

Εισαγωγή στους γενετικούς αλγόριθμους

Ευάγγελος Π. Φινδανής
Τοπογράφος Μηχανικός MSc

Εισαγωγή

Οι γενετικοί αλγόριθμοι είναι μέθοδοι βελτιστοποίησης βασισμένοι στις αρχές της φυσικής εξέλιξης. Οι μέθοδοι αυτοί λειτουργούν διατηρώντας ένα πληθυσμό δυνατών λύσεων του προβλήματος και εφαρμόζοντας πάνω σε αυτόν διάφορες διαδικασίες εμπνευσμένες από την εξέλιξη των ειδών. Έτσι μέσω αυτών των διαδικασιών δημιουργούνται νέοι πληθυσμοί καλύτερων λύσεων που προσεγγίζουν την τελική λύση του προβλήματος.

Αλγόριθμοι αυτού του τύπου δανείζονται ορολογία από το χώρο της γενετικής. Κάθε άτομο του πληθυσμού αποτελείται από χρωμοσώματα τα οποία με την σειρά τους αποτελούνται από γονίδια διατεταγμένα σε γραμμική ακολουθία. Επίσης, ο πληθυσμός αυτός μέσω του αλγορίθμου βελτιώνεται καθώς με το πέρασμα των γενεών προσαρμόζεται καλύτερα στο περιβάλλον του (το ρόλο του περιβάλλοντος παίζει η συνάρτηση στόχου που θέλουμε να βελτιστοποιήσουμε). Έτσι μέσω του αλγορίθμου τα άτομα του πληθυσμού που αποτελούν καλή λύση του προβλήματος αναπαράγονται ενώ τα άτομα τα οποία δεν είναι καλές λύσεις απομακρύνονται.

Μερικά από τα πλεονεκτήματα των γενετικών αλγορίθμων είναι τα εξής:

1. Λύνουν δύσκολα προβλήματα γρήγορα και αξιόπιστα.
2. Συνεργάζονται εύκολα με μοντέλα και συστήματα.
3. Είναι εξελίξιμοι και επεκτάσιμοι.
4. Είναι δυνατόν να συμμετέχουν σε υβριδικές μορφές με άλλες μεθόδους.
5. Οι συναρτήσεις που επεξεργάζονται δεν απαιτούν περιορισμούς.

Οι διαδικασίες που υφίσταται η κάθε γενιά είναι οι ακόλουθες: Κατά την διάρκεια μίας γενιάς ο αλγόριθμος αξιολογεί κάθε άτομο του πληθυσμού βάσει της προσαρμοστικότητας στο περιβάλλον και αφού ολοκληρωθεί η αξιολόγηση όλων των μελών δημιουργείται ένας νέος πληθυσμός με τα πιο κατάλληλα άτομα του πληθυσμού της προηγούμενης γενιάς. Έπειτα, αυτά τα άτομα υπόκεινται τις διαδικασίες της διασταύρωσης και της μετάλλαξης με αποτέλεσμα να δημιουργούνται νέες πιθανές λύσεις στο πρόβλημα. Η διαδικασία της διασταύρωσης σχηματίζει νέους απογόνους

ανταλλάσσοντας στοιχεία από τους γονείς. Συνεπώς αυτή η διαδικασία εξυπηρετεί την ανταλλαγή πληροφορίας μεταξύ των πιθανών λύσεων. Η διαδικασία της μετάλλαξης πραγματοποιεί τυχαία αλλαγή των γονιδίων των ατόμων και εξυπηρετεί την εισαγωγή νέων πιθανών λύσεων.

Τα βασικά βήματα ενός γενετικού αλγόριθμου που επιλύει ένα πρόβλημα βελτιστοποίησης μίας συνάρτησης στόχου f συνοψίζονται στα εξής:

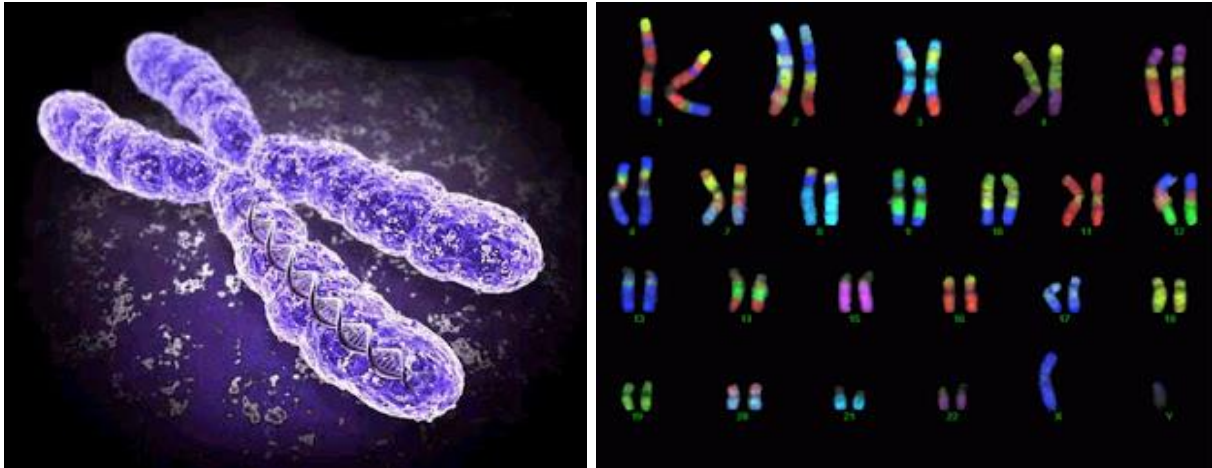
1. Δημιουργείται ένας τυχαίος πληθυσμός πιθανών λύσεων (αρχικοποίηση).
2. Κάθε λύση αξιολογείται με την βοήθεια της συνάρτησης f .
3. Δημιουργείται ένας νέος πληθυσμός βάσει της απόδοσης του κάθε μέλους του προηγούμενου πληθυσμού.
4. Εφαρμόζονται στον νέο πληθυσμό οι τελεστές της διασταύρωσης και της μετάλλαξης.
5. Με την ολοκλήρωση του βήματος 4 δημιουργείται η νέα γενιά και επιστρέφουμε στο βήμα 2.
6. Αν μετά από κάποιο αριθμό γενεών καμία βελτίωση δεν παρουσιάζεται, ο αλγόριθμος τερματίζεται και το καλύτερο χρωμόσωμα αποτελεί την λύση του προβλήματος.

Το βιολογικό υπόβαθρο

Καταρχήν, η ονομασία ενός γενετικού αλγορίθμου παραπέμπει στην γενετική, δηλαδή σε έναν κλάδο της βιολογίας ο οποίος ασχολείται με την μελέτη του γονιδιώματος των οργανισμών. Το γονιδίωμα είναι το σύνολο του γενετικού υλικού που φέρει ένα άτομο, δηλαδή ουσιαστικά είναι το σύνολο των πληροφοριών που καθορίζουν τα χαρακτηριστικά ενός ατόμου. Άτομα του ίδιου είδους φέρουν την ίδια ποσότητα γονιδιώματος. Συγκεκριμένα, το γονιδίωμα οργανώνεται σε χρωμοσώματα και κάθε χρωμόσωμα αποτελείται από έναν συγκεκριμένο αριθμό γονιδίων. Έτσι τα άτομα του ίδιου είδους έχουν τον ίδιο αριθμό χρωμοσωμάτων. Παραδείγματος χάριν, το ανθρώπινο γονιδίωμα αποτελείται από 46 χρωμοσώματα, το γονιδίωμα της γάτας οργανώνεται σε 38 χρωμοσώματα ενώ αυτό ενός σκύλου έχει 78 χρωμοσώματα.

Όμως στην αρχή της ζωής στον πλανήτη μας δεν υπήρχαν ούτε σκύλοι, ούτε άνθρωποι. Υπήρχαν προκαρυωτικά κύτταρα τα οποία με το πέρασμα του χρόνου εξελίχθηκαν στα είδη που γνωρίζουμε σήμερα. Δηλαδή, με την διαδικασία της εξέλιξης η αρχική μορφή ζωής οδηγήθηκε σε άλλες πιο εξελιγμένες μορφές οι οποίες προσαρμόστηκαν στο γήινο

περιβάλλον. Επομένως, συμπεραίνουμε πως η διαδικασία της εξέλιξης είναι μία διαδικασία **βελτιστοποίησης** των μορφών ζωής στο περιβάλλον τους.



Εικόνα 1: a) Ένα ζεύγος χρωμοσωμάτων. b) Τα 46 χρωμοσώματα ενός ανθρώπου.

Ας δούμε όμως πως λειτουργεί η διαδικασία της εξέλιξης. Έστω πως έχουμε μία αγέλη λύκων που ζουν σε ένα συγκεκριμένο οικοσύστημα (περιβάλλον). Η αγέλη αυτή αποτελεί τον αρχικό μας πληθυσμό. Ορισμένοι από αυτούς τους λύκους είναι πιο έξυπνοι και πιο γρήγοροι με αποτέλεσμα να βρίσκουν ευκολότερα τροφή. Αυτό έχει ως αποτέλεσμα ορισμένοι από τους πιο αργούς λύκους να μην βρουν φαγητό και να πεθάνουν. Συνεπώς, γίνεται επιλογή των καλύτερων ατόμων του πληθυσμού, δηλαδή των ατόμων που προσαρμόστηκαν καλύτερα στο περιβάλλον μειωμένης τροφής και επιβίωσαν (*φυσική επιλογή*). Έπειτα, οι λύκοι που επιβίωσαν θα ζευγαρώσουν μεταξύ τους (*διασταύρωση*) με αποτέλεσμα να δημιουργήσουν μία νέα γενιά η οποία κατά μέσο όρο θα είναι πιο γρήγορη και έξυπνη από την προηγούμενη γενιά. Επιπλέον, εξαιτίας διάφορων παραγόντων η νέα γενιά θα υποστεί τυχαίες μεταλλάξεις και συνεπώς ορισμένα χαρακτηριστικά των νέων ατόμων θα αλλάξουν τυχαία (*μετάλλαξη*). Έτσι από την αρχική γενιά προέκυψε μία νέα γενιά η οποία είναι γενικά πιο προσαρμοσμένη στο περιβάλλον, περιέχει χαρακτηριστικά των προγόνων της αλλά μέσω της διαδικασίας της μετάλλαξης έχουν προστεθεί σε αυτή ορισμένα νέα χαρακτηριστικά (είτε θετικά είτε αρνητικά) τα οποία δεν υπήρχαν στους προγόνους. Η νέα γενιά θα υποστεί πάλι τις διαδικασίες της φυσικής επιλογής, της διασταύρωσης και της μετάλλαξης μέχρι να προκύψει η επόμενη γενιά. Με αυτόν τον τρόπο η αγέλη των λύκων εξελίσσεται και προσαρμόζεται στο περιβάλλον της.

Για να δούμε τώρα τι ακριβώς κάνει ένας γενετικός αλγόριθμος ας κάνουμε την εξής σκέψη: Το πρόβλημα βελτιστοποίησης μίας συνάρτησης στόχου f ουσιαστικά είναι ένα

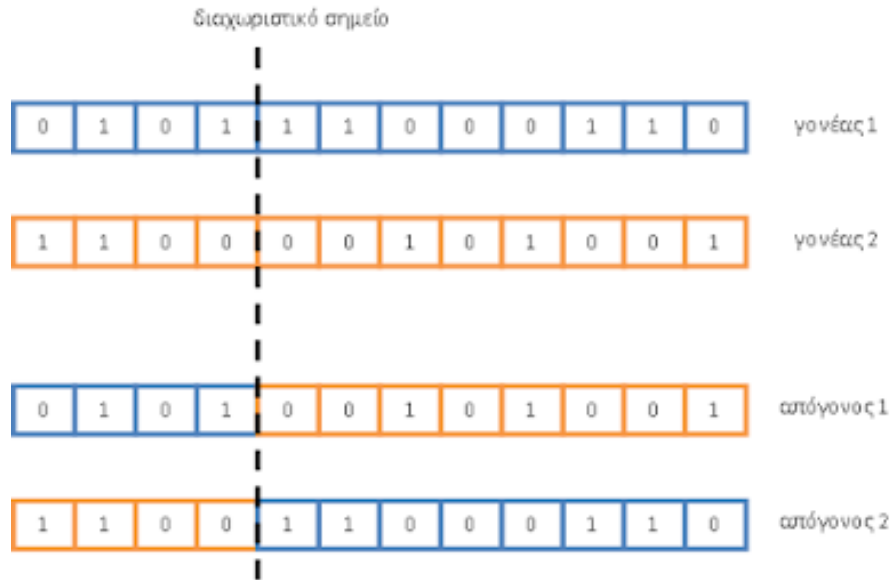
πρόβλημα εύρεσης μίας αριθμητικής τιμής που είναι βέλτιστη για την συγκεκριμένη συνάρτηση. Οπότε μέσα από διάφορες αριθμητικές τιμές καλούμαστε να βρούμε αυτή που προσαρμόζεται καλύτερα στην συνάρτηση f . Η αναλογία είναι εμφανής: **οι αριθμητικές τιμές είναι τα άτομα τα οποία πρέπει να εξελιχθούν και να προσαρμοστούν στην συνάρτηση στόχου** (περιβάλλον).

Όμως η πραγματική διαδικασία της εξέλιξης έχει να κάνει με το γονιδίωμα των οργανισμών. Δηλαδή το γονιδίωμα είναι αυτό που υφίσταται τις διαδικασίες της φυσικής επιλογής, της διασταύρωσης και της μετάλλαξης. Αν θέλουμε να εφαρμόσουμε την διαδικασία της εξέλιξης πάνω σε αριθμούς τι γονιδίωμα θα χρησιμοποιήσουμε;

Η λύση σε αυτό το ζήτημα είναι απλή. Καταρχήν θεωρούμε πως όλα τα άτομα (αριθμοί) που θα προσπαθήσουμε να εξελίξουμε αποτελούνται μόνο από ένα χρωμόσωμα. Έπειτα, ορίζουμε τον αριθμό των γονιδίων αυτού του μοναδικού χρωμοσώματος. Και τέλος, μετατρέπουμε τις αριθμητικές τιμές που μας ενδιαφέρουν στο δυαδικό σύστημα. Έτσι παραδείγματος χάριν αν επιλέξουμε πως τα χρωμοσώματα του πληθυσμού μας αποτελούνται από 8 γονίδια, ο αριθμός 10 αντιστοιχεί στο χρωμόσωμα 00001010 ενώ ο αριθμός 142 αντιστοιχεί στο χρωμόσωμα 10001110.

Οπότε ένας γενετικός αλγόριθμος ακολουθεί τις εξής διαδικασίες:

- 1) Δημιουργεί έναν τυχαίο αρχικό πληθυσμό αριθμών (πχ 1000 τυχαίοι αριθμοί).
- 2) Από τον παραπάνω πληθυσμό επιλέγει τους αριθμούς x_i που αντιστοιχούν σε μεγάλες τιμές $f(x_i)$
- 3) Γίνεται εφαρμογή του τελεστή της διασταύρωσης στον αρχικό πληθυσμό. Μία διασταύρωση φαίνεται στην εικόνα 2 όπου οι δύο γονείς διασταυρώνονται ώστε να δώσουν δύο απογόνους.
- 4) Γίνεται εφαρμογή του τελεστή της μετάλλαξης στους απογόνους. Στην διαδικασία αυτή τα γονίδια στα χρωμοσώματα ορισμένων απογόνων αλλάζουν τυχαία, δηλαδή εάν ένα γονίδιο είναι 0 γίνεται 1 και το αντίστροφο (εικόνα 3).
- 5) Ο νέος πληθυσμός (απόγονοι) παίζει τον ρόλο του αρχικού πληθυσμού με αποτέλεσμα να επιστρέφουμε στο βήμα 2.
- 6) Η παραπάνω διαδικασία πραγματοποιείται για έναν συγκεκριμένο αριθμό γενιών που επιλέγει ο χρήστης. Στο τέλος επιλέγουμε ως λύση τον αριθμό (άτομο) που μεγιστοποιεί την συνάρτηση στόχου f . Να σημειωθεί πως ο αριθμός αυτός δεν είναι απαραίτητο να βρίσκεται ανάμεσα στα άτομα της τελευταίας γενιάς αλλά μπορεί να βρίσκεται σε οποιαδήποτε γενιά.



Εικόνα 2: Τα χρωμοσώματα δύο γονέων συνδυάζονται για να δώσουν δύο απογόνους.



Εικόνα 3: Η διαδικασία της μετάλλαξης σε ένα χρωμόσωμα 5 γονιδίων.

Με αυτόν τον τρόπο, λοιπόν, ένας γενετικός αλγόριθμος μιμείται την διαδικασία της εξέλιξης ώστε να βελτιστοποιήσει μία συνάρτηση. Η ευελιξία και η ταχύτητα των γενετικών αλγορίθμων τους έχει εδραιώσει ως ένα σημαντικό κομμάτι της επιστήμης των υπολογιστών. Έτσι ολοένα και περισσότερα προβλήματα αντιμετωπίζονται με τέτοιες τεχνικές.

Παράδειγμα κώδικα ενός γενετικού αλγόριθμου

Για να παρουσιάσουμε την δομή ενός κώδικα γενετικού αλγορίθμου, θα επιλύσουμε το πρόβλημα βελτιστοποίησης της συνάρτησης στόχου $f(x) = |5 - x^2 - y^2|$. Η συνάρτηση αυτή είναι δύο μεταβλητών κάτι που θα πρέπει να λάβουμε υπόψιν στον κώδικα. Τα δεδομένα που θα χρειαστούμε είναι το πεδίο ορισμού της συνάρτησης f , τα δεκαδικά ψηφία που θα κρατήσουμε και οι πιθανότητες διασταύρωσης και μετάλλαξης. Θεωρούμε ότι το πεδίο ορισμού είναι το $x \in [-2, 2]$ και $y \in [-2, 2]$, οι υπολογισμοί θα γίνουν με ακρίβεια 4 δεκαδικών ψηφίων και οι πιθανότητες διασταύρωσης και μετάλλαξης είναι 25% και 1% αντίστοιχα. Ο κώδικας που επιλύει το παραπάνω πρόβλημα είναι ο ακόλουθος:

Private Sub Command1_Click()

'δήλωση των μεταβλητών

Dim xk(20) As Double, ev(20) As Double, p(20) As Double, m(20) As Double
Dim z(20) As Double, B() As Integer, val1(20) As Double, y1() As Double, y2() As Double
Dim val2(20) As Double, A() As Integer, cross(20) As Double, c() As Integer
Dim id(20) As Integer, xs() As Integer, xf() As Integer, random() As Double
Dim a1v As Double, b1v As Double, a2v As Double, b2v As Double, q As Integer
Dim pc As Double, pm As Double, valn1(20) As Double, valn2(20) As Double

'εισαγωγή των δεδομένων από την φόρμα (εικόνα 4)

a1v = Text1.Text
b1v = Text2.Text
q = Text3.Text
a2v = Text4.Text
b2v = Text5.Text
pc = Text6.Text
pm = Text7.Text

'υπολογισμός του μεγέθους του χρωμοσώματος

Do Until $2^{n1} > (b1v - a1v) * (10^q)$
n1 = n1 + 1
Loop

Do Until $2^{n2} > (b2v - a2v) * (10^q)$
n2 = n2 + 1
Loop

'επαναδήλωση του μεγέθους ορισμένων πινάκων

ReDim B(20, n1 + n2) As Integer, y1(n1) As Double, y2(n2) As Double, A(20, n1 + n2) As Integer
ReDim c(20, n1 + n2) As Integer, xs(20, n1 + n2) As Integer, xf(20, n1 + n2) As Integer
ReDim random(20, n1 + n2) As Double

'δημιουργία τυχαίων χρωμοσωμάτων (αρχικοποίηση)

For i = 1 To 20
For j = 1 To n1 + n2
A(i, j) = Int(Rnd + 0.5)
Next j
Next i

'έναρξη των επαναλήψεων για 10000 γενιές

For gen = 1 To 10000

'μετατροπή των χρωμοσωμάτων στο δεκαδικό σύστημα

For i = 1 To 20

```

valn1(i) = 0
For j = 1 To n1
valn1(i) = valn1(i) + A(i, j) * (2 ^ (n1 - j))
val1(i) = a1v + valn1(i) * (b1v - a1v) / (2 ^ n1 - 1)
Next j
Next i

```

```

For i = 1 To 20
valn2(i) = 0
For j = n1 + 1 To n1 + n2
valn2(i) = valn2(i) + A(i, j) * (2 ^ (n2 + n1 - j))
val2(i) = a2v + valn2(i) * (b2v - a2v) / (2 ^ n2 - 1)
Next j
Next i

```

Υπολογισμός της απόδοσης των χρωμοσωμάτων βάση της συνάρτησης στόχου

```

For i = 1 To 20
ev(i) = 5 - (val1(i)) ^ 2 - (val2(i)) ^ 2
F = F + ev(i)
Next i

```

Επιλογή του χρωμοσώματος που έχει προσαρμοστεί καλύτερα στο περιβάλλον του

```

For i = 1 To 20
If ev(i) > final Then
final = ev(i)
End If
Next i

```

Υπολογισμός της απόδοσης του κάθε χρωμοσώματος

```

For i = 1 To 20
p(i) = ev(i) / F
Next i

```

Δημιουργία αθροιστικών πιθανοτήτων και 20 τυχαίων αριθμών στο διάστημα [0,1]

```

For i = 1 To 20
m(i) = m(i - 1) + p(i)
z(i) = Rnd
Next i

```

Επιλογή των χρωμοσωμάτων βάση της απόδοσης τους

```

For i = 1 To 20
For j = 1 To 20
If z(i) < m(j) And z(i) > m(j - 1) Then
For k = 1 To n1 + n2

```

```
B(i, k) = A(j, k)
Next k
End If
Next j
Next i
```

'επιλογή χρωμοσωμάτων για την διασταύρωση (crossover)

```
kn = 0
For i = 1 To 20
cross(i) = Rnd
If cross(i) < pc Then
kn = kn + 1
For j = 1 To n1 + n2
c(kn, j) = A(i, j)
id(kn) = i
Next j
End If
Next i
```

'αν ο αριθμός των χρωμοσωμάτων που επιλέχθηκε είναι περιττός απορρίπτουμε το τελευταίο

```
If kn Mod 2 = 1 Then
kn = kn - 1
End If
```

'ανταλλαγή των γονιδίων των χρωμοσωμάτων

```
For i = 1 To kn / 2
pos = Int(Rnd * (n1 + n2 - 2) + 1)
For j = 1 To pos
xf(i, j) = c(i, j)
xs(i, j) = c(i + kn / 2, j)
Next j
Next i
```

```
For i = 1 To kn / 2
For j = 1 To pos
c(i, j) = xs(i, j)
c(i + kn / 2, j) = xf(i, j)
Next j
Next i
For i = 1 To kn
For j = 1 To n1 + n2
A(id(i), j) = c(i, j)
Next j
Next i
```


'δημιουργία τυχαίων αριθμών για την μετάλλαξη των γονιδίων (mutation)

```
For i = 1 To 20  
For j = 1 To n1 + n2  
random(i, j) = Rnd  
Next j  
Next i
```

'μετάλλαξη των γονιδίων

```
For i = 1 To 20  
For j = 1 To n1 + n2  
If random(i, j) < pm Then  
If A(i, j) = 1 Then  
A(i, j) = 0  
ElseIf A(i, j) = 0 Then  
A(i, j) = 1  
End If  
End If  
Next j  
Next i
```

'ξεκινάει η επόμενη γενιά

```
Next gen
```

'εκτύπωση του καλύτερου χρωμοσώματος από όλες τις γενιές στην φόρμα

```
Text8.Text = final
```

```
End Sub
```

Αποτελέσματα του κώδικα

The image shows two side-by-side screenshots of a Windows application window titled "Form1". Each window contains a set of input fields and a "resolve" button. The left window (a) shows empty input fields for parameters: a1, a2, b1, b2, q, pc, pm, and maxima. The right window (b) shows the same form after execution, with numerical values entered in the fields: a1: -2, a2: -2, b1: 2, b2: 2, q: 4, pc: .25, pm: .01, and maxima: 4.9999970.

Εικόνα 4: a) Η φόρμα εισαγωγής των δεδομένων. b) Η φόρμα μετά την εκτέλεση του κώδικα.

Τα δεδομένα του προβλήματος, θα εισαχθούν μέσω της φόρμας της εικόνας 4a. Στην φόρμα αυτή, a_1 και b_1 είναι τα άκρα του διαστήματος τιμών του x , a_2 και b_2 είναι τα άκρα του διαστήματος τιμών του y , q είναι το επιθυμητό πλήθος των δεκαδικών, p_c είναι η πιθανότητα διασταύρωσης και p_m είναι η πιθανότητα μετάλλαξης. Όταν συμπληρώσουμε τα παραπάνω πεδία και πατήσουμε το "resolve", στο πεδίο maxima λαμβάνουμε την τιμή του μέγιστου της συνάρτησης στο επιλεγμένο πεδίο ορισμού (εικόνα 4b). Οπότε μέσω του γενετικού αλγόριθμου βρίσκουμε ότι το μέγιστο στην συγκεκριμένη περιοχή του χώρου είναι 4.999997 . Πράγματι η μέγιστη τιμή της συνάρτησης $f(x,y)$ ισούται με 5 .