

Γενετικοί αλγόριθμοι - ΓΑ

Genetic algorithms - GA

ΕΦΑΡΜΟΓΗ

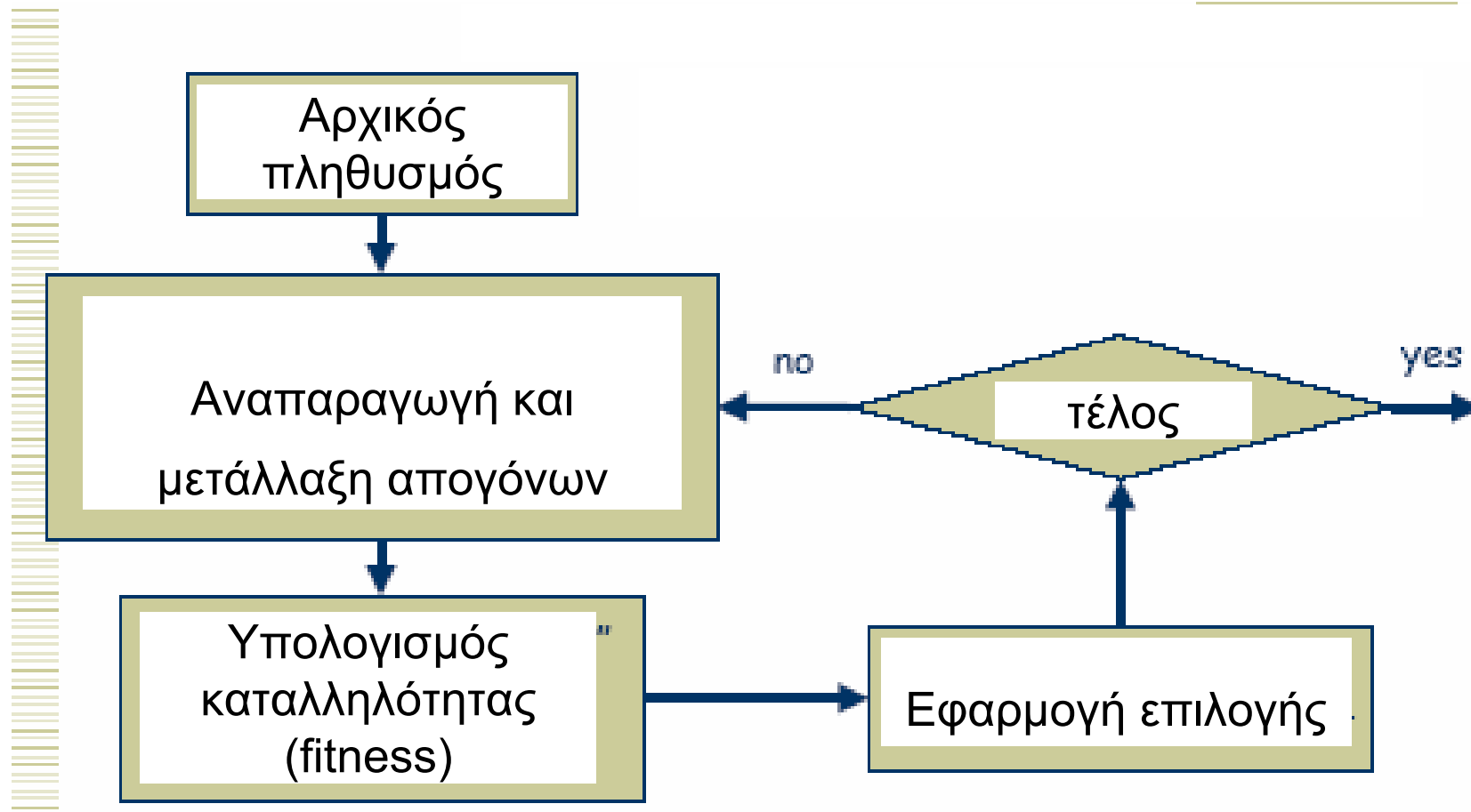
στην ΕΠΕΞΕΡΓΑΣΙΑ ΣΗΜΑΤΟΣ και στην ΑΣΑΦΗ ΛΟΓΙΚΗ

ΓΑ - Εισαγωγικά

- Γενετικοί αλγόριθμοι (Genetic algorithms)
- Εξελικτικοί αλγόριθμοί (Evolutionary algorithms)
- Εξελικτικά προγράμματα (Evolutionary programs)
- Περιοχή της Τεχνητής Νοημοσύνης
- **Μιμείται** τη βιολογική διεργασία της εξέλιξης:
«Τα είδη του ζωικού και φυτικού βασιλείου μεταλλάσσονται και στις επόμενες γενεές επικρατούν τα ισχυρότερα»

The survival of the fittest

ΓΑ - διάγραμμα



ΓΑ και Τεχνικές "βελτιστοποίησης" (Optimization techniques)

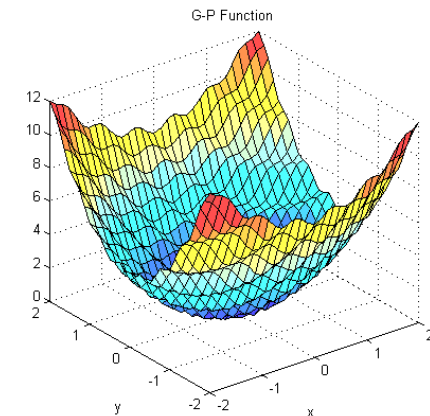
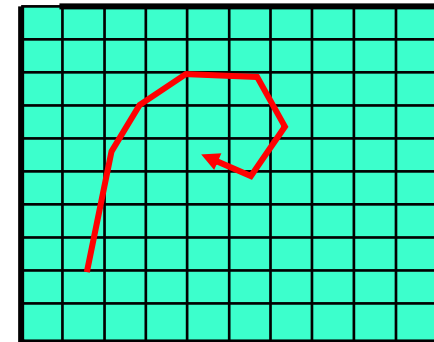
Εξαντλητική διαδικασία
(Exhaustive search)

Gradient methods - steepest descent

Neural networks

SA –simulated annealing

.....



ΓΕΝΕΤΙΚΗ

1. Αρχή της **ΠΡΟΣΑΡΜΟΓΗΣ**
2. Υλοποιείται με τρεις πράξεις:
 - **ΑΝΑΠΑΡΑΓΩΓΗ**
 - **ΔΙΑΣΤΑΥΡΩΣΗ**
 - **ΜΕΤΑΛΛΑΞΗ**

"ΜΕΤΑΦΟΡΑ" ΣΕ ΑΛΓΟΡΙΘΜΟ

- Η αρχή της προσαρμογής υλοποιείται με μία σχετική συνάρτηση προσαρμοστικότητας – καταλληλότητας (fitness function)
- Όλα τα μέλη του πληθυσμού αξιολογούνται με την **τιμή καταλληλότητας** που λαμβάνουν (fitness value)

(συνέχεια)

- ΓΕΝΕΤΙΚΟΣ ΚΩΔΙΚΑΣ → ΔΥΑΔΙΚΟΣ ΚΩΔΙΚΑΣ
- ΓΟΝΙΔΙΟ → bit
- ΧΡΩΜΟΣΩΜΑ → δυαδικός αριθμός B_i , πχ. 10011000 (=152)
- Γίνεται **απεικόνιση** του αριθμού B_i στον C_i σύμφωνα με τον αριθμό L των **διαθεσίμων bits** (γραμμική παρεμβολή):

$$C_i = C_{\min} + \frac{B_i}{2^L - 1} (C_{\max} - C_{\min})$$

Όλες οι παράμετροι θεωρούνται δυαδικοί αριθμοί και τοποθετούνται σε μία σειρά σχηματίζοντας ένα τελικό **αριθμό - «χρωμόσωμα»** για επεξεργασία

παράδειγμα

$$C_i = C_{\min} + \frac{B_i}{2^L - 1} (C_{\max} - C_{\min})$$

Σε ένα πρόβλημα επιζητείται η εύρεση δύο παραμέτρων.

Για κάθε ένα θεωρούμε 6 bits δηλαδή το χρωμόσωμα θα έχει μήκος =12.

Για την απεικόνιση των παραμέτρων δίνονται: $C_{\max}=5$, $C_{\min}=-2$, $L=6$ και $B_i=7,18,21,36$

Οι τιμές C_i υπολογίζονται:

$$C_1 = C_{\min} + \frac{B_1}{2^L - 1} (C_{\max} - C_{\min}) = -2 + \frac{7}{2^6 - 1} (5 + 2) = -1.22$$

$$C_2 = C_{\min} + \frac{B_2}{2^L - 1} (C_{\max} - C_{\min}) = -2 + \frac{18}{2^6 - 1} (5 + 2) = 0.0$$

$$C_3 = C_{\min} + \frac{B_3}{2^L - 1} (C_{\max} - C_{\min}) = -2 + \frac{21}{2^6 - 1} (5 + 2) = 0.33$$

$$C_4 = C_{\min} + \frac{B_4}{2^L - 1} (C_{\max} - C_{\min}) = -2 + \frac{36}{2^6 - 1} (5 + 2) = 2.0$$

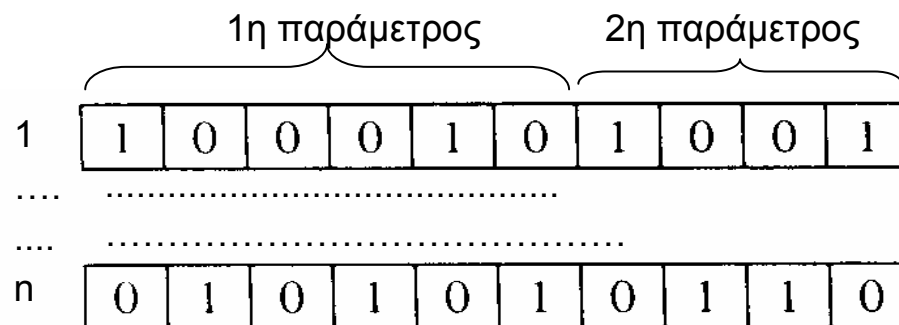
ΓΑ - Αρχικές τιμές

ΑΥΘΑΙΡΕΤΗ ΠΑΡΑΓΩΓΗ n ΑΡΧΙΚΩΝ ΤΙΜΩΝ

Δηλαδή παράγονται n «χρωμοσώματα» - δυαδικοί αριθμοί που περιλαμβάνουν τις ζητούμενες παραμέτρους (λύσεις).

Η διαδικασία είναι τυχαία (random numbers)

Στη συνέχεια γίνεται αποκωδικοποίηση των σε δεκαδικούς αριθμούς που αντιστοιχούν στις παραμέτρους



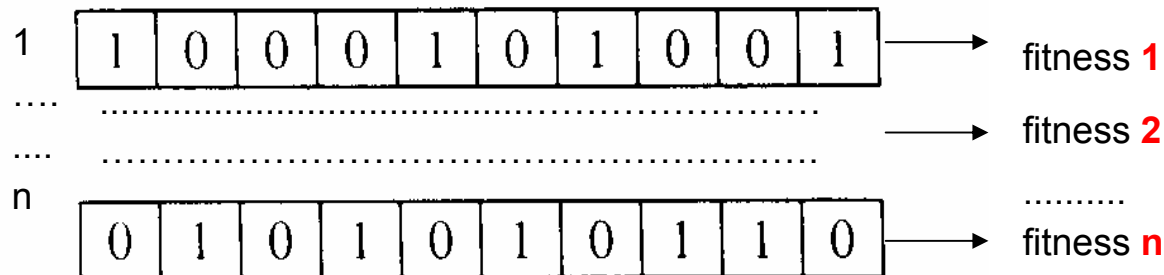
ΓΑ – Αναπαραγωγή –επιλογή γονέων

Με τις αποκωδικοποιημένες (δεκαδικές) τιμές των παραμέτρων υπολογίζεται η τιμή προσαρμοστικότητας- καταλληλότητας (**fitness value**) για κάθε ένα χρωμόσωμα

Οι λύσεις με υψηλή τιμή αναπαράγονται σε μεγαλύτερο ποσοστό στο νέο πληθυσμό.

Η ισοδύναμα επιλέγονται με μεγαλύτερη πιθανότητα

Ο συνολικός αριθμός των χρωμοσωμάτων παραμένει ίδιος

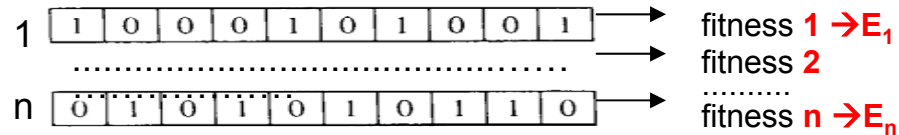


(συνέχεια)

Η συνάρτηση καταλληλότητας (**fitness**) χρησιμοποιείται για την επιλογή των γονέων (αναπαραγωγή) με τους εξής τρόπους:

1. Σαν σχετική συχνότητα (πιθανότητα) επιλογής (Αναλογική μέθοδος – proportionate). Μία υλοποίησή γίνεται με την μέθοδο του «τροχού της τύχης» - (**roulette wheel method**)
2. Για δημιουργία διάταξης και επιλογής βάσει του **βαθμού** κάθε γονέα
3. Με διαδικασία τυχαίας επιλογής από ένα σύνολο εκ των γονέων (Tournament selection)

1.α Αναπαραγωγή με σχετική συχνότητα



Στην απλούστερη μορφή η τιμή καταλληλότητας – fitness χρησιμοποιείται άμεσα για να δώσει τον αριθμό E_i των “αντιγράφων” ενός γονέα i σε ένα πληθυσμό n

$$E_i = \frac{\text{fitness}(n_i)}{\langle \text{fitness}(n_i) \rangle}$$

$\langle \text{fitness}(n_i) \rangle$ είναι η μέση τιμή της fitness στον πληθυσμό

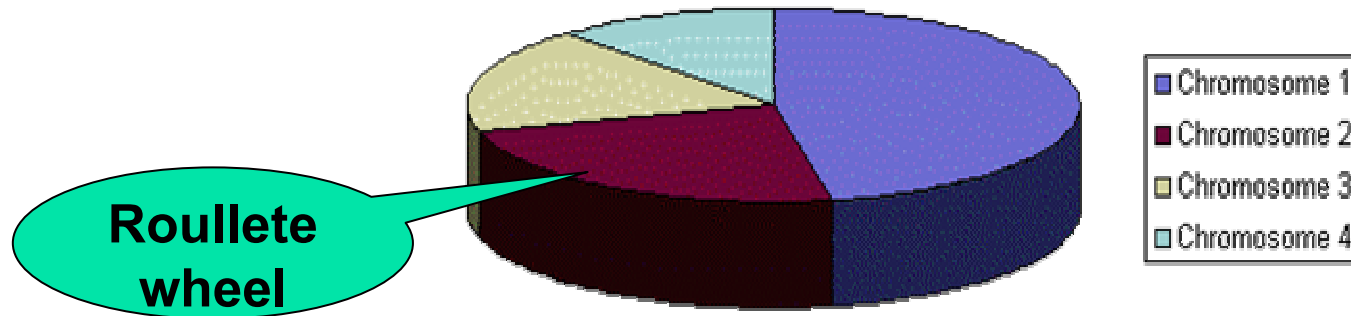
Στην διαδικασία αυτή συνήθως τίθεται και ένα κατώφλιο στην τιμή της E_i . Χρωμοσώματα κάτω της τιμής αυτής δεν επιλέγονται

Αναπαραγωγή με σχετική συχνότητα - παράδειγμα

- Το πρόβλημα: $\max x^2$ στο σύνολο $\{0,1,\dots,31\}$
- Λύση με ΓΑ. Χρωμόσωμα 5 ψηφίων πχ. 01101 \leftrightarrow 13
- Μέγεθος πληθυσμού : 4

String no.	Initial population	x Value	Fitness $f(x) = x^2$	$Prob_i$	Expected count	Actual count
1	0 1 1 0 1	13	169	0.14	0.58	1
2	1 1 0 0 0	24	576	0.49	1.97	2
3	0 1 0 0 0	8	64	0.06	0.22	0
4	1 0 0 1 1	19	361	0.31	1.23	1
Sum			1170	1.00	4.00	4
Average			293	0.25	1.00	1
Max			576	0.49	1.97	2

1.β Αναπαραγωγή με σχετική συχνότητα -*roulette wheel*



Η περιφέρεια χωρίζεται σε μήκη ανάλογα του $s_i = \frac{\text{fitness}(n_i)}{\sum_i \text{fitness}(n_i)}$

Παράγεται ένας τυχαίος αριθμός ομοιόμορφης κατανομής και επιλέγεται το αντίστοιχο χρωμόσωμα για αναπαραγωγή

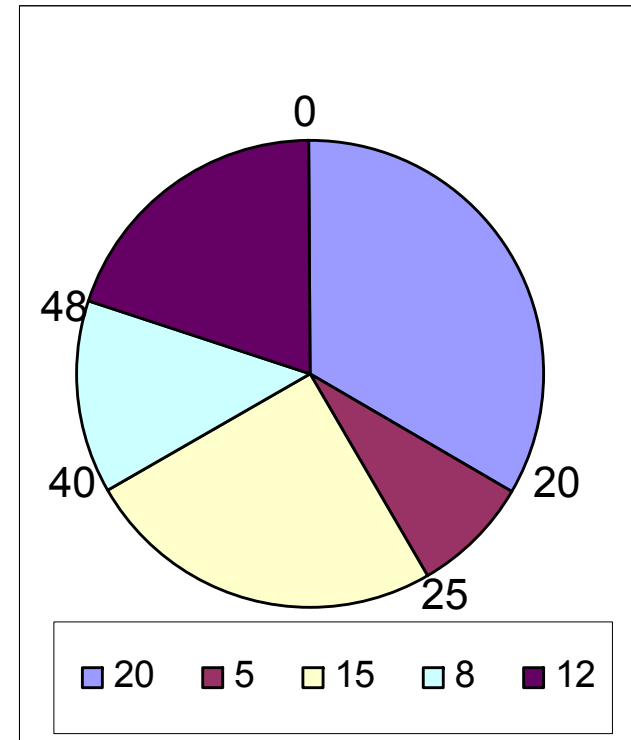
Roulette wheel

παράδειγμα

Οι αριθμοί: 0 -60

Chromosome	Fitness
10110110	20
10000000	5
11101110	15
10010011	8
10100010	12

Random Numbers	Father	Mother
44, 31	10010011	11101110
5, 32	10110110	11101110
49, 3	10100010	10110110
18, 27	10110110	11101110
22, 54	10000000	10100010



Αναπαραγωγή με τον βαθμό διάταξης (rank)

- Κάθε χρωμόσωμα i λαμβάνει ένα βαθμό r_i σύμφωνα με την fitness value
- Στη συνέχεια υπολογίζεται η καινούρια (υποκειμενική) fitness value sf_i :

$$sf_i = (P-r_i)(max-min)/(P-1) + min$$

P is the population size,

Max είναι η τιμή (καταλληλότητας) του καλύτερου χρωμοσώματος

Min είναι η τιμή (καταλληλότητας) του χειρότερου χρωμοσώματος

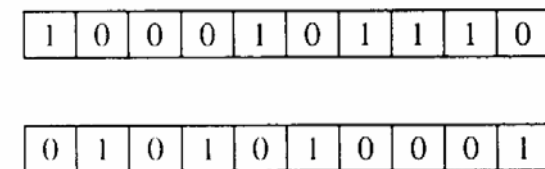
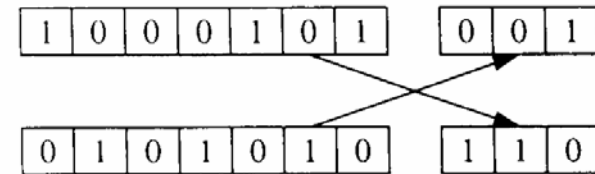
- Η τιμή της sf_i μπορεί να χρησιμοποιηθεί στη διαδικασία επιλογής roulette wheel: $s_i = \frac{sf_i}{\sum_i sf_i}$

ΓΑ - Διασταύρωση

- Μετά την επιλογή των γονέων γίνεται διασταύρωση (crossover) για παραγωγή καινούριων χρωμοσωμάτων

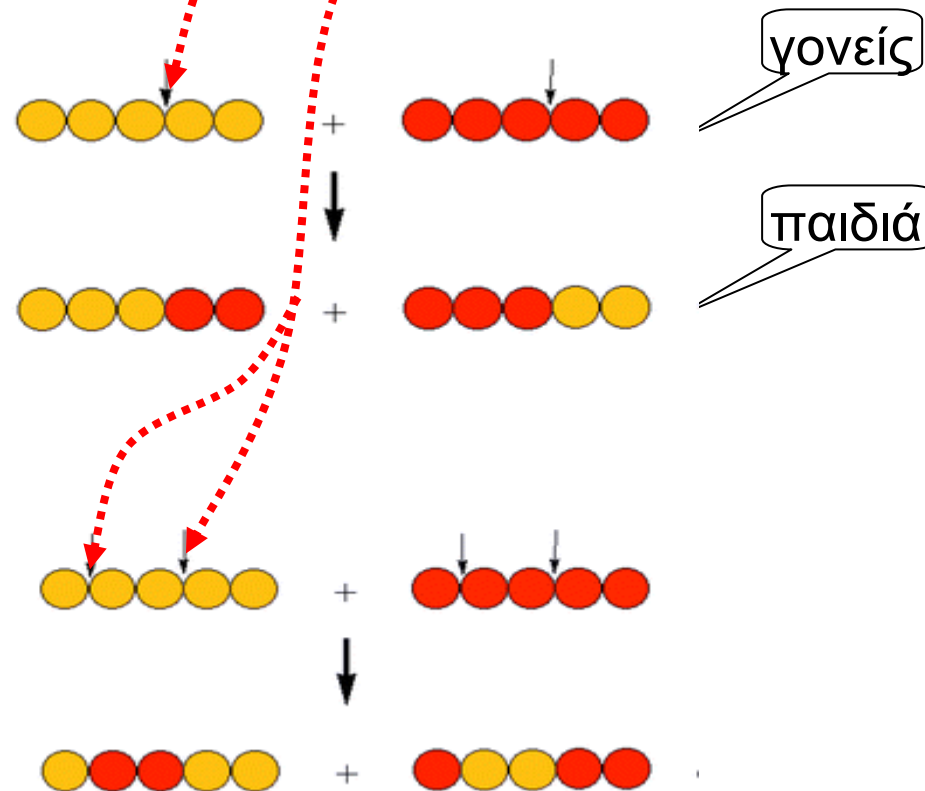
Βήματα:

1. Τυχαία επιλογή ζεύγους «γονέων»
2. Τυχαία επιλογή του σημείου διασταύρωσης
3. Διασταύρωση του «γενετικού» κώδικα



Τα νέα χρωμοσώματα αποτελούν την καινούρια γενεά, υπολογίζεται συνάρτηση καταλληλότητας (*fitness*) και συνεχίζεται η διαδικασία αναπαραγωγής.

Διασταύρωση σε 1 ή 2 σημεία



Διασταύρωση - παράδειγμα

- Το πρόβλημα: $\max x^2$ στο σύνολο $\{0,1,\dots,31\}$
- Λύση με ΓΑ. Χρωμόσωμα 4 ψηφίων πχ. 01101 \leftrightarrow 13
- Μέγεθος πληθυσμού : 4
- Διασταύρωση σε ένα σημείο

String no.	Mating pool	Crossover point	Offspring after xover	x Value	Fitness $f(x) = x^2$
1	0 1 1 0 1	4	0 1 1 0 0	12	144
2	1 1 0 0 0	4	1 1 0 0 1	25	625
2	1 1 0 0 0	2	1 1 0 1 1	27	729
4	1 0 0 1 1	2	1 0 0 0 0	16	256
Sum					1754
Average					439
Max					729

ΓΑ. – Μετάλλαξη (mutation)

Στη διαδικασία αυτή γίνεται αυθαίρετα και τυχαία η αλλαγή μερικών ψηφίων (γονιδίων) στον πληθυσμό

Η συχνότητα αυτή είναι πολύ μικρή. Πχ. 0.005/bit/γενεά

Σκοπός είναι να «απεγκλωβίσει» την διαδικασία από μερική λύση.

Μία προτεινόμενη τιμή p_m ορίζεται ως εξής:

$$\frac{1}{\text{μέγεθος πληθυσμού}} < p_m < \frac{1}{\text{μήκος χρωμοσώματος}}$$

Μετάλαξη - παράδειγμα

- Το πρόβλημα: $\max x^2$ στο σύνολο $\{0,1,\dots,31\}$
- Λύση με ΓΑ. Χρωμόσωμα 5 ψηφίων πχ. 01101 \leftrightarrow 13
- Μέγεθος πληθυσμού : 4
- Διασταύρωση σε ένα σημείο

# Χρωμοσωμ	Απόγονοι μετά από διασταύρωση	Απόγονοι μετά από μετάλαξη	Τιμή x	Fitness $f(x) = x^2$
1	0 1 1 0 0	1 1 1 0 0	26	676
2	1 1 0 0 1	1 1 0 0 1	25	625
2	1 1 0 1 1	1 1 0 1 1	27	729
4	1 0 0 0 0	1 0 1 0 0	18	324
Sum				2354
Average				588.5
Max				729

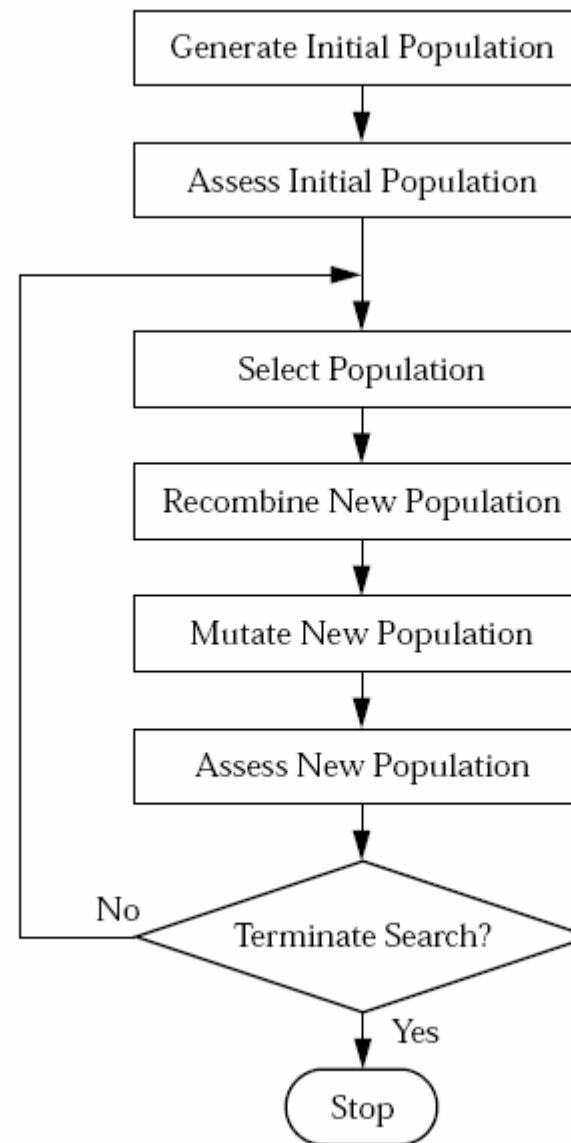
Τερματισμός του γενετικού αλγορίθμου

- Όταν το ποσοστό των χρωμοσωμάτων που δεν μεταβάλλεται έχει ξεπεράσει ένα συγκεκριμένο όριο .
- Όταν ο μέσος όρος της τιμής καταλληλότητας για όλα τα χρωμοσώματα ενός πληθυσμού και για έναν αριθμό διαδοχικών πληθυσμών δεν βελτιώνεται.
- Επειδή ο GA είναι μια πιθανολογική μέθοδος αναζήτησης, είναι δύσκολο να διευκρινιστούν τυπικά κριτήρια σύγκλισης.
- Η ικανότητα ενός πληθυσμού μπορεί να παραμείνει στατική για έναν αριθμό γενεών προτού να βρεθεί ένα ανώτερο άτομο, έτσι η εφαρμογή των συμβατικών κριτηρίων λήξης γίνεται προβληματική.
- Μια κοινή πρακτική είναι να τερματίζεται ο GA ύστερα από έναν αριθμό γενεών. Στη συνέχεια εξετάζεται η ποιότητα των καλύτερων μελών του πληθυσμού.
- Εάν καμία αποδεκτή λύση δε βρεθεί, ο GA μπορεί να ξαναξεκινήσει

Απλός Γενετικός Αλγόριθμος - ΣΥΝΟΨΗ

-
- | | |
|--------|---|
| Step 1 | Let $i = 0$; |
| Step 2 | Initialize a population $P(i)$; |
| Step 3 | Evaluate $P(i)$; |
| Step 4 | If SGA reaches termination conditions, then go to Step 8; |
| Step 5 | Perform reproduction on $P(i)$ to generate $P(i + 1)$; |
| Step 6 | Perform crossover and mutation on $P(i + 1)$; |
| Step 7 | Let $i = i + 1$ and goto Step 3; |
| Step 8 | Stop. |
-

Απλός Γενετικός Αλγόριθμος - διάγραμμα



Παράδειγμα 1^ο

Προσαρμογή γραμμής σε δεδομένα

x	y
1.0	1.0
2.0	2.0
3.0	3.0
6.0	6.0

Για τις τιμές x,y του πίνακα να γίνει η προσαρμογή της ευθείας $y=C_1x+C_2$

Παράδειγμα 1 -Πρώτη επανάληψη

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
String number	String	C_1 (bin.)	C_1	C_2 (bin.)	C_2	y_1	y_2	y_3	y_4	$f(x) =$ $400 - \sum (y_i - y_r)^2$	Expected count = f/average	Actual count
1	000111 010100	7	-1.22	20	0.22	-1.00	-2.22	-4.67	-7.11	147.49	0.48	0
2	010010 001100	18	0.00	12	-0.67	-0.67	-0.67	-0.67	-0.67	332.22	1.08	1
3	010101 101010	21	0.33	42	2.67	3.00	3.33	4.00	4.67	391.44	1.27	2
4	100100 001001	36	2.00	9	-1.00	1.00	3.00	7.00	11.00	358.00	1.17	1
Sum										1229.15		
Average										307.29		
Maximum										391.44		

Παράδειγμα 1 – Δεύτερη επανάληψη

Selected strings	New strings	C_1 (bin.)	C_1	C_2 (bin.)	C_2	y_1	y_2	y_3	y_4	$f(x) =$ $400 - \sum (y_i - y_r)^2$	Expected count = f/average	Actual count
0101 01 101010	010110 001100	22	0.44	12	-0.67	-0.22	0.22	1.11	2.00	375.78	1.15	1
0100 10 001100	010001 101010	17	-0.11	42	2.67	2.56	2.44	2.22	2.00	380.78	1.17	2
010101 101 010	010101 101001	21	0.33	41	2.56	2.89	3.22	3.89	4.56	292.06	0.90	1
100100 001 001	100100 001010	36	2.0	10	-0.89	1.11	3.11	7.11	11.11	255.73	0.78	0
Sum										1304.35		
Average										326.09		
Maximum										380.78		

Παράδειγμα 2^ο

Υπολογισμός της συνάρτησης συμμετοχής

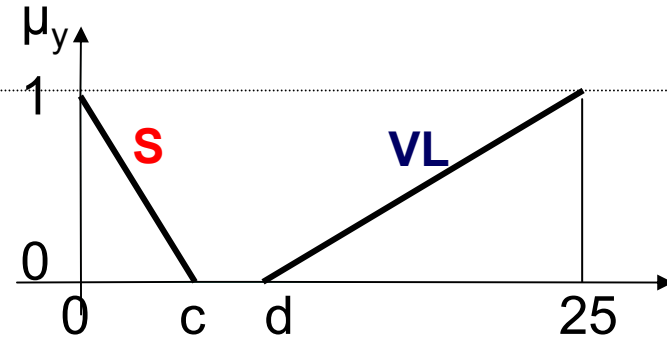
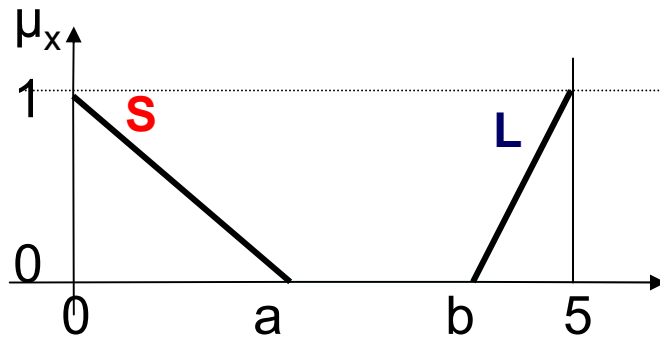
Δίνονται τα δεδομένα εισόδου εξόδου:

x	1	2	3	4	5
y	1	4	9	16	25

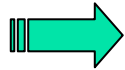
Και η σχέση των συνόλων (κανόνες):

x	S _{mall}	L _{arge}
y	S _{mall}	V _{ery} L _{arge}

Ζητούνται οι παράμετροι των ασαφών συνόλων



έχουμε 4 παραμέτρους (a,b,c,d) και χρησιμοποιούμε 6bits για κάθε μία - χρωμόσωμα $24^{\omega v}$ γονιδίων –bits.



1 Αρχικός πληθυσμός 4 χρωμοσωμάτων:

2 Οι 4 binary μεταβλητές (a,b,c,d)

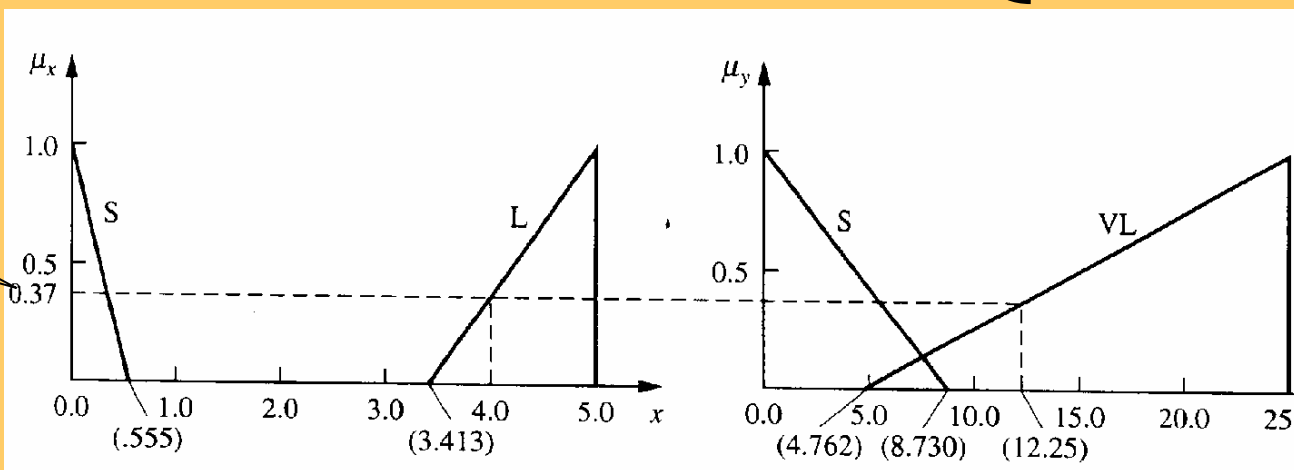
3 Οι 4 αντίστοιχοι δεκαδικοί αριθμοί

4 Απεικόνιση των a,b στο διάστημα 0-5

5 Απεικόνιση των c,d στο διάστημα 0-25

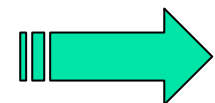
$$C_i = C_{\min} + \frac{B_i}{2^l - 1} (C_{\max} - C_{\min})$$

1η γενειά
1ο
χρωμόσωμα



6 Εφαρμογή των κανόνων –υπολογισμός του y

7 Υπολογισμός της τιμής καταλληλότητας (fitness)



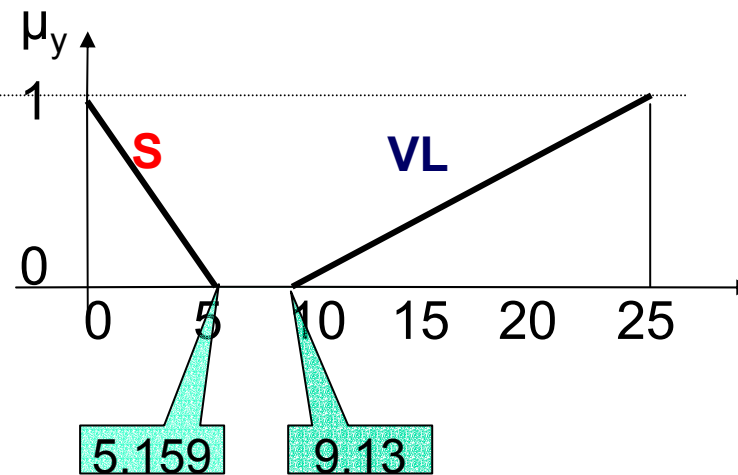
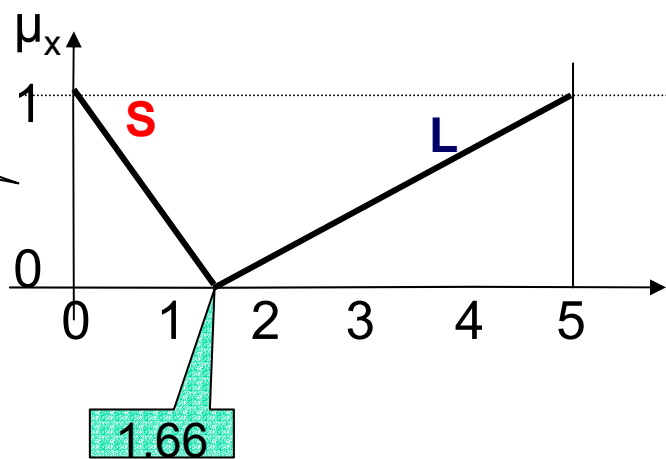
Κατώφλιο=0.75



	a	b	c	d	a	b	c	d	a	b	c	d	y(1)	y(2)	y(3)	y(4)	y(5)	1000-Σe ²	f/f _a	
1	000111	010100	010110	110011	7	20	22	51	0.56	1.59	8.73	20.24	0	0	0	7.49	25	829.58	1.18	1
2	010010	001100	101100	100110														521.11	0.74	0
3	010101	101010	001101	101000														890.46	1.27	2
4	100100	001001	101100	100011														559.67	0.8	1
άθροισμα																		2800.82		
Μέση τιμή f _a																		700.2		

	Προηγούμενη γενεά	Νέα γενεά (διασταύρωση)																1000-Σe ²	f/f _a	1
1	000111 0101 00 010110 110011	000111 010110 001101 101000																900.98	1.1	2
2	010101 1010 10 001101 101000	010101 1010 00 010110 110011																961.3	1.18	1
3	010101 101010 001101 1 01000	010101 101010 001101 1 00011																840.78	1.03	0
4	100100 001001 101100 1 00011	100100 001001 101100 1 01000																569.32	0.7	
άθροισμα																		3272.36		
Μέση τιμή f _a																		818.09		

1η γενειά
το καλύτερο
χρωμόσωμα



Παράδειγμα 3^ο

Επεξεργασία εικόνας

Δίνεται η **g** και ζητείται η **f**

$$\|g - \hat{f} * h\|^2 = \varepsilon,$$

- a) Θολωμένη εικόνα με θόρυβο Gauss
- b) Αποτέλεσμα μετά την 1η γενεά
- c) «» μετά την 200η γενεά
- d) «» μετά την 1000 γενεά



(a)

(b)



(c)



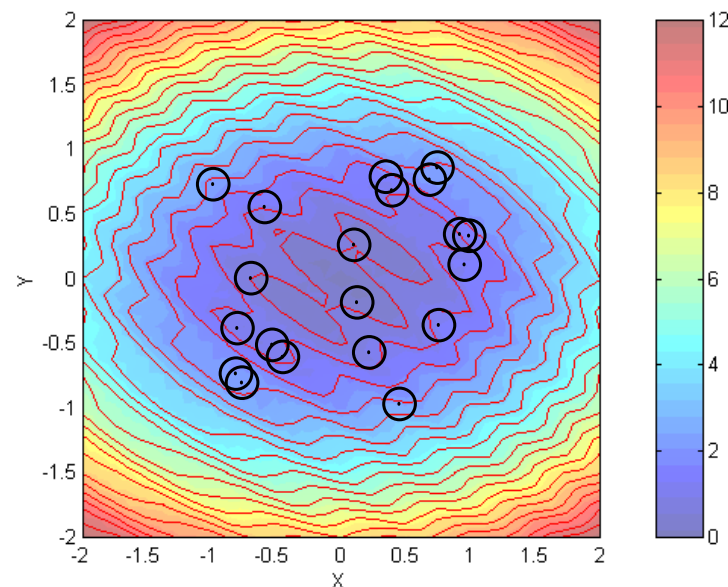
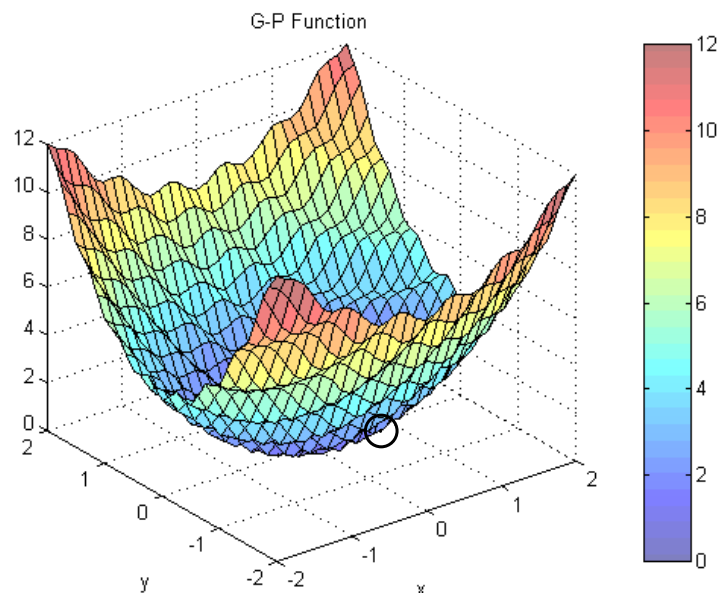
(d)



(e)

Παράδειγμα 4^ο: Εύρεση του ελαχίστου της

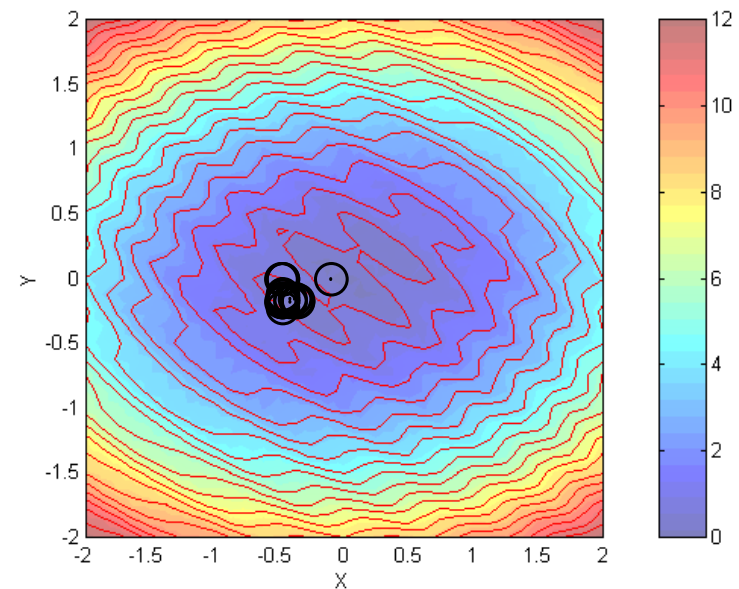
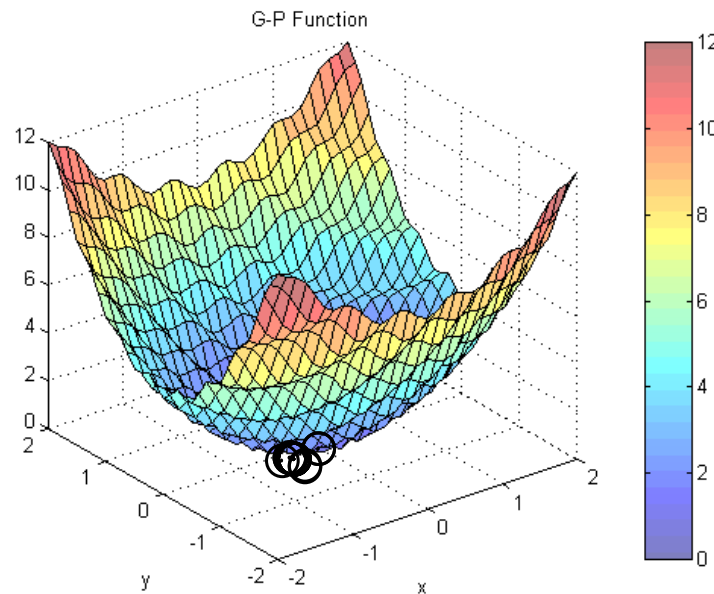
$$z = x^2 + 2y^2 - 0.3 \cos(3\pi x + 4\pi y) + 0.3$$



Γραφική παράσταση της z για τον αρχικό πληθυσμό

Κάτοψη της z για τον αρχικό πληθυσμό

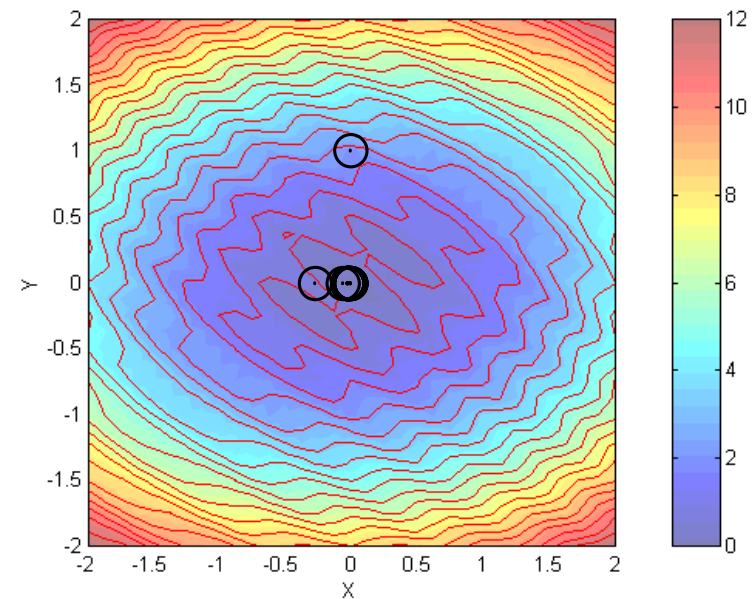
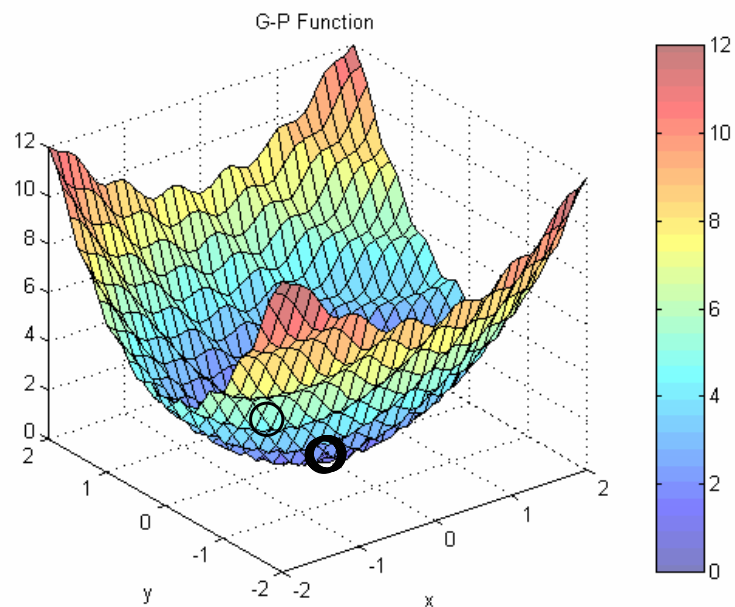
Παράδειγμα 4 (συν)



Γραφική παράσταση της z για την 10η γενεά

Κάτοψη της z για την 10η γενεά

Παράδειγμα 4 (συν)

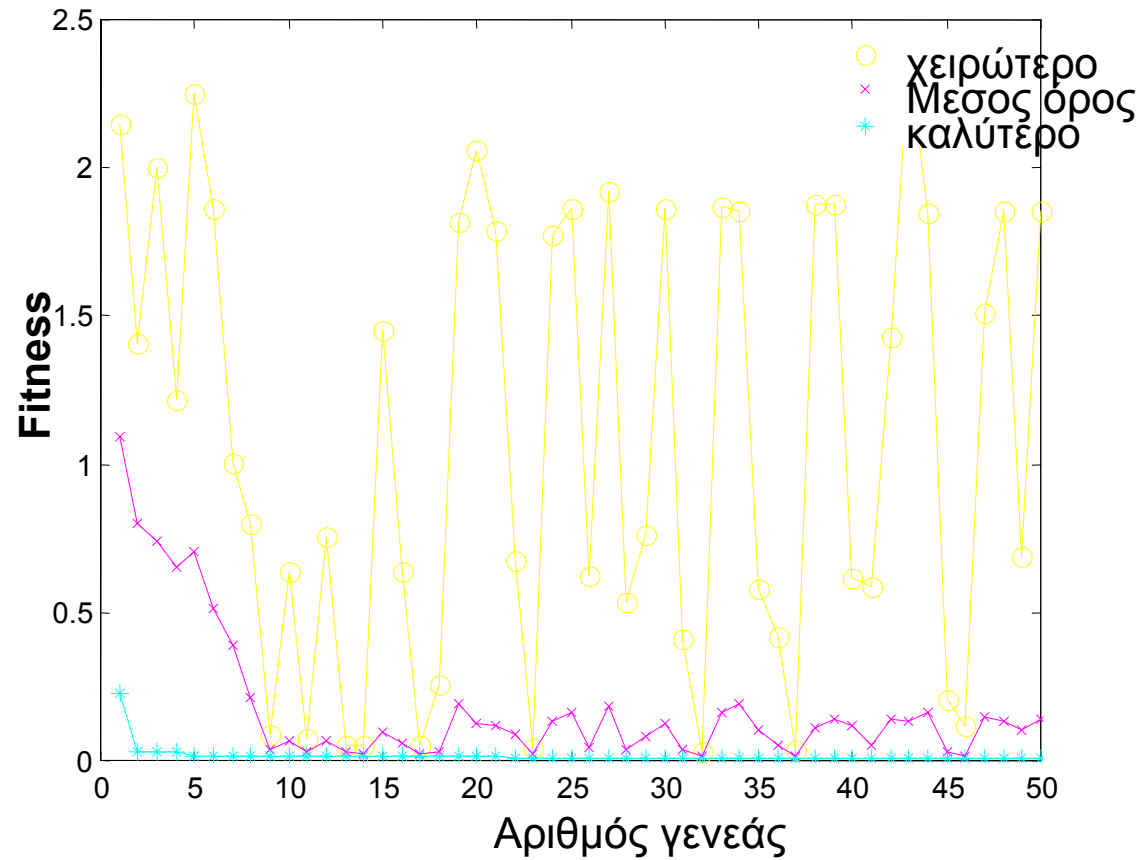


Γραφική παράσταση της z για την 50η γενεά

Κάτοψη της z για την 50η γενεά

Παράδειγμα 4 (συν)

Γραφική παράσταση της εξέλιξης του γενετικού αλγορίθμου



Βιβλιογραφία - διαδίκτυο

1. <http://cs.felk.cvut.cz/~xobitko/ga/>
2. <http://www.rennard.org/alife/english/gavintrgb.html#rTit01>
3. <http://www.sci.hkbu.edu.hk/msc/full/regina/node1.html>
4. <http://www.aboriginemundi.com/genelab/>
5. Goldberg, D.E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Redwood City, 1998
6. Holland, J.H., *Adaptation in Natural and Artificial Systems*, Univ. of Michigan Press, Ann Arbor, Mich., 1975
7. Heistermann, Jochen, *Genetische Algorithmen*, B.G.Teubner, Stuttgart, 1994
8. T. Ross *Fuzzy Logic with Engineering Applications* McGraw-Hill, Inc., USA, 1995

ασκήσεις – εργασίες

1. μελέτη του “traveler salesman problem”
(<http://www.generation5.org/content/2001/tspapp.asp>)
2. Λεπτομέρειες στη λύση του 2ου παραδείγματος
3. Παρουσίαση του σχετικού toolbox of matlab (5 GUI tools)
4. Υπολογείστε το ελάχιστο της εξίσωσης Rosenbrock:

$$f = (1 - x_1)^2 + 100(x_2 - x_1^2)^2$$