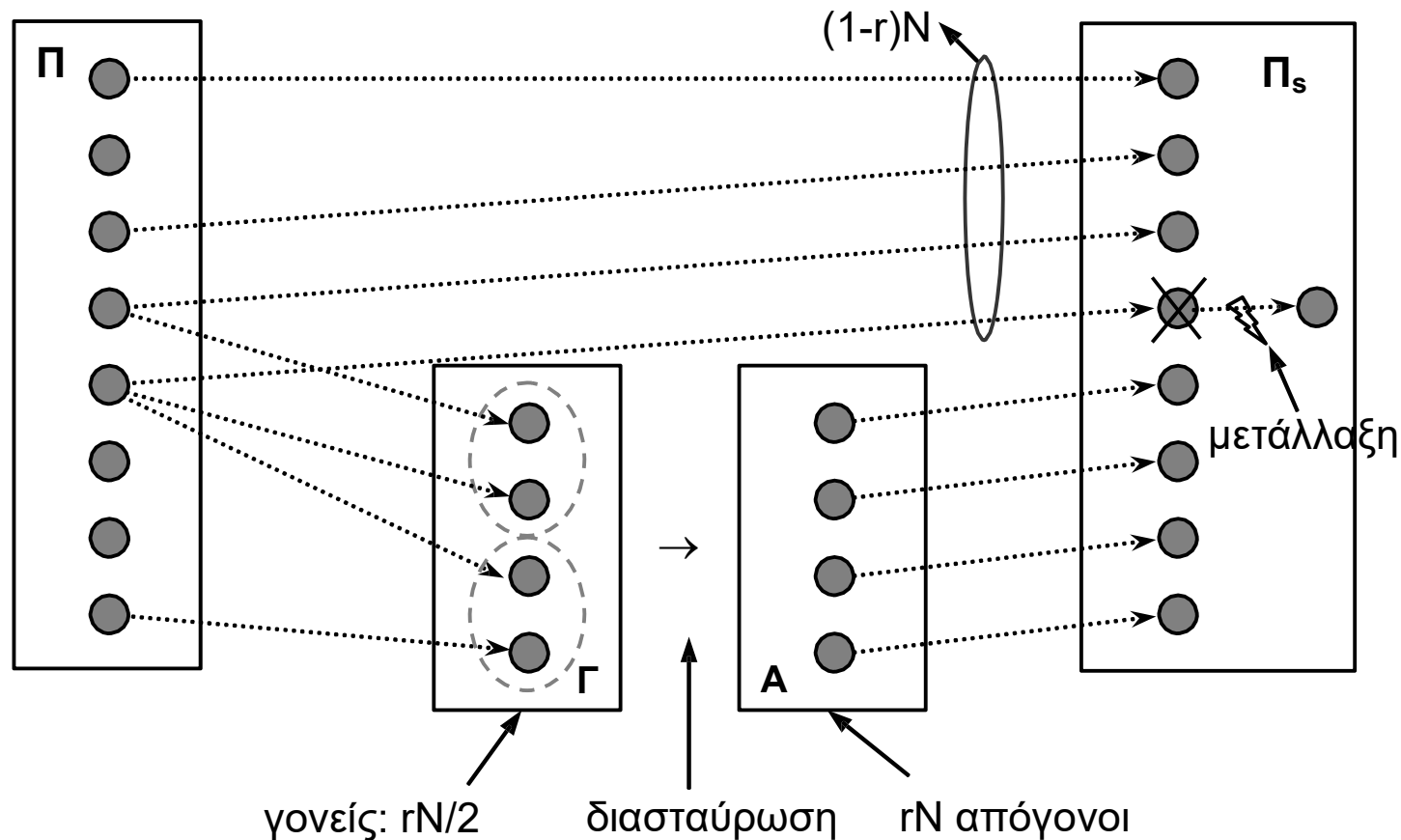
Σύγκλιση Αλγορίθμου

- **Σύγκλιση:** Η επικράτηση ενός χρωμοσώματος ή μικρών παραλλαγών του, σε μεγάλο ποσοστό στον πληθυσμό.
- Με έναν αποδοτικό γενετικό αλγόριθμο, ο πληθυσμός θα πρέπει μετά από αρκετές επαναλήψεις να συγκλίνει προς το ολικό βέλτιστο.
- Ένα γονίδιο συγκλίνει, όταν έχει την ίδια τιμή στο 95% των χρωμοσωμάτων.
- Ένας πληθυσμός συγκλίνει, όταν όλα τα γονίδιά του έχουν συγκλίνει.
- Εναλλακτικά κριτήρια τερματισμού:
 - Συγκεκριμένος αριθμός επαναλήψεων
 - Μη εύρεση καλύτερης λύσης για συνεχόμενες επαναλήψεις

Μερική Ανανέωση Πληθυσμού

- Εναλλακτική προσέγγιση στο θέμα της εξέλιξης του πληθυσμού είναι η μη ανανέωση ολόκληρου του πληθυσμού σε κάθε γενιά, αλλά ενός μέρους αυτού.
- Η μέθοδος της μερικής ανανέωσης είναι πιο κοντά στην πραγματικότητα, αφού εκεί συνυπάρχουν πάντα σε κάποιο βαθμό οι διαφορετικές γενεές.
- Μάλιστα δίνεται η δυνατότητα στους απογόνους να ανταγωνιστούν τους γονείς τους, επικρατώντας και πάλι ο καλύτερος.
- Εξασφαλίζει ότι το καταλληλότερο χρωμόσωμα θα επιβιώσει στην επόμενη γενιά.
- Διάφορες μέθοδοι επιλογής μεταξύ των απογόνων και των γονέων για την επόμενη γενιά.
- Ο πληθυσμός των χρωμοσωμάτων παραμένει σταθερός

Σχηματική λειτουργία γενικού γενετικού αλγορίθμου



Παράμετροι Γενετικού Αλγορίθμου

- Αναπαράσταση λύσης (δυαδική, ακέραια, συνεχής, μεικτή)
- Μέγεθος πληθυσμού (N χρωμοσώματα)
- Αρχικοποίηση πληθυσμού (συνήθως τυχαία)
- Τρόπος επιλογής γονέων για διασταύρωση
- Είδος διασταύρωσης (ενός ή περισσοτέρων σημείων – ομοιόμορφη)
- Ποσοστό διασταύρωσης (crossover rate – πιθανότητα οι επιλεγμένοι γονείς να ζευγαρώσουν: συνήθως κοντά στο 1)
- Ποσοστό μετάλλαξης (mutation rate – πόσα χρωμοσώματα θα επιλεγούν για να τροποποιήσουν ένα ή περισσότερα γονίδια τους)
- Στρατηγική μετανάστευσης πληθυσμού στην επόμενη γενιά
- Κριτήριο τερματισμού

Βήματα Γενετικού Αλγορίθμου

- Αρχικά καθορίζεται ο τρόπος κωδικοποίησης (encoding) των λύσεων στο χρωμόσωμα (π.χ. δυαδική ή ακέραια αναπαράσταση)
 1. Ο αρχικός πληθυσμός των n χρωμοσωμάτων δημιουργείται τυχαία, εκτός αν γνωρίζουμε κάποιες καλές λύσεις.
 2. Αξιολογείται η καταλληλότητα των χρωμοσωμάτων του αρχικού πληθυσμού
 3. Επιλέγονται οι γονείς και υλοποιείται η διασταύρωση
 4. Στους απογόνους που δημιουργούνται εφαρμόζεται η μετάλλαξη.
 5. Αξιολογείται η καταλληλότητα του νέου πληθυσμού που προκύπτει (γονείς + απόγονοι)
 6. Επιλέγονται τα n χρωμοσώματα (συνήθως τα καταλληλότερα) τα οποία επιβιώνουν στην επόμενη γενιά
 7. Η διαδικασία επαναλαμβάνεται από το βήμα 2 μέχρι να ικανοποιηθεί η συνθήκη τερματισμού.

Ψευδοκώδικας Γενετικού Αλγόριθμου

1. Δημιούργησε έναν αρχικό πληθυσμό $P(t=0)$, με N χρωμοσώματα.
2. Υπολόγισε την καταλληλότητα κάθε χρωμοσώματος.
3. Όσο δεν ισχύει κάποια συνθήκη τερματισμού:
 - α. Διασταύρωση:
 - i. Επέλεξε τα χρωμοσώματα-γονείς (parents) από τον πληθυσμό P .
 - ii. Με πιθανότητα Cr ζευγάρωσε 2 γονείς για να δημιουργήσεις 2 χρωμοσώματα-απογόνους (offspring).
 - β. Μετάλλαξη: Επέλεξε κάθε χρωμόσωμα με πιθανότητα Pr και τροποποίησε το γενετικό του υλικό.
 - γ. Υπολόγισε την καταλληλότητα κάθε νέου χρωμοσώματος
 - δ. Δημιούργησε το νέο πληθυσμό $P(t+1)$ επιλέγοντας N χρωμοσώματα από τους απογόνους (και τους γονείς) βάση (και) της καταλληλότητάς τους
4. Επέστρεψε το (τα) χρωμόσωμα(τα) με τη μεγαλύτερη καταλληλότητα.

Πλεονεκτήματα Γενετικών Αλγορίθμων

- Μπορούν να λύσουν είτε διακριτά είτε συνεχή προβλήματα.
- Δεν χρειάζονται πληροφορίες από παραγώγους.
- Μπορούν να κάνουν ταυτόχρονα αναζήτηση σε ένα πολύ μεγάλο μέρος του χώρου λύσεων.
- Μπορούν να λύσουν γρήγορα το πρόβλημα ανεξάρτητα από το μέγεθός του (πλήθος μεταβλητών ή λύσεων).
- Μπορούν να ξεφύγουν εύκολα από κάποιο τοπικό ελάχιστο.
- Μπορούν να δώσουν παραπάνω από μία τοπικά βέλτιστες λύσεις.

Παράδειγμα Αναπαραγωγής (1/2)

- Έστω ένας πληθυσμός αποτελείται από 4 χρωμοσώματα.
- Στην 3^η στήλη φαίνονται οι τιμές της συνάρτησης καταλληλότητας, ενώ στην 4^η οι πιθανότητες επιλογής προς αναπαραγωγή κάθε χρωμοσώματος, ανηγμένες στο διάστημα $[0, 1]$.
- Οι τιμές της 4^{ης} στήλης προκύπτουν από το λόγο της καταλληλότητας του χρωμοσώματος προς το άθροισμα των καταλληλοτήτων όλων των χρωμοσωμάτων του πληθυσμού.

Χρωμοσώματα		Καταλληλότητα	Πιθανότητα επιλογής
A	000110010111	8	0.32
B	111010101100	6	0.24
Γ	001110101001	6	0.24
Δ	111011011100	5	0.20

- Έστω ότι το αποτέλεσμα της διαδικασίας επιλογής είναι το $[A, B, B, \Gamma]$ και τα ζευγάρια που σχηματίζονται τυχαία είναι τα (B, A) και (B, Γ)

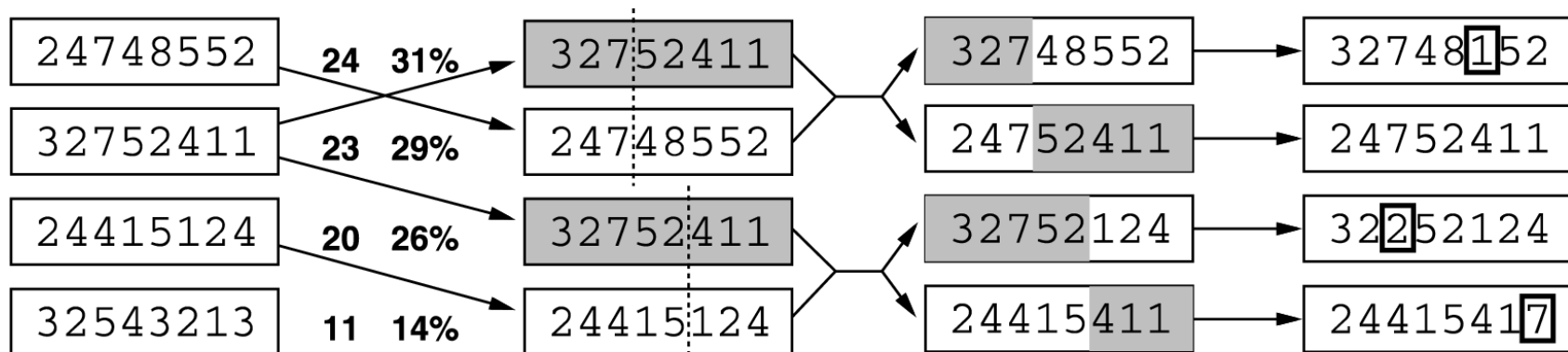
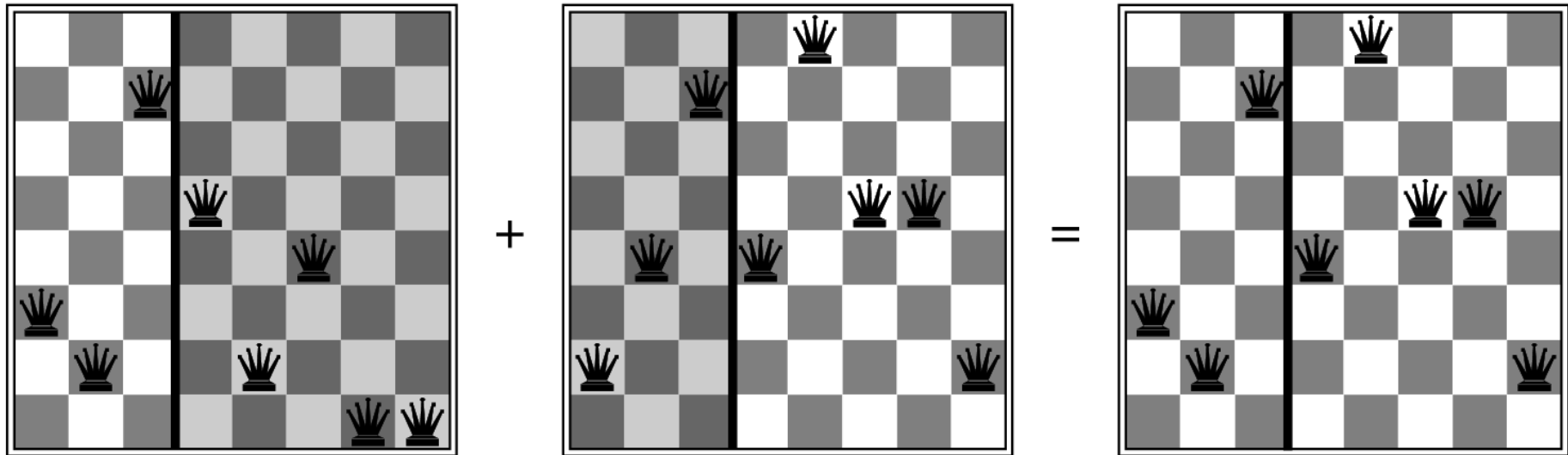
Παράδειγμα Αναπαραγωγής (2/2)

- Ως τελεστής αναπαραγωγής του πρώτου ζευγαριού επιλέγεται η διασταύρωση ενός σημείου με μάσκα διασταύρωσης "111100000000"
- Ενώ για το δεύτερο ο ίδιος τελεστής με μάσκα διασταύρωσης "111111111000"

Γονέας 1	(B) 1110-10101100	(B) 1110-10101100
Γονέας 2	(A) 0001-10010111	(A) 0001-10010111
Απόγονος	1 ^{ος}	2 ^{ος}
Διασταύρωση	1110-10010111	0001-10101100
Μετάλλαξη	1110-10010111	0001-10101100

Γονέας 3	(B) 111010101-100	(B) 111010101-100
Γονέας 4	(Γ) 001110101-001	(Γ) 001110101-001
Απόγονος	3 ^{ος}	4 ^{ος}
Διασταύρωση	111010101-001	001110101-100
Μετάλλαξη	111 <u>1</u> 10101-001	001110101-10 <u>1</u>

8 Βασίλισσες



(α)
Αρχικός
πληθυσμός

(β)
Συνάρτηση
καταλληλότητας

(γ)
Επιλογή

(δ)
Διασταύρωση

(ε)
Μετάλλαξη