

# ΠΛΗ31

## ΕΝΟΤΗΤΑ 4: ΓΕΝΕΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

### Μάθημα 4.3: Θεμέλια των Γενετικών Αλγορίθμων

Δημήτρης Ψούνης



[www.psounis.gr](http://www.psounis.gr)



# ΠΕΡΙΕΧΟΜΕΝΑ

## **A.Θεωρία**

### **1. Εισαγωγικοί Ορισμοί**

1. Ο ορισμός του σχήματος
2. Τάξη και ορισμός του σχήματος
3. Συνολική και Μέση Απόδοση του Πληθυσμού

### **2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου**

1. Απόδοση σχήματος σε μία χρονική στιγμή
2. Διαίσθηση πίσω από τους τύπους
3. Επίδραση του τελεστή επιλογής
4. Επίδραση του τελεστή διασταύρωσης
5. Επίδραση του τελεστή μετάλλαξης
6. Το θεώρημα των σχημάτων

### **3. Επικράτηση - Εξαφάνιση Σχημάτων**

1. Ορισμός
2. Παραδείγματα

## **B.Ασκήσεις**



# A. Θεωρία

## 1. Εισαγωγικοί Ορισμοί

### 1. Ο ορισμός του σχήματος

**Ορισμός:** Έστω  $\Sigma$  το αλφάβητο των συμβόλων που χρησιμοποιεί ο γενετικός αλγόριθμος για την κωδικοποίηση των χρωμοσωμάτων του πληθυσμού. Ένα σχήμα  $S$  (ή πρότυπο  $S$ ) είναι ένα χρωμόσωμα που χρησιμοποιεί το  $*$  (διαβάζεται αδιάφορο σύμβολο) το οποίο μπορεί να αντικαταστασθεί από οποιοδήποτε σύμβολο του αλφαβήτου.

Έτσι ένα σχήμα αναπαριστά με έναν ενιαίο τρόπο όλα τα χρωμοσώματα που μπορούν να ταιριάξουν με το σχήμα αντικαθιστώντας το  $*$  με τα σύμβολα που απαιτούνται από το αλφάβητο.

### Παράδειγμα

Στο σχήμα  $S=11*10$  ταιριάζουν οι δύο συμβολοσειρές  $\{11010, 11110\}$

Στο σχήμα  $S=*1*1$  ταιριάζουν οι τέσσερις συμβολοσειρές  $\{0101, 0111, 1101, 1111\}$



# A. Θεωρία

## 1. Εισαγωγικοί Ορισμοί

### 1. Ο ορισμός του σχήματος

Ισχύουν τα ακόλουθα για τον ορισμό του σχήματος:

Έστω ότι έχουμε ένα πρόβλημα, που το χρωμόσωμα αναπαρίσταται με μια δυαδική συμβολοσειρά μήκους  $n$ :

- Ένα σχήμα με κανένα  $*$  θα αναπαριστά μία συμβολοσειρά
- Ένα σχήμα με  $k$   $*$  θα αναπαριστά  $2^k$  συμβολοσειρές
- Ένα σχήμα που αποτελείται μόνο από  $*$  θα αναπαριστά  $2^n$  συμβολοσειρές

### Παράδειγμα

Στο σχήμα  $S=11*100**1$  ταιριάζουν  $2^3=8$  συμβολοσειρες

Στο σχήμα  $S=*****$  ταιριάζουν  $2^9=512$  συμβολοσειρες



# A. Θεωρία

## 1. Εισαγωγικοί Ορισμοί

### 1. Ο ορισμός του σχήματος

Συμβολίζουμε με  $c$  τον πληθάρημο του αλφαβήτου ( $c=|\Sigma|$  δηλαδή το πλήθος των συμβόλων του αλφαβήτου).

Έστω ότι έχουμε ένα πρόβλημα, που το χρωμόσωμα αναπαρίσταται με μια δυαδική συμβολοσειρά μήκους  $n$ :

- Τα δυνατά σχήματα που μπορούν να κατασκευαστούν είναι  $(c+1)^n$ 
  - Αφού κάθε θέση μπορεί να πάρει οποιοδήποτε σύμβολο του αλφαβήτου και των χαρακτήρα \*
- Ενώ μία συμβολοσειρά ταιριάζει σε  $2^n$  διαφορετικά σχήματα.
  - Αφού η συμβολοσειρά θα ταιριάζει μόνο σε σχήματα που στην αντίστοιχη θέση ενός συμβόλου θα πρέπει να έχει ή το ίδιο σύμβολο ή το \*

### Παράδειγμα

Η συμβολοσειρά 101 ταιριάζει στα 8 σχήματα: 101, 10\*, 1\*1, \*01, 1\*\*, \*0\*, \*\*1, \*\*\*



# A. Θεωρία

## 1. Εισαγωγικοί Ορισμοί

### 2. Τάξη και οριστικό μήκος σχήματος

#### Τάξη ενός σχήματος $o(S)$ :

- είναι αριθμός των θέσεων με 0 και 1
- (δηλαδή το πλήθος των θέσεων που δεν έχουν \*)
  - Προσδιορίζει πόσο ειδικό είναι το σχήμα (μεγάλη τάξη=>μεγάλη εξειδίκευση, μικρή τάξη=>μικρή εξειδίκευση των συμβολοσειρών)

#### Οριστικό μήκος σχήματος $\delta(S)$ :

- είναι η απόσταση της πρώτης και της τελευταίας σταθερής θέσης
- (υπολογίζεται ως η διαφορά της τελευταίας θέσης που έχουμε σύμβολο μείον την πρώτη θέση που έχουμε σύμβολο)
  - Προσδιορίζει την πυκνότητα της πληροφορίας που περιέχεται στο σχήμα

### Παράδειγμα

Το σχήμα  $S=11^{**}1$  έχει τάξη 3 και οριστικό μήκος  $6-1=5$

Το σχήμα  $S=*0*111*$  έχει τάξη 4 και οριστικό μήκος  $6-2=4$

# A. Θεωρία

## 1. Εισαγωγικοί Ορισμοί

### 3. Συνολική και Μέση Απόδοση Πληθυσμού

Έστω  $N$  το μέγεθος του πληθυσμού (pop\_size) και  $f(v)$  η αντικειμενική συνάρτηση ενός γενετικού αλγορίθμου:

Αν  $t$  είναι μία χρονική στιγμή της εκτέλεσης του αλγορίθμου:

- Συνολική απόδοση του πληθυσμού σε μία χρονική στιγμή ορίζεται ως το άθροισμα όλων των αξιολογήσεων του πληθυσμού:
  - $F(t) = \sum_{i=1}^N f(v_i)$
- Μέση απόδοση του πληθυσμού σε μία χρονική στιγμή ορίζεται ως η συνολική απόδοση προς το μέγεθος του πληθυσμού
  - $\bar{F}(t) = \frac{F(t)}{N}$



# A. Θεωρία

## 2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου

### 1. Απόδοση σχήματος

Σε μία χρονική στιγμή (επαναληπτικό κύκλο του αλγόριθμου)  $t$  έχουμε pop\_size (=N) χρωμοσώματα.

- Έστω ένα σχήμα  $S$  και έστω ότι είναι  $p$  το πλήθος των χρωμοσωμάτων που ταιριάζουν με το σχήμα  $S$  την χρονική στιγμή  $t$ .

Ορίζουμε ως:

- $eval(S,t)$  την μέση απόδοση ενός σχήματος την χρονική στιγμή  $t$  ως

$$eval(S, t) = \frac{1}{p} \sum_{j=1}^p eval(v_j)$$

- Με απλά λόγια η μέση απόδοση του σχήματος είναι ο μέσος όρος των αποδόσεων των χρωμοσωμάτων που ταιριάζουν με το σχήμα την τρέχουσα χρονική στιγμή.

Ορίζουμε επίσης ως:

- Η πιθανότητα επιλογής μίας συμβολοσειράς που ταιριάζει στο σχήμα  $S$ , ως:

$$\frac{eval(S, t)}{F(t)}$$





# A. Θεωρία

## 2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου

### 2. Διάισθηση πίσω από τους τύπους

- Ορίζουμε ως  $\xi(S, t)$  το πλήθος των χρωμοσωμάτων που ταιριάζουν με το σχήμα  $S$  την χρονική στιγμή  $t$ .
- Θα αποδείξουμε ότι το πλήθος των χρωμοσωμάτων που ταιριάζουν με το σχήμα  $S$  την χρονική στιγμή  $t+1$  δίνεται από τον τύπο:

$$\xi(S, t + 1) \geq \xi(S, t) \cdot \frac{eval(S, t)}{\bar{F}(t)} \cdot \left[ 1 - p_c \frac{\delta(S)}{m - 1} - o(S) \cdot p_m \right]$$

- Η παραπάνω σχέση θα κατασκευαστεί σταδιακά μελετώντας την επίδραση των 3 γενετικών τελεστών.
  - Του τελεστή επιλογής
  - Του τελεστή διασταύρωσης
  - Του τελεστή μετάλλαξης.



# A. Θεωρία

## 2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου

### 3. Διαδικασία της Επιλογής

Μετά το βήμα της επιλογής ο αριθμός των συμβολοσειρών που ταιριάζουν στο σχήμα  $S$  είναι  $\xi(S, t+1)$ .

Ισχύει ότι:

$$\xi(S, t + 1) = \xi(S, t) \cdot N \cdot \frac{eval(S, t)}{F(t)}$$

που πρακτικά μας λέει ότι:

- Το  $eval(S, t)$  είναι η μέση απόδοση των συμβολοσειρών του σχήματος
- Το  $\xi(S, t)$  είναι το πλήθος των συμβολοσειρών του σχήματος  $S$ 
  - Άρα το γινόμενο τους είναι η απόδοση των συμβολοσειρών
- Το  $F(t)$  είναι η συνολική απόδοση του πληθυσμού
  - Άρα διαιρώντας με το  $F(t)$  παίρνουμε την πιθανότητα επιλογής κάποιου μέλους του σχήματος
- Η ρουλέτα θα γυριστεί  $N$  φορές, άρα ο τύπος μας δίνει πόσα χρωμοσώματα θα επιλεγθούν από την ρουλέτα που ταιριάζουν στο σχήμα  $S$ .



# A. Θεωρία

## 2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου

### 3. Διαδικασία της Επιλογής

Αποδεικνύεται επίσης ότι:

- $$\xi(S, t + 1) = \xi(S, t) \cdot \frac{eval(S, t)}{\bar{F}(t)}$$

Συνεπώς αν  $\frac{eval(S, t)}{\bar{F}(t)} > 1$  (αρα  $eval(S, t) > \bar{F}(t)$  δηλαδή η μέση απόδοση του σχήματος είναι μεγαλύτερη από την μέση απόδοση του πληθυσμού), τότε αυξάνει το πλήθος των συμβολοσειρών του σχήματος στην επόμενη γενιά.

Αντίθετα αν  $eval(S, t) < \bar{F}(t)$  τότε το πλήθος των συμβολοσειρών του σχήματος αναμένεται να μειωθεί στην επόμενη γενιά.



# A. Θεωρία

## 2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου

### 3. Διαδικασία της Διασταύρωσης

Τα ακόλουθα αφορούν την γνωστή διαδικασία της διασταύρωσης μονού σημείου σε συμβολοσειρές που έχουν  $m$  bits στην κωδικοποίησή τους.

Ορίζουμε τις πιθανότητες επιβίωσης και καταστροφής ενός σχήματος ως εξής:

Πιθανότητα Καταστροφής Σχήματος:

$$p_D(S) = p_c \frac{\delta(S)}{m-1}$$

Όπου:  $p_c$  είναι η πιθανότητα διασταύρωσης και  $\delta(S)$  είναι το οριστικό μήκος του σχήματος.

Πιθανότητα Επιβίωσης Σχήματος:

$$p_S(S) = 1 - p_D(S) = 1 - p_c \frac{\delta(S)}{m-1}$$

Όπου:  $p_c$  είναι η πιθανότητα διασταύρωσης και  $\delta(S)$  είναι το οριστικό μήκος του σχήματος.

# A. Θεωρία

## 2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου

### 3. Διαδικασία της Διασταύρωσης

Από τα παραπάνω αποδεικνύεται ότι ο αναμενόμενος αριθμός συμβολοσειρών του σχήματος  $S$  μετά την διαδικασία της διασταύρωσης είναι::

$$\xi(S, t + 1) \geq \xi(S, t) \cdot \frac{eval(S, t)}{\bar{F}(t)} \cdot p_s(S)$$

Δηλαδή πολλαπλασιάζουμε το αναμενόμενο πλήθος των συμβολοσειρών μετά τη διαδικασία της επιλογής με την πιθανότητα επιβίωσης των συμβολοσειρών του σχήματος.

Ισοδύναμα η παραπάνω σχέση γράφεται:

$$\xi(S, t + 1) \geq \xi(S, t) \cdot \frac{eval(S, t)}{\bar{F}(t)} \cdot \left[ 1 - p_c \frac{\delta(S)}{m - 1} \right]$$

Συνεπώς σχήματα με μικρό οριστικό μήκος και άνω του μέσου όρου θα αυξάνουν το πλήθος των συμβολοσειρών τους με εκθετικό ρυθμό.



# A. Θεωρία

## 2. Σχήματα κατά την εκτέλεση του γενετικού αλγορίθμου

### 4. Διαδικασία της Μετάλλαξης

Η πιθανότητα επιβίωσης ενός σχήματος  $S$  κατά τη διαδικασία της μετάλλαξης είναι:

$$p_S(S) = (1 - p_m)^{o(S)} \approx 1 - o(S) \cdot p_m$$

Όπου:  $p_m$  είναι η πιθανότητα μετάλλαξης και  $o(S)$  η τάξη του σχήματος.

Συνεπώς ισχύει για τον αναμενόμενο αριθμό συμβολοσειρών του σχήματος  $S$  μετά από την διαδικασία της μετάλλαξης ότι::

$$\xi(S, t + 1) \geq \xi(S, t) \cdot \frac{eval(S, t)}{\bar{F}(t)} \cdot \left[ 1 - p_c \frac{\delta(S)}{m - 1} - o(S) \cdot p_m \right]$$

Από τον παραπάνω τύπο αναδεικνύεται το ακόλουθο (σημαντικό) **θεώρημα των σχημάτων**:

Σχήματα άνω του μέσου όρου απόδοσης, με μικρό οριστικό μήκος και μικρή τάξη λαμβάνουν εκθετικά αυξανόμενες συμβολοσειρές σε διαδοχικές γενιές ενός γενετικού αλγορίθμου.



# A. Θεωρία

## 3. Επικράτηση – Εξαφάνιση Σχημάτων

### 1. Ορισμός

Ο παρακάτω τύπος δίνει το πλήθος των συμβολοσειρών των συμβολοσειρών που ταιριάζουν στο σχήμα  $S$  μετά από  $K$  γενιές, αν εφαρμόζεται μόνο επιλογή (όχι διασταύρωση και μετάλλαξη)

$$\xi(S, t + K) = \xi(S, t) \cdot (1 + \varepsilon)^K$$

Όπου  $\varepsilon$  είναι η επί τοις εκατό απόκλιση της μέσης απόδοσης του πληθυσμού σε σχέση με την μέση απόδοση του σχήματος και δίνεται από τον τύπο:

$$\varepsilon = \frac{eval(S)}{\bar{F}} - 1$$

Όπου  $eval(S)$  είναι η μέση απόδοση του σχήματος και  $\bar{F}$  είναι η μέση απόδοση του πληθυσμού.



# A. Θεωρία

## 3. Επικράτηση – Εξαφάνιση Σχημάτων

### 2. Παραδείγματα

**Παράδειγμα 1:** Έστω η αντικειμενική συνάρτηση  $f$ , που όταν δέχεται είσοδο τη δυαδική συμβολοσειρά  $x$ , με μήκος  $\lambda = 4$ , να δίνει αποτέλεσμα που ισούται με τον ακέραιο που αναπαριστάται από το δυαδικό αριθμό  $x$  (π.χ.  $f(0011)=3$ ,  $f(1111)=15$  κτλ.).

1. Ποια είναι η μέση απόδοση του σχήματος  $S_1=1^{****}$ , σε σχέση με την αντικειμενική συνάρτηση  $f$ ; Ποια είναι η μέση απόδοση του σχήματος  $S_2=0^{****}$  σε σχέση με τη συνάρτηση  $f$ ; Να υποθέσετε ότι ο πληθυσμός αποτελείται 16 χρωμοσώματα, ένα για κάθε διαφορετικό δυαδικό αριθμό από το 0000 έως το 1111
2. Ποιο από τα δύο σχήματα θα επικρατήσει σε ολόκληρο τον πληθυσμό μετά από κάποιες γενιές και ποιο θα εξαφανιστεί από τον πληθυσμό; Να υπολογίσετε μετά από πόσες γενιές θα συμβεί αυτό, για κάθε περίπτωση, αν θεωρήσουμε την επίδραση της διασταύρωσης και της μετάλλαξης είναι αμελητέα.





# A. Θεωρία

## 3. Επικράτηση – Εξαφάνιση Σχημάτων

### 2. Παραδείγματα

#### Λύση Ερώτημα 1:

Στο σχήμα  $S_1=1^{***}$  ταιριάζουν οι παρακάτω συμβολοσειρές:

$$1000 \Rightarrow \text{eval}(1000) = 8$$

$$1001 \Rightarrow \text{eval}(1001) = 9$$

$$1010 \Rightarrow \text{eval}(1010) = 10$$

$$1011 \Rightarrow \text{eval}(1011) = 11$$

$$1100 \Rightarrow \text{eval}(1100) = 12$$

$$1101 \Rightarrow \text{eval}(1101) = 13$$

$$1110 \Rightarrow \text{eval}(1110) = 14$$

$$1111 \Rightarrow \text{eval}(1111) = 15$$

Επομένως η μέση απόδοση του σχήματος  $S_1$  ισούται με:

$$\bar{F}(S_1) = (8 + 9 + 10 + 11 + 12 + 13 + 14 + 15) / 8 = 11.5$$

Στο σχήμα  $S_2=0^{***}$  ταιριάζουν οι παρακάτω συμβολοσειρές:

$$0000 \Rightarrow \text{eval}(0000) = 0$$

$$0001 \Rightarrow \text{eval}(0001) = 1$$

$$0010 \Rightarrow \text{eval}(0010) = 2$$

$$0011 \Rightarrow \text{eval}(0011) = 3$$

$$0100 \Rightarrow \text{eval}(0100) = 4$$

$$0101 \Rightarrow \text{eval}(0101) = 5$$

$$0110 \Rightarrow \text{eval}(0110) = 6$$

$$0111 \Rightarrow \text{eval}(0111) = 7$$

Επομένως η μέση απόδοση του σχήματος  $S_2$  ισούται με:

$$\bar{F}(S_2) = (0 + 1 + 2 + 3 + 4 + 5 + 6 + 7) / 8 = 3.5$$



# A. Θεωρία

## 3. Επικράτηση – Εξαφάνιση Σχημάτων

### 2. Παραδείγματα

#### Λύση Ερώτημα 2:

Η συνολική απόδοση του πληθυσμού είναι:

$F=0+1+2+\dots+15=120$  και η μέση απόδοση είναι 7,5

Άρα, η μέση απόδοση του  $S_1$  είναι 53.33% μεγαλύτερη της μέσης απόδοσης του πληθυσμού, άρα  $\varepsilon=0.53$  και του  $S_2$ , είναι 53.33% μικρότερη, άρα  $\varepsilon=-0.53$ .

Χρησιμοποιούμε την εξίσωση:  $\xi(S, t+k) = \xi(S, t) \cdot (1+\varepsilon)^k$

Για  $t=0$  έχουμε  $\xi(S_1, 0)=8$ , άρα πρέπει να υπολογίσουμε το:  $\xi(S_1, t+k) = 8 \cdot (1+\varepsilon)^k$ .

A) Για  $\varepsilon=0.53$  και  $\xi(S, t+k)=16$  (γιατί το  $S_1$ , θα επικρατήσει σε όλο τον πληθυσμό) έχουμε:

$$16 = 8 \cdot (1+0.53)^k = 8 \cdot 1.53^k \rightarrow 2 = 1.53^k$$

Αν λογαριθμίσουμε και για τα δύο μέρη έχουμε:

$$\log 2 = k \cdot \log (1.53) \rightarrow$$

$k = (0.301/0.185) \rightarrow k= 1.63$ . Συνεπώς, μετά από 2 γενιές αναμένουμε ότι το σχήμα  $S_1$  θα επικρατήσει στον πληθυσμό.

B) Για  $\varepsilon= -0.53$  και  $\xi(S_1, t+k) < 1$  έχουμε:

$$1 > 8 \cdot (1-0.53)^k = 8 \cdot (0.47)^k$$

Αν λογαριθμίσουμε και για τα δύο μέρη έχουμε:

$$0 > \log 8 + k \cdot \log (0.47) \rightarrow 0.903 + k (-0.327) < 0 \rightarrow (-0.327) k < -0.903 \rightarrow k > 2.73.$$

Άρα, μετά από 3 γενιές αναμένουμε ότι το σχήμα  $S_2$  θα εξαφανιστεί από τον πληθυσμό.



# A. Θεωρία

## 3. Επικράτηση – Εξαφάνιση Σχημάτων

### 2. Παραδείγματα

**Παράδειγμα 2:** Έστω το σχήμα  $S$  που αντιπροσωπεύεται, κατά τη γενιά 0, σε 60% ενός πληθυσμού μεγέθους  $n=100$  ατόμων και με αμελητέες πιθανότητες καταστροφής από διασταύρωση και μετάλλαξη. Να υπολογίσετε μετά από πόσες γενιές το σχήμα  $S$  θα ...  
(1) ... επικρατήσει σε ολόκληρο τον πληθυσμό αν τα μέλη του  $S$  έχουν καταλληλότητα 25% μεγαλύτερη από τη μέση καταλληλότητα του τρέχοντα πληθυσμού

**Λύση:** Χρησιμοποιούμε την εξίσωση:

$$\xi(S, t+k) = \xi(S, t) \cdot (1+\varepsilon)^k$$

Για  $t=0$  έχουμε  $\xi(S,0)=60$ , άρα πρέπει να υπολογίσουμε το:  $\xi(S, t+k) = 60 \cdot (1+\varepsilon)^k$ .

Για  $\varepsilon=0.25$  και  $\xi(S, t+k)=100$  έχουμε:

$$100 = 60 \cdot (1+0.25)^k = 60 \cdot 1.25^k$$

Αν πάρουμε το Log και για τα δύο μέρη έχουμε:

$$\text{Log } 100 = \text{Log } 60 + k \cdot \text{Log } (1.25) \rightarrow k=2.28, \text{ που σημαίνει } k=3.$$

Άρα, μετά από 3 γενιές το σχήμα  $S$  θα επικρατήσει στον πληθυσμό.



# A. Θεωρία

## 3. Επικράτηση – Εξαφάνιση Σχημάτων

### 2. Παραδείγματα

**Παράδειγμα 3:** Έστω το σχήμα  $S$  που αντιπροσωπεύεται, κατά τη γενιά 0, σε 60% ενός πληθυσμού μεγέθους  $n=100$  ατόμων και με αμελητέες πιθανότητες καταστροφής από διασταύρωση και μετάλλαξη. Να υπολογίσετε μετά από πόσες γενιές το σχήμα  $S$  θα ... (2)... εξαφανιστεί από τον πληθυσμό αν τα μέλη του  $S$  έχουν καταλληλότητα 10% μικρότερη από τη μέση καταλληλότητα του τρέχοντα πληθυσμού

**Λύση:** Χρησιμοποιούμε την εξίσωση:

Για  $\varepsilon = -0.10$  και  $\xi(S, t+k) < 1$  έχουμε:

$$1 > 60 \cdot (1-0.10)^k = 60 \cdot 0.9^k$$

Αν πάρουμε το Log και για τα δύο μέρη έχουμε:

$$1 > \text{Log } 60 + k \cdot \text{Log } (0.9) \rightarrow k=38.8, \text{ που σημαίνει } k=39.$$

Άρα, μετά από 39 γενιές το σχήμα  $S$  θα εξαφανιστεί από τον πληθυσμό.



## B. Ασκήσεις

### Εφαρμογή 1

Θεωρήστε τρεις δυαδικές συμβολοσειρές  $A_1=11101111$ ,  $A_2=00010100$ ,  $A_3=01000011$  και έξι σχήματα:

$H1=1*****$

$H2=0*****$

$H3=*****11$

$H4=***0*00*$

$H5=1*****1*$

$H6=1110**1*$

1. Ποιες συμβολοσειρές ταιριάζουν σε κάθε σχήμα;
2. Βρείτε την τάξη και το οριστικό μήκος κάθε σχήματος
3. Υπολογίστε την πιθανότητα επιβίωσης κάθε σχήματος, αν γίνεται διασταύρωση με πιθανότητα διασταύρωσης  $p_c=0.85$
4. Υπολογίστε την πιθανότητα επιβίωσης κάθε σχήματος, αν γίνεται μετάλλαξη με πιθανότητα μετάλλαξης  $p_m=0.001$



## Β. Ασκήσεις

### Εφαρμογή 2

Έστω ότι ο πληθυσμός που δίνεται στον παρακάτω πίνακα είναι ο αρχικός πληθυσμός ενός Γενετικού Αλγορίθμου που χρησιμοποιείται για τη μεγιστοποίηση της συνάρτησης:  $f(x_1, x_2) = x_2^2 - x_1^2$  όπου τα  $x_1$  και  $x_2$  είναι ακέραιοι αριθμοί στο διάστημα  $[0, 7]$ . Για την αναπαράσταση κάθε ανεξάρτητης μεταβλητής χρησιμοποιούνται 3 bits ( $2^3 = 8$ ), οπότε για την αναπαράσταση κάθε χρωμοσώματος του πληθυσμού χρησιμοποιούνται 6 bits με τα 3 αριστερά να αναπαριστούν την μεταβλητή  $x_1$  και τα 3 δεξιά τη μεταβλητή  $x_2$ . Απαντήστε στα ζητήματα που ακολουθούν και συμπληρώστε τον πίνακα (κάντε τις πράξεις με ακρίβεια 4 δεκαδικών ψηφίων).

	ΑΤΟΜΟ - ΧΡΩΜΟΣΩΜΑ	ΑΠΟΔΟΣΗ - ΙΚΑΝΟΤΗΤΑ (FITNESS)	ΠΙΘΑΝΟΤΗΤΑ ΕΠΙΛΟΓΗΣ	ΑΝΑΜΕΝΟΜΕΝ ΟΣ ΑΡΙΘΜΟΣ ΑΝΤΙΓΡΑΦΩΝ	ΠΛΗΘΥΣΜΟΣ ΜΕΤΑ ΤΗΝ ΕΦΑΡΜΟΓΗ ΤΩΝ ΓΕΝΕΤΙΚΩΝ
A	101011				
B	111000				
Γ	111011				
Δ	101000				
E	111001				

**(α) (5/15)**

**(α1) (3/5)** Να υπολογίσετε την απόδοση (Fitness) κάθε μέλους του πληθυσμού (Να γράψετε την απάντησή σας στον παραπάνω πίνακα). Επειδή η καταλληλότητα (ικανότητα) περιέχει αρνητικές τιμές, δεν μπορούν να υπολογιστούν οι πιθανότητες επιλογής.

Πώς μπορεί να αντιμετωπιστεί αυτό το πρόβλημα; Να αιτιολογήσετε την απάντησή σας και να συμπληρώσετε την κενή στήλη του παρακάτω πίνακα, με τη νέα ικανότητα. Ποιο πρόβλημα υπάρχει με τον υπολογισμό της πιθανότητας επιλογής του 2<sup>ου</sup> ατόμου; Πως μπορεί να αντιμετωπιστεί;

**(α2) (2/5)** Να υπολογίσετε την πιθανότητα επιλογής κάθε ατόμου, χρησιμοποιώντας επιλογή εξαναγκασμένης ρουλέτας (Να γράψετε την απάντησή σας στον παραπάνω πίνακα).

**(β) (3/15)** Να υπολογίσετε τον αναμενόμενο αριθμό αντιγράφων κάθε ατόμου στην επόμενη γενιά (Να γράψετε την απάντησή σας στον παραπάνω πίνακα).



**(γ) (7/15)** Έστω ότι η 1<sup>η</sup> γενιά που προκύπτει μετά την εφαρμογή των γενετικών τελεστών της επιλογής, διασταύρωσης και μετάλλαξης είναι αυτή που απεικονίζεται στον πίνακα.

	<b>ΑΤΟΜΟ – ΧΡΩΜΟΣΩΜΑ</b>
A	111   100
B	110   110
Γ	111   000
Δ	110   011
E	111   011

**(γ1) (2/7)** Έχει βελτιωθεί η μέση απόδοση του πληθυσμού στη γενιά 1 σε σχέση με τη γενιά 0; Πόσο έχει μεταβληθεί η απόδοση του καλύτερου μέλους του πληθυσμού στη γενιά 1 σε σχέση με τη γενιά 0;





**(γ2) (5/7)** Ποιος θα είναι ο αναμενόμενος αριθμός ατόμων, σύμφωνα με το θεώρημα σχημάτων (Schema Theorem), που θα αντιστοιχούν στο σχήμα  $S_1=111^{***}$  στη γενιά 2 αν θεωρήσουμε ότι εφαρμόζεται διασταύρωση με πιθανότητα  $P_c=0.5$  και μετάλλαξη με πιθανότητα  $P_m=0.1$ ;