

Θέμα 1: Ένας Απλός Γενετικός Αλγόριθμος

Θεωρούμε τη γενική γενετική δεξαμενή τεσσάρων βασιλόσαυρων που ανήκουν σε αυτόν το κόσμο. Θα θεωρήσουμε τα χρωμοσώματα που κωδικοποιούν τα χαρακτηριστικά των ποδιών των προγόνων τους. Το μήκος του πέλματος, το μήκος των δακτύλων και το μήκος των νυχιών κωδικοποιείται από έξι γονίδια: τα δύο πρώτα κωδικοποιούν το πέλμα, τα άλλα δύο τα δάκτυλα και τα δύο τελευταία τα νύχια. Στην αναπαράσταση του γονότυπου, το O δείχνει την ενεργοποίηση ενός χαρακτηριστικού (γονιδίου) και το X την απενεργοποίηση ενός γονιδίου. Ο ιδεατός γονότυπος (κοντά πέλματα, μακριά δάκτυλα και μακριά νύχια) είναι: **XXOOOO**.

Στον παρακάτω πίνακα δίνεται η γενετική δεξαμενή (αρχικός πληθυσμός). Μπορούμε να παρατηρήσουμε ότι μερικά άτομα (A, B) είναι αρκετά κοντά στα χαρακτηριστικά των προγόνων τους και το D είναι το πλησιέστερο. Στον κόσμο των βασιλόσαυρων η ικανότητα της κίνησης είναι το κύριο κριτήριο επιβίωσης και αναπαραγωγής. Έτσι, κανένα θηλυκό δεν θα επιθυμούσε να παντρευτεί ένα βασιλόσαυρο του οποίου τα πέλματα θα ήταν σαν του A, αλλά όλα θα ονειρευόντουσαν να συναντήσουν το D κάποια μέρα. Θέλουμε να δούμε, αν στην επόμενη γενιά ένας θηλυκός βασιλόσαυρος θα συναντήσει ένα αρσενικό με τα χαρακτηριστικά του D.

1. Να σχεδιάσετε ένα κατάλληλο Γενετικό Αλγόριθμο για το παραπάνω πρόβλημα. Συγκεκριμένα:

- Να προτείνετε την κατάλληλη συνάρτηση αξιολόγησης.
- Να προτείνετε την κατάλληλη κωδικοποίηση και να τεκμηριώσετε την επιλογή σας.
- Να ορίσετε όσες άλλες παραμέτρους χρειάζονται.

Απάντηση:

- Αφού το πλέον κατάλληλο (ιδεατό) άτομο είναι το XXOOOO, για κάθε άτομο μετράμε σε πόσες θέσεις ταιριάζει με το ιδεατό. Ο αριθμός αυτών των θέσεων είναι η καταλληλότητά του, π.χ. το άτομο B= XOXOXO έχει Eval(B)=2.
- Αφού κάθε γονίδιο παίρνει δύο τιμές, X ή O (δυαδική) η κωδικοποίηση του πίνακα είναι η πλέον κατάλληλη, γιατί έχει τον ελάχιστο αριθμό συμβόλων.

Άτομο (XXOOOO - ιδεατό)	Καταλληλότητα	Ζεύγη για Ζευγάριωμα	Νέος Πληθυσμός
A = OXXXXX			
B = XOXOXO			
C = XOXOXO			
D = XXOXOO			
E = OXOXOX			
F = OOXOXX			
Συνολική Καταλληλότητα:			

- Επειδή έχουμε ένα απλό Γ.Α. προτείνεται επιλογή εξαναγκασμένης ρουλέτας, διασταύρωση μονού σημείου και απλή μετάλλαξη ($X \rightarrow O$ και $O \rightarrow X$).
- Οι άλλες παράμετροι είναι το μέγεθος του πληθυσμού που δίνεται pop_size=6, και οι πιθανότητες διασταύρωσης P_C και μετάλλαξης P_m .

2. Να εξελίξετε τον παραπάνω πληθυσμό, κατά μία γενιά και να συμπληρώσετε τον παρακάτω πίνακα, με τις εξής προϋποθέσεις.

- Το μέγεθος του πληθυσμού παραμένει σταθερό.
- Η διασταύρωση γίνεται με πιθανότητα $P_C=0.80$.
- Η μετάλλαξη γίνεται με πιθανότητα $P_m=0.30$.
- Τα άτομα ζευγαρώνουν με τη σειρά που επιλέγονται από τη ρουλέτα (δεν χρειάζεται άλλος έλεγχος).

Αν χρειαστείτε τυχαίους αριθμούς, χρησιμοποιείτε τους παρακάτω, με τη σειρά που δίνονται:

0.12, 0.86, 0.21, 0.67, 0.64, 0.34, 0.30, 0.85, 0.56, 0.63, 0.47, 0.19, 0.80, 0.98, 0.58, 0.03, 0.57, 0.49, 0.92 .

Επαναλάβετε από την αρχή, αν χρειαστείτε περισσότερους. Για τους υπολογισμούς χρησιμοποιείτε ακρίβεια δύο δεκαδικών ψηφίων.

Απάντηση:

Άτομο	Γονότυπος	Καταλληλότητα (Συνολική καταλ.=15)		
A	OXXXXX	1	$p_1=1/15$	$q_1=1/15=0.07$
B	XOXXOX	2	$p_2=2/15$	$q_2=3/15=0.2$
C	XOXOXO	3	$p_3=3/15$	$q_3=6/15=0.4$
D	XXOXOO	5	$p_4=5/15$	$q_4=11/15=0.73$
E	OXOXOX	3	$p_5=3/15$	$q_5=14/15=0.93$
F	OOXOXX	1	$p_6=1/15$	$q_6=15/15=1$

Η μέση καταλληλότητα του παραπάνω πληθυσμού είναι 2.5

Χρησιμοποιώ τους τυχαίους αριθμούς για να τους θεωρήσω ως το αποτέλεσμα της ρουλέτας.

Τυχαίοι αριθμοί		Άτομο που επιλέγω
0.12	$0.07 < 0.12 < 0.2$	B
0.86	$0.73 < 0.86 < 0.9$	E
0.21	$0.2 < 0.21 < 0.4$	C
0.67	$0.4 < 0.67 < 0.73$	D
0.64	$0.4 < 0.64 < 0.73$	D
0.34	$0.2 < 0.34 < 0.4$	C

Χρησιμοποιώ και πάλι τους τυχαίους αριθμούς και αν αυτοί είναι μικρότεροι του 0.8 (που είναι η πιθανότητα διασταύρωσης) διασταυρώνω το άτομο, αλλιώς όχι.

Για το B έχω την πιθανότητα 0.3, άρα το διασταυρώνω.

Για το E	-//-	0.85, άρα ΔΕΝ το διασταυρώνω.
Για το C	-//-	0.56, άρα το διασταυρώνω.
Για το D	-//-	0.63, άρα το διασταυρώνω.
Για το D	-//-	0.47, άρα το διασταυρώνω.
Για το C	-//-	0.19, άρα το διασταυρώνω.

Παρατηρώ ότι κατ' αυτό τον τρόπο μου προέκυψε πληθυσμός με 5 άτομα, άρα κάποιο θα πρέπει να το απορρίψω. Χωρίζω το διάστημα (0,1] σε 5 ίσα μέρη, όσα και τα άτομα που έχω τώρα για διασταύρωση. Παίρνω τον επόμενο τυχαίο αριθμό, που είναι ο 0.80, ο οποίος ανήκει στο προτελευταίο διάστημα (0.6, 0.8]. Άρα θα αποκλείσω το δεύτερο εμφανιζόμενο D. Επειδή τα άτομα επιλέχθηκαν με τυχαία σειρά, μπορώ να δημιουργήσω ζεύγη με τη σειρά που εμφανίζονται, δηλαδή (B,C) και (D,C).

Παρατηρώ επίσης ότι διαχωρισμός στο γονότυπο μπορεί να συμβεί σε 5 θέσεις. Άρα και πάλι χωρίζω το (0,1] σε 5 τμήματα, όπου κάθε τμήμα αντιστοιχεί σε μία θέση. Ο επόμενος τυχαίος αριθμός είναι το 0.98, ανήκει στο 5^ο τμήμα άρα το σημείο διαχωρισμού θα είναι μετά το 5^ο γονίδιο του B και του C.

Άρα το B: XOXXOX γίνεται XOXXOO
και το C: XOXXOXO γίνεται XOXXOX

Ο επόμενος τυχαίος αριθμός είναι 0.58, ο οποίος αντιστοιχεί στο 3^ο τμήμα, άρα ο διαχωρισμός θα γίνει στην 3^η θέση για το δεύτερο ζευγάρι, δηλαδή το D:XXOXOO γίνεται XXOOXO
και το C: XOXXOXO γίνεται XOXXOO

Στη συνέχεια, τα άτομα που προέκυψαν από τη διασταύρωση θα αντικαταστήσουν τα άτομα από τα οποία προήλθαν, οπότε η τρέχουσα μορφή του πληθυσμού έχει ως εξής: B', E, C₁', D', D και C₂' και επιτρέπεται να μεταλλαχθούν με πιθανότητα $p_m=0.3$. Χρησιμοποιούμε για κάθε γονίδιο κάθε ατόμου έναν τυχαίο αριθμό και αν αυτός είναι μικρότερος από 0.3 το γονίδιο μεταλλάσσεται, διαφορετικά όχι. Άρα έχουμε:

B': X : 0.03 (<0.3) → O
O : 0.57 → O
X : 0.49 → X
X : 0.92 → X
O : 0.12 (<0.3) → X
O : 0.86 → O

E: O : 0.21 (<0.3) → X
X : 0.67 → X
O : 0.64 → O
X : 0.34 → X
O : 0.30 → O
X : 0.85 → X

C₁': X : 0.56 → X
O : 0.63 → O
X : 0.47 → X

O : 0.19 (<0.3) → X
 X : 0.8 → X
 X : 0.98 → X

D': X : 0.58 → X
 X : 0.03 (<0.3) → O
 O : 0.57 → O
 O : 0.49 → O
 X : 0.92 → X
 O : 0.12 (<0.3) → X

D: X: 0.86 → X
 X: 0.21 (<0.3) → O
 O: 0.67 → O
 X: 0.64 → X
 O: 0.34 → O
 O: 0.30 → O

C₂': X: 0.85 → X
 O: 0.56 → O
 X: 0.63 → X
 X: 0.47 → X
 O: 0.19 (<0.3) → X
 O: 0.80 → O

Συνεπώς ο νέος πληθυσμός (με μέση καταλληλότητα = 2.5) δεν είναι καλύτερος από τον προηγούμενο και αποτελείται από τα εξής άτομα:

B': OOXXXO με καταλληλότητα 1
 E : XXOXOX με καταλληλότητα 4
 C₁': XOXXXX με καταλληλότητα 1
 D': XOOOXX με καταλληλότητα 3
 D: XOOXOO με καταλληλότητα 4
 C₂': XOXXCO με καταλληλότητα 2

3. Στη νέα γενιά του πληθυσμού, ποιο είναι το καταλληλότερο άτομο; Είναι το ιδεατό; Αν όχι, σε πόσα βήματα μπορείτε να βρείτε το ιδεατό άτομο, αν επιτρέπεται να χρησιμοποιηθούν οι γενετικοί τελεστές ντετερμινιστικά (δηλαδή με πιθανότητα =1).

Απάντηση:

Στη νέα γενιά τα καταλληλότερα άτομα είναι δύο, το E και το D, τα οποία δεν είναι ιδεατά και θα εφαρμόσουμε του γενετικούς τελεστές ντετερμινιστικά στον πληθυσμό που προέκυψε από τη μετάλλαξη.

Άτομο	Γονότυπος	Καταλληλότητα (Συνολική καταλ.=15)		
B	OOXXCO	1	p1=1/15	q1=1/15=0.07

E	XXOXOX	4	$P_2=4/15$	$q_2=5/15=0.33$
C ₁	XOXXXX	1	$P_3=1/15$	$q_3=6/15=0.40$
D ₁	XOOOXX	3	$p_4=3/15$	$q_4=9/15=0.60$
D ₂	XOOXOO	4	$p_5=4/15$	$q_5=13/15=0.86$
C ₂	XOXXOX	2	$p_6=2/15$	$q_6=15/15=1$

Η μέση καταλληλότητα του παραπάνω πληθυσμού είναι 2.5.

Χρησιμοποιώ τους τυχαίους αριθμούς για να τους θεωρήσω ως το αποτέλεσμα της ρουλέτας.

Τυχαίοι αριθμοί		Άτομο που επιλέγω
0.98	$0.86 < 0.98 < 1$	C ₂
0.58	$0.40 < 0.58 < 0.60$	D ₁
0.03	$0 < 0.03 < 0.07$	B
0.57	$0.40 < 0.57 < 0.60$	D ₁
0.49	$0.40 < 0.49 < 0.60$	D ₁
0.92	$0.86 < 0.92 < 1$	C ₂

Ο νέος πληθυσμός αποτελείται από τα παραπάνω νέα άτομα. Στη συνέχεια εφαρμόζουμε τον τελεστή διασταύρωσης με πιθανότητα $p_c=1$, γεγονός που σημαίνει ότι θα επιλεγούν όλα τα άτομα για διασταύρωση. Τα ζευγάρια που προκύπτουν είναι τα C₂-D₁, B-D₁ και D₁-C₂.

Εφόσον η διασταύρωση μπορεί να συμβεί σε 5 θέσεις του γονότυπου, χωρίζω το (0,1] σε 5 τμήματα, όπου κάθε τμήμα αντιστοιχεί σε μία θέση. Ο πρώτος τυχαίος αριθμός για το πρώτο ζευγάρι είναι 0.30, οπότε το σημείο διασταύρωσης είναι η δεύτερη θέση ($pos=2$).

Οπότε το C₂ από XOXXOX γίνεται XOOOXX

D₁ από XOOOXX γίνεται XOXXOX.

Για το δεύτερο ζευγάρι ο δεύτερος τυχαίος αριθμός είναι το 0.85, άρα το σημείο διασταύρωσης είναι η πέμπτη θέση.

Το B από OXXXXO γίνεται OXXXXX

και το D₁ από XOOOXX γίνεται XOOOXO.

Για το τρίτο ζευγάρι ο τρίτος τυχαίος αριθμός είναι το 0.56, άρα το σημείο διασταύρωσης είναι η τρίτη θέση.

Το D₁ από XOOOXX γίνεται XOOXOX

και το C₂ από XOXXOX γίνεται XOXOXX.

Έπειτα εφαρμόζουμε τον τελεστή μετάλλαξης με πιθανότητα $p_m=1$. Επειδή όλοι οι τυχαίοι αριθμοί είναι μικρότεροι του 1, αντιστρέφονται όλα τα γονίδια. Συνεπώς ο τελικός πληθυσμός της δεύτερης γενιάς είναι ο εξής:

C₂: OXXXOO με καταλληλότητα 3
D₁: OXOOOX με καταλληλότητα 4
B: XXOOOO με καταλληλότητα 6
D₁: OXXXOX με καταλληλότητα 2
D₁: OXXOOX με καταλληλότητα 3
C₂: OXOXOO με καταλληλότητα 4

Παρατηρούμε ότι μας προέκυψε ιδεατό άτομο και είναι το B.

4. Αν αντί για επιλογή εξαναγκασμένης ρουλέτας χρησιμοποιήσετε ελιτισμό για να εξελίξετε τον αρχικό πληθυσμό, η επόμενη γενιά είναι καλύτερη ή χειρότερη από αυτήν που έχει προκύψει από το ερώτημα 1.

Απάντηση:

Εδώ χρησιμοποιούμε τον ελιτισμό, οπότε επιτρέπουμε στο πιο κατάλληλο άτομο να περάσει απευθείας στο νέο πληθυσμό. Από όλα τα άτομα το D είναι το καταλληλότερο, άρα θα γυρίσουμε τη ρουλέτα 5 φορές αντί για 6. Χρησιμοποιώντας τους τυχαίους αριθμούς από την αρχή, επιλέγουμε τα εξής άτομα:

	D	(λόγω ελιτισμού)
$r = 0.12$ οπότε επιλέγω το B		
$r = 0.86$ οπότε επιλέγω το E		
$r = 0.21$ οπότε επιλέγω το C		
$r = 0.67$ οπότε επιλέγω το D		
$r = 0.64$ οπότε επιλέγω το D		

Χρησιμοποιώ και πάλι τους τυχαίους αριθμούς και αν αυτοί είναι μικρότεροι του 0.8 (που είναι η πιθανότητα διασταύρωσης) διασταυρώνω το άτομο, αλλιώς όχι.

Για το B έχω την πιθανότητα 0.34, άρα το διασταυρώνω.

Για το E -//- 0.30, άρα το διασταυρώνω.

Για το C -//- 0.85, άρα ΔΕΝ το διασταυρώνω.

Για το D -//- 0.56, άρα το διασταυρώνω.

Για το D -//- 0.63, άρα το διασταυρώνω.

Παρατηρώ ότι κατ' αυτό τον τρόπο μου προέκυψε πληθυσμός με 4 άτομα, άρα με αυτά δημιουργώ τα ζεύγη μου, που θα είναι τα (B, E) και (D, D).

Για να βρω το σημείο διασταύρωσης και πάλι χωρίζω το (0,1] σε 5 ίσα τμήματα και χρησιμοποιώ τους τυχαίους αριθμούς. Ο πρώτος τυχαίος είναι το 0.47, ανήκει στο 3^ο διάστημα οπότε το πρώτο ζευγάρι θα διασταυρωθεί μετά το 3^ο γονίδιο και θα προκύψουν τα άτομα B': ΧΟΧΧΟΧ Ε': ΟΧΟΧΟΧ (και έτυχε να παραμείνουν ως είχαν). Ο δεύτερος τυχαίος αριθμός είναι το 0.19, το οποίο ανήκει στο 1^ο διάστημα οπότε το δεύτερο ζευγάρι θα διασταυρωθεί μετά το 1^ο γονίδιο και θα προκύψουν τα εξής άτομα: D¹': ΧΧΟΧΟΟ και D²': ΧΧΟΧΟΟ αλλά επειδή τα άτομα ήταν ίδια, παρέμειναν κι αυτά ως είχαν!

Στο επόμενο βήμα επιλέγουμε τα γονίδια που θα μεταλλαχθούν.

B' :	X : 0.8	→ X
	O : 0.98	→ O
	X : 0.58	→ X
	X : 0.03	→ O
	O : 0.57	→ O
	X : 0.49	→ X

E': O : 0.92 → O
 X : 0.12 → O
 O : 0.86 → O
 X : 0.21 → O
 O : 0.67 → O
 X : 0.64 → X

C: X : 0.34 → X
 O : 0.30 → O
 X : 0.85 → X
 O : 0.56 → O
 X : 0.63 → X
 O : 0.49 → O

D₁': X : 0.19 → O
 X : 0.80 → X
 O : 0.98 → O
 X : 0.58 → X
 O : 0.03 → X
 O : 0.57 → O

D₂': X: 0.49 → X
 X: 0.92 → X
 O: 0.12 → X
 X: 0.86 → X
 O: 0.21 → X
 O: 0.67 → O

Άρα ο νέος πληθυσμός με τη μέθοδο του ελιτισμού είναι ο:

D : XXOXOO με καταλληλότητα 5
 B' : XOXOOX με καταλληλότητα 3
 E' : OOOOOX με καταλληλότητα 3
 C : XOXOXO με καταλληλότητα 3
 D₁' : OXOXOX με καταλληλότητα 3
 D₂' : XXXXOX με καταλληλότητα 3

και έχει μέση καταλληλότητα = 3.33, συνεπώς είναι καλύτερος ο πληθυσμός από αυτόν που είχε προκύψει στο ερώτημα 1.