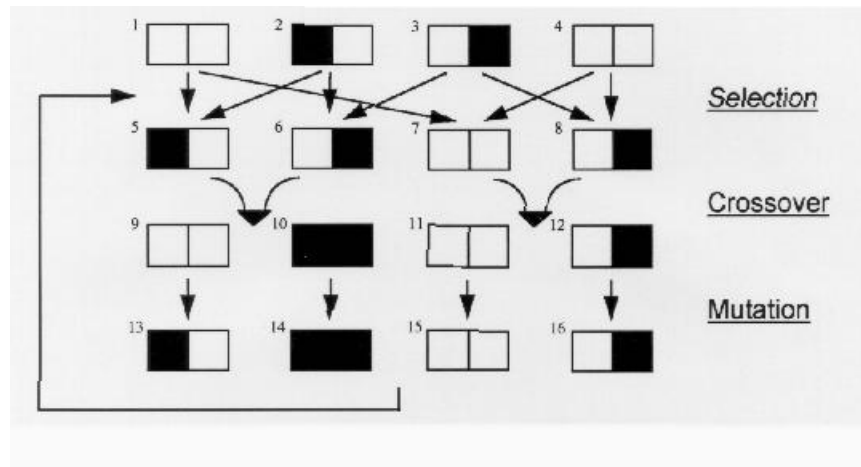


- **ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ II**
- **(ΓΕΝΕΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ - ΕΞΕΛΙΚΤΙΚΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ)**
- *Κεφάλαια 5,6,7,8*



Παράλληλοι ΕΑ

- Όλοι οι ΕΑ, επειδή διατηρούν έναν πληθυσμό από λύσεις, μπορούν από τη φύση τους να υλοποιηθούν παράλληλα.
- Επειδή οι ΓΑ χρησιμοποιούν διασταύρωση, η οποία αποτελεί ένα τρόπο διανομής της πληροφορίας, υπάρχουν δύο παραλλαγές των ΓΑ που τους επιτρέπουν πλήρη παραλληλία:
 - ✓ 1 . Μέθοδος: Υπάρχει ένας γενικός πληθυσμός με πολλούς επεξεργαστές για την αποτίμηση των λύσεων των ατόμων του πληθυσμού (global model).
 - ✓ 2 . Μέθοδος: Διατηρούνται ξεχωριστοί υπο-πληθυσμοί. Κάθε τόσο ένα άτομο από έναν από τους υπο-πληθυσμούς επιτρέπεται να μεταναστεύσει σε κάποιον άλλο υπο-πληθυσμό (island model).

Παράλληλοι ΕΑ

- Υπάρχει 3. Μέθοδος, η οποία διατηρεί επικαλυπτόμενες γειτονιές. (diffusion model, neighbourhood model)



- ✓ Εξαιτίας της έμφυτης φυσικής ικανότητας παραλληλίας μέσα στους ΕΑ, πάρα πολλές πρόσφατες εργασίες έχουν επικεντρωθεί στην εφαρμογή ΕΑ σε παράλληλες μηχανές.
- ✓ Χαρακτηριστικά είτε ένας επεξεργαστής αποτελεί μια ατομικότητα του πληθυσμού (στις μηχανές SIMD), είτε ένα υποσύνολο πληθυσμού (στις μηχανές MIMD).

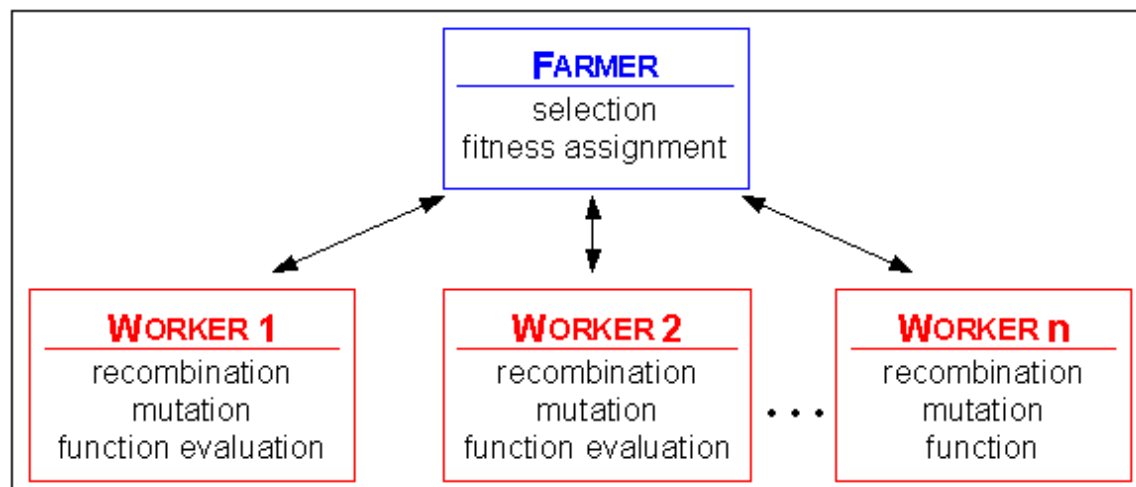
Σαφώς, τέτοιες εφαρμογές κρατούν την υπόσχεση των χρονικών μειώσεων εκτέλεσης.

Παράλληλες Εφαρμογές Αλγορίθμων

- 1. Σφαρικό Μοντέλο – Global Model
- 2. Μοντέλου νησιού – Island Model
- 3. Μοντέλο Διάχυσης – Diffusion Model

1. Σφαιρικό Μοντέλο - Εργαζόμενος/Αγρότης

- ✓ Το σφαιρικό μοντέλο υιοθετεί τον έμφυτο παραλληλισμό των γενετικών αλγορίθμων (πληθυσμός των ατόμων).
- ✓ Μια πιθανή εφαρμογή είναι ο αλγόριθμος εργαζομένων / αγροτών.



2. Μοντέλο Νησιού - Island Model

- Διαιρεί τον πληθυσμό σε πολλά υποσύνολα του πληθυσμού.
- Αυτά τα υποσύνολα πληθυσμού εξελίσσονται ανεξάρτητα το ένα από το άλλο για ορισμένες γενεές (Χρόνος Απομόνωσης – Isolation Time).
- Μετά από το χρόνο απομόνωσης διάφορα άτομα διανέμονται μεταξύ των υποσυνόλων πληθυσμού (Μετανάστευση - Migration).

Μετανάστευση - Migration

Το μέγεθος της γενετικής ποικιλομορφίας που μπορεί να εμφανιστεί στα υποσύνολα πληθυσμού και την ανταλλαγή των πληροφοριών μεταξύ των υποσυνόλων πληθυσμού καθορίζεται από:

- Τον αριθμό ανταλλαγμένων ατόμων (Ποσοστό Μετανάστευσης – Migration Rate)
- Την μέθοδο επιλογής των ατόμων για τη μετανάστευση και
- Το σχέδιο της μετανάστευσης

Μετανάστευση - Migration

- Σημαντική επιτάχυνση στο χρόνο υπολογισμού, αλλά και μείωση του αριθμού υπολογισμού της τιμής της αντικειμενικής συνάρτησης.
- Καλύτερα αποτελέσματα, εφόσον ο αλγόριθμος βρίσκει το ολικό βέλτιστο συχνότερα ή με τους λιγότερους υπολογισμούς της συνάρτησης αξιολόγησης.
 - ✓ Ακόμη και για έναν απλό υπολογιστή, που εφαρμόζει παράλληλους αλγόριθμους κατά τρόπο τμηματικό (Ψευδό-Παράλληλοι Αλγόριθμοι).

Μετανάστευση - Migration

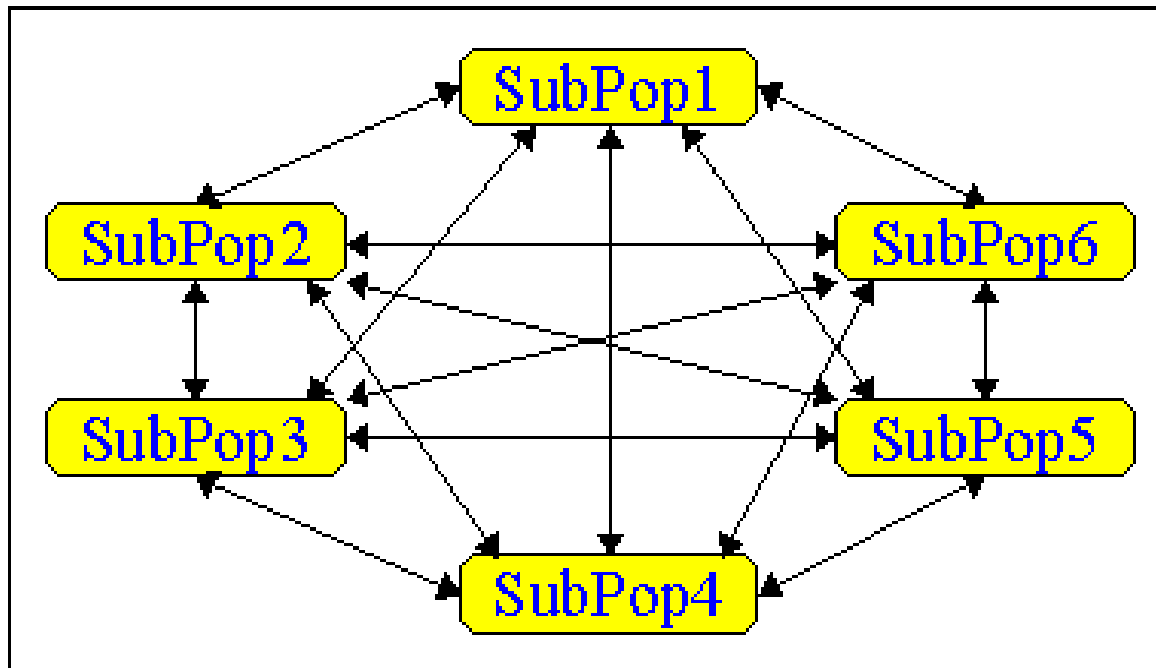
- Η επιλογή των ατόμων για τη μετανάστευση
 - Ομοιόμορφα τυχαία (επιλογή οντοτήτων για τη μετανάστευση κατά τρόπο τυχαίο),
 - Βασισμένη στην Καταλληλότητα (επιλογή των καλύτερων οντοτήτων για τη μετανάστευση).

Μετανάστευση - Migration

- Η Δομή των ατόμων για τη μετανάστευση
- Υπάρχουν πολλές δυνατότητες για τη δομή της μετανάστευσης των οντοτήτων μεταξύ των υποσυνόλων πληθυσμού:
 1. Πλήρης δικτυακή τοπολογία
 2. Τοπολογία δαχτυλιδιών
 3. Τοπολογία γειτονιάς

Μετανάστευση - Migration

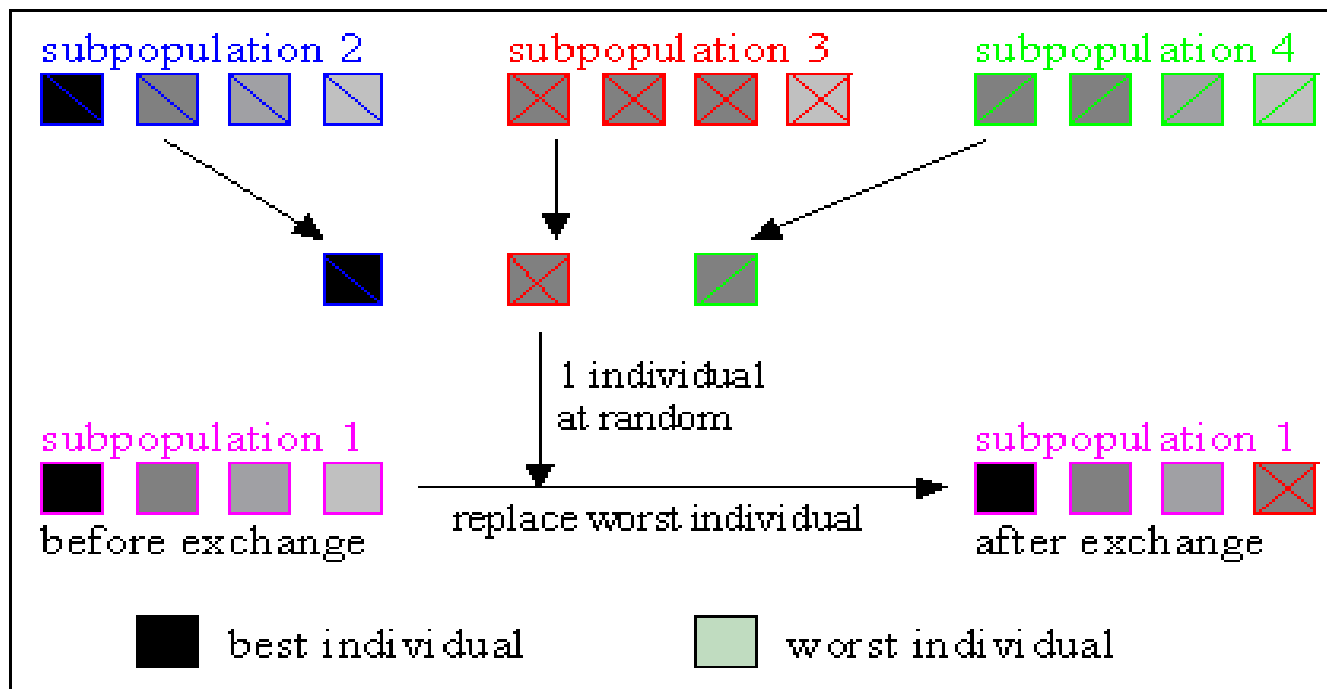
1. Πλήρης Δικτυακή Τοπολογία (Απεριόριστη τοπολογία μετανάστευσης)



✓ Τα άτομα μεταναστεύουν από οποιοδήποτε υποσύνολο πληθυσμού σε άλλο

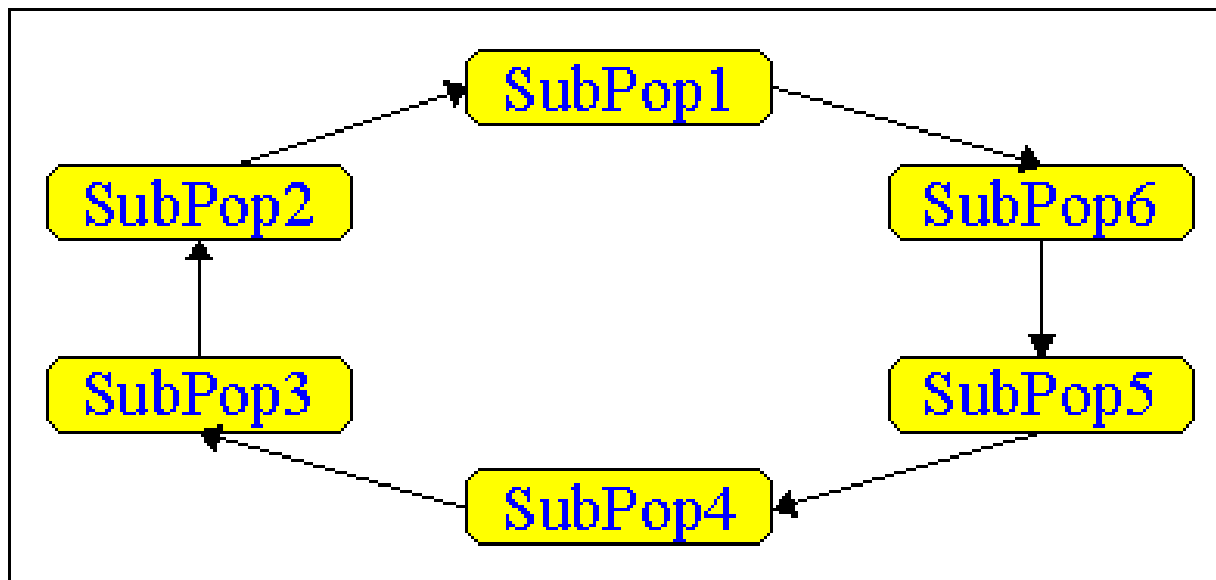
Μετανάστευση - Migration

1. Πλήρης Δικτυακή Τοπολογία (Απεριόριστη τοπολογία μετανάστευσης)



Μετανάστευση - Migration

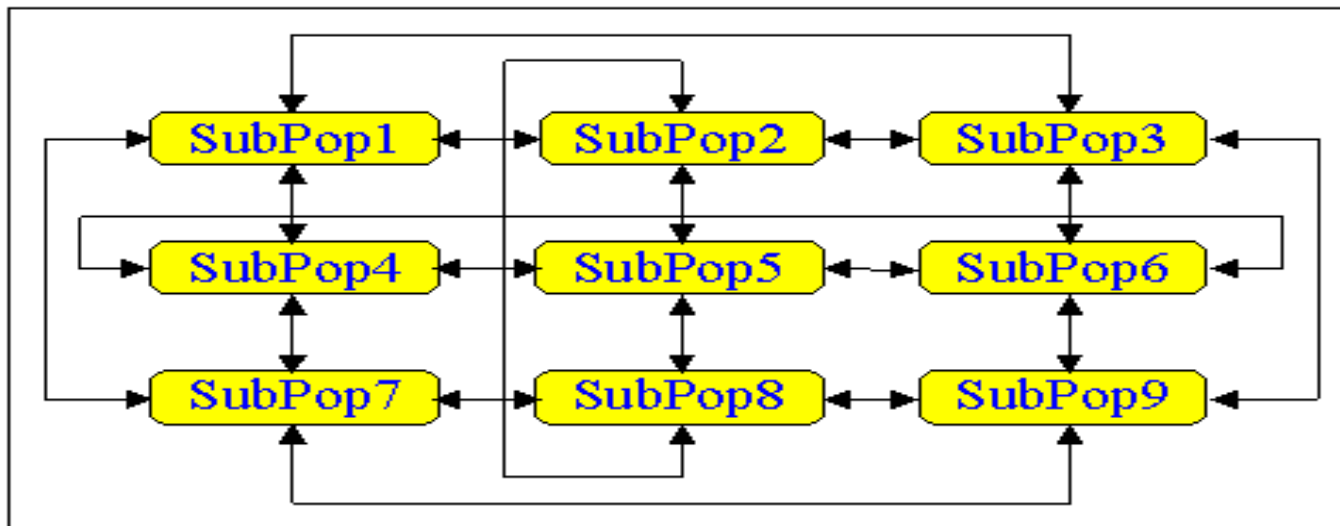
2. Τοπολογία Δαχτυλιδιών.



✓ Τα άτομα μεταφέρονται μεταξύ των κατευθυντικά παρακείμενων υποσυνόλων πληθυσμού.

Μετανάστευση - Migration

3. Τοπολογία γειτονιάς (Δακτύλιο – Torus)

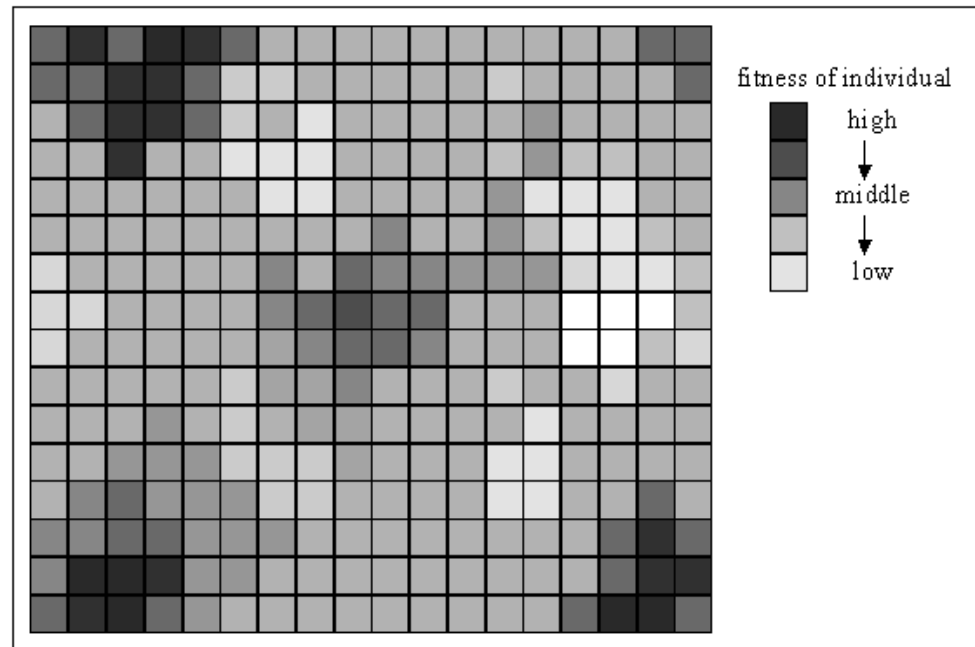


✓ Μετανάστευση μόνο μεταξύ των κοντινότερων γειτόνων, αλλά σε καθεμία κατεύθυνση μεταξύ των υποσυνόλων

3. Μοντέλο Διάχυσης

- ✓ Χειρίζεται κάθε άτομο χωριστά και επιλέγει το συνεργάτη ζευγαρώματος σε μια τοπική γειτονιά, παρόμοια με την τοπική επιλογή.
- ✓ Πραγματοποιείται μια διάχυση των πληροφοριών μέσω του πληθυσμού.

*τα εικονικά νησιά
αναζήτησης που θα
εξελιχθούν*



Γνωστά επίσης ως:
Celluar GAs ή massively-parallel GAs

- **Σύγχρονες τάσεις**
- Με μια καλύτερη κατανόηση των ομοιοτήτων και των διαφορών μεταξύ των διαφόρων υλοποιήσεων των ΕΑ, έχει αρχίσει να επικεντρώνεται σε γενικευμένα αποτελέσματα που αρχικά παρουσιάζονται μόνο για συγκεκριμένους ΕΑ.
- Παραδείγματος χάριν, οι Grefenstette και Baker (1989) διευκρινίζουν ότι πολλά χαρακτηριστικά γνωρίσματα των ΕΑ δεν αλλάζουν όταν τεθούν ορισμένες ιδιότητες στους τελεστές Επιλογής και Διαβάθμισης.
- Δείχνουν επίσης, πότε τα χαρακτηριστικά γνωρίσματα αλλάζουν, εάν οι ιδιότητες δεν συναντιούνται.
- Αν και αυτή η εργασία είναι προκαταρκτική, βοηθά στο να εξηγήσει γιατί τα αποτελέσματα μιας ευρείας ποικιλίας ΕΑ επιτυγχάνονται.
- Ο Back ερευνά επίσης τις διαφορές μεταξύ των ΓΑ και των ΕΣ και προσπαθεί να συγχωνεύσει τα καλύτερα χαρακτηριστικά γνωρίσματα κάθε ενός, προκειμένου να έχει έναν πιο εύρωστο ΕΑ.

- Δεδομένου ότι καταλαβαίνουμε καλύτερα τις δυνάμεις και τις αδυναμίες των τρεχόντων εξελικτικών μοντέλων, είναι επίσης σημαντικό να επισκεφτούμε ξανά τη βιολογική και εξελικτική βιβλιογραφία για νέες ιδέες και εμπνεύσεις.
- Ο Booker (1992) έχει πρόσφατα επισημάνει τις συνδέσεις της θεωρίας Διασταύρωσης των ΓΑ με τη γενικότερη θεωρία κατανομής του Γενετικού Συνδυασμού πληθυσμών.
- Ο Mühlenbein (1993) έχει επικεντρωθεί στους ΕΑ, που διαμορφώνονται μετά από πρακτικές αναπαραγωγής.
- Στην κοινότητα του ΕΠ, ο Atmar (1992) δίνει έμφαση σε μερικά κοινά λάθη για την εξελικτική θεωρία και τους ΕΑ.
- Τα τελευταία χρόνια το ενδιαφέρον των επιστημόνων έχει προσανατολιστεί στην εύρεση αποδοτικών μεθόδων αποθήκευσης και ανάκτησης πληροφοριών.

- Οι ΕΑ έχουν χρησιμοποιηθεί στην ερμηνεία και την αποθήκευση χημικών δομών και στην ανάκτηση από βάσεις δεδομένων μορίων που περιέχουν συγκεκριμένες δομές (Jones κ.α. 1993).
- Η ανάκτηση εγγράφων τα οποία χαρακτηρίζονται από συγκεκριμένα γνωρίσματα γίνεται όλο και περισσότερο σημαντική καθώς όλο και μεγαλύτερος είναι ο όγκος των πληροφοριών που ανταλλάσσονται μέσω υπολογιστών.
- Εργαλεία για την ανάκτηση εγγράφων τα οποία περιέχουν συγκεκριμένες λέξεις έχουν κατασκευαστεί εδώ και πολλά χρόνια.
- Όμως τα εργαλεία αυτά υπόκεινται σε διάφορους περιορισμούς, κυρίως όταν κάποιος επιθυμεί να κατασκευάσει την κατάλληλη ερώτηση αναζήτησης, με αποτέλεσμα να μην ανταποκρίνονται στην πλειοψηφία των απαιτήσεων των χρηστών.
- Οι ερευνητές τα τελευταία χρόνια έχουν ασχοληθεί με τη χρήση ΕΑ για την κατασκευή ερωτήσεων αναζήτησης (Yang και Korfhage 1993).

Γενετικοί Αλγόριθμοι και Αριθμητική Βελτιστοποίηση

AN INTRODUCTION TO GENETIC ALGORITHMS
FOR NUMERICAL OPTIMIZATION¹

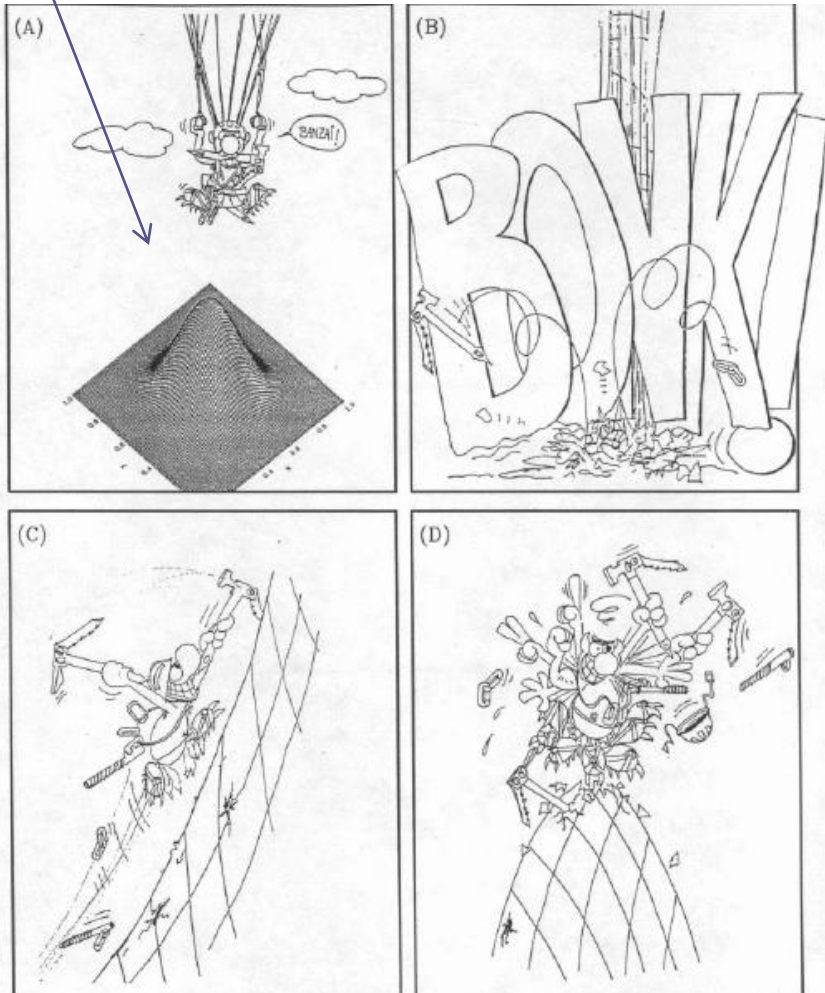
Paul Charbonneau

High Altitude Observatory, NCAR, P.O. Box 3000, Boulder, CO 80307, USA

paulchar@hao.ucar.edu

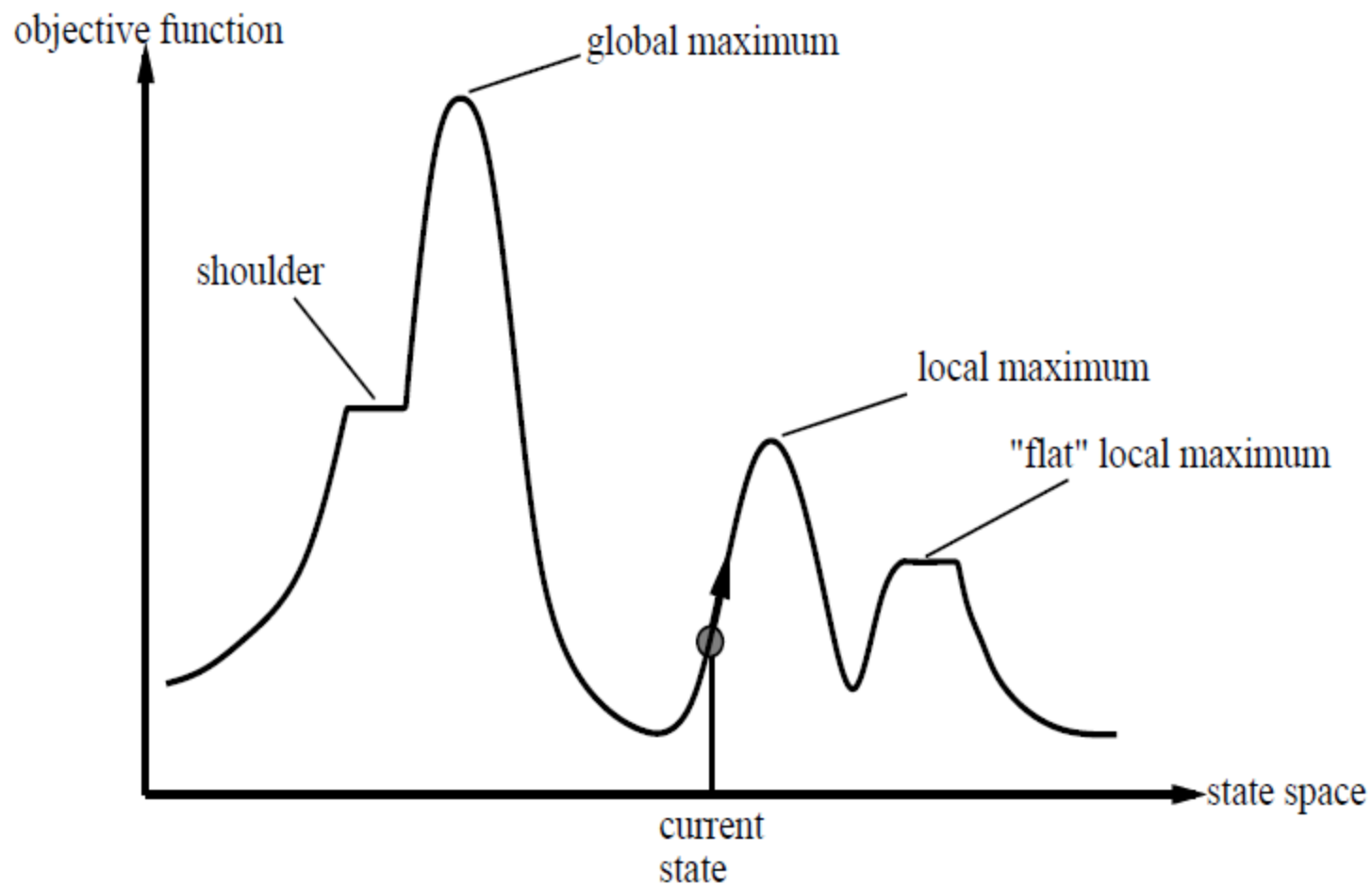
Αναρρίχηση λόφων (hill climbing)

Μία κορυφή

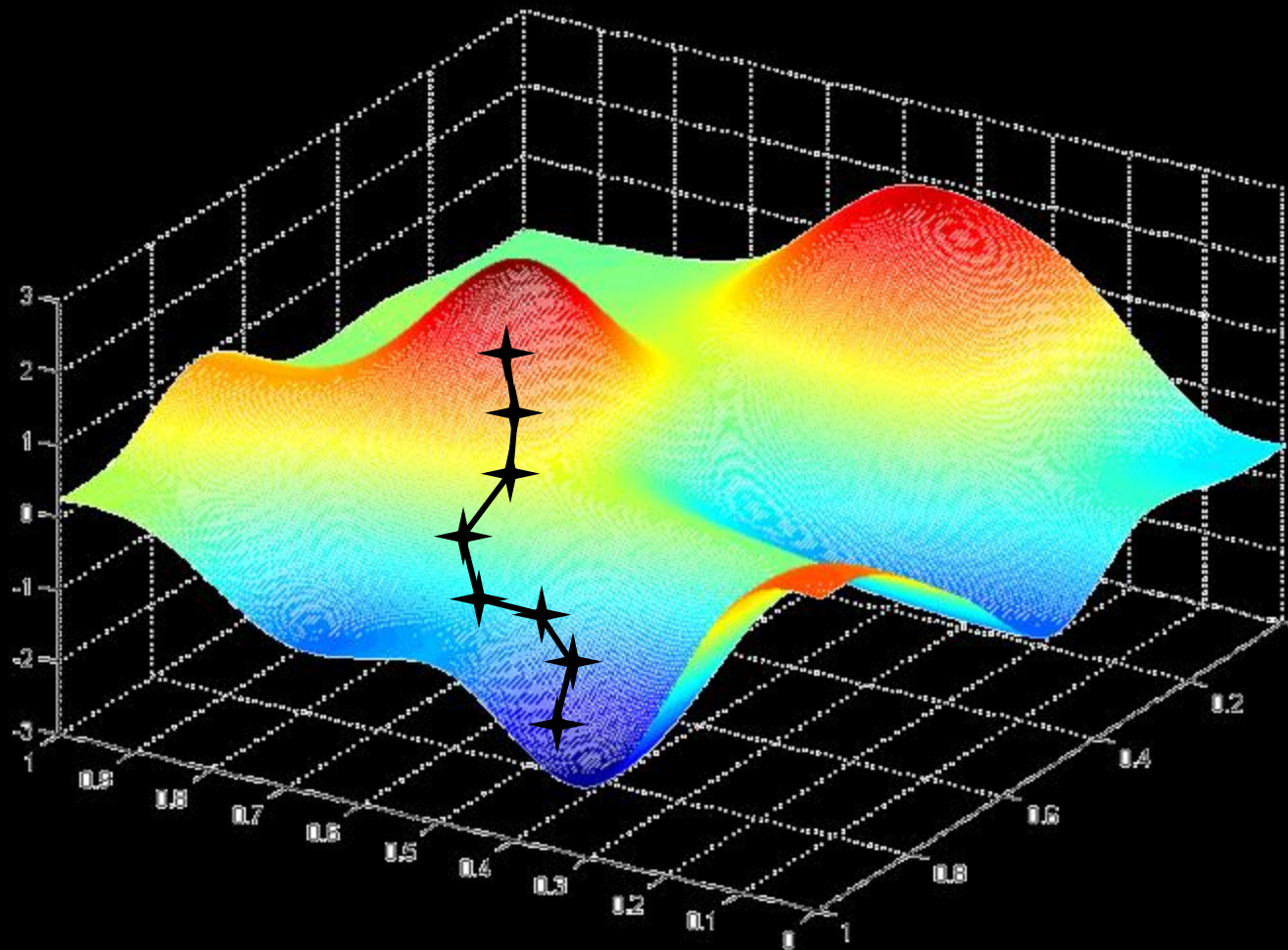


- Περισσότερες μέθοδοι βελτιστοποίησης που υπάρχουν στα βιβλία, λειτουργουν κατ'αυτόν τον τρόπο και διαφέρουν στο:
1. πώς καθορίζουν την κατεύθυνση ανάβασης,
 2. πώς επιλέγουν το μέγεθος του βήματος που θα γίνει σε αυτή την κατεύθυνση
 3. αν χρησιμοποιούν ή όχι πληροφορία κλίσης στα προηγούμενα βήματα.

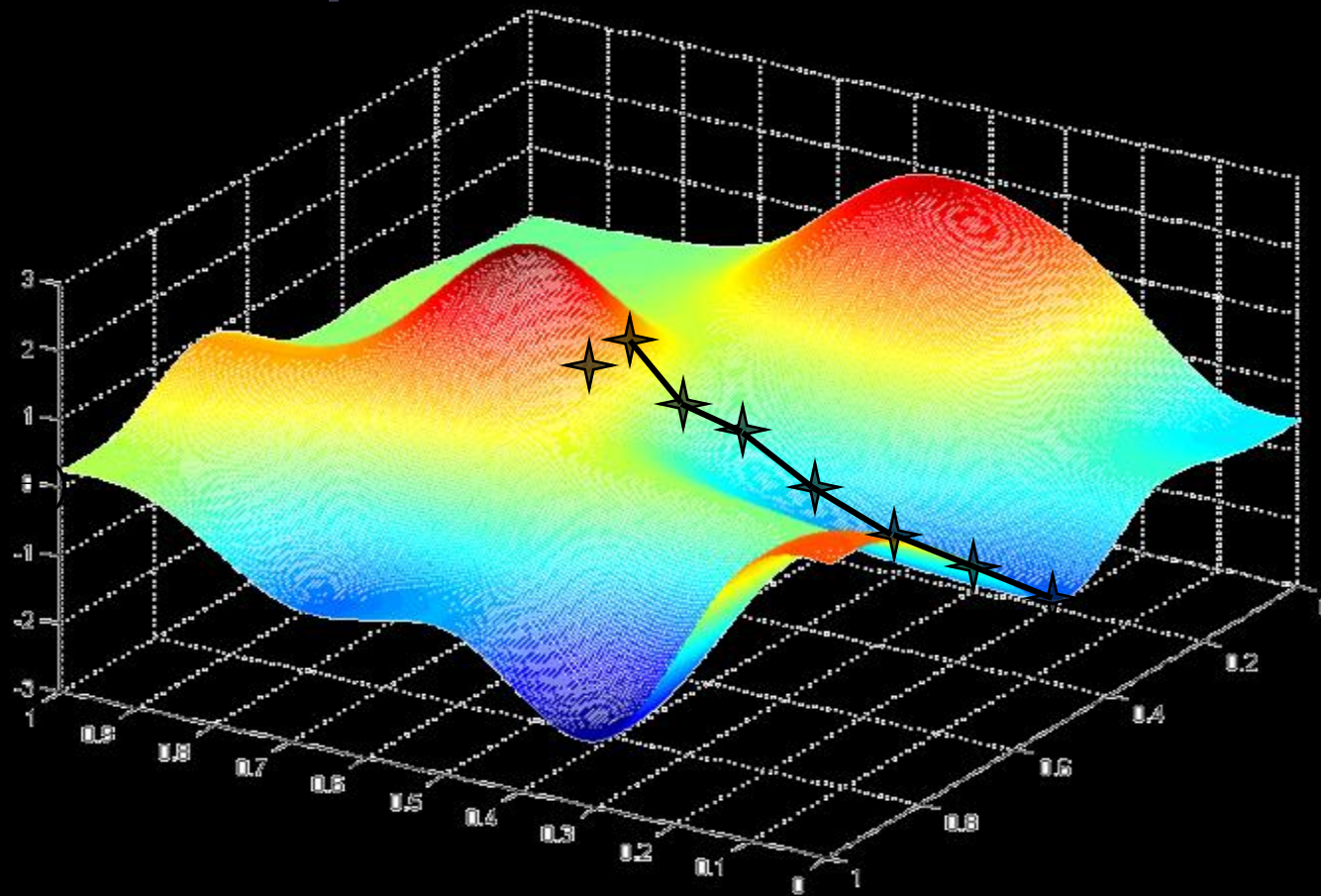
Χώρος καταστάσεων



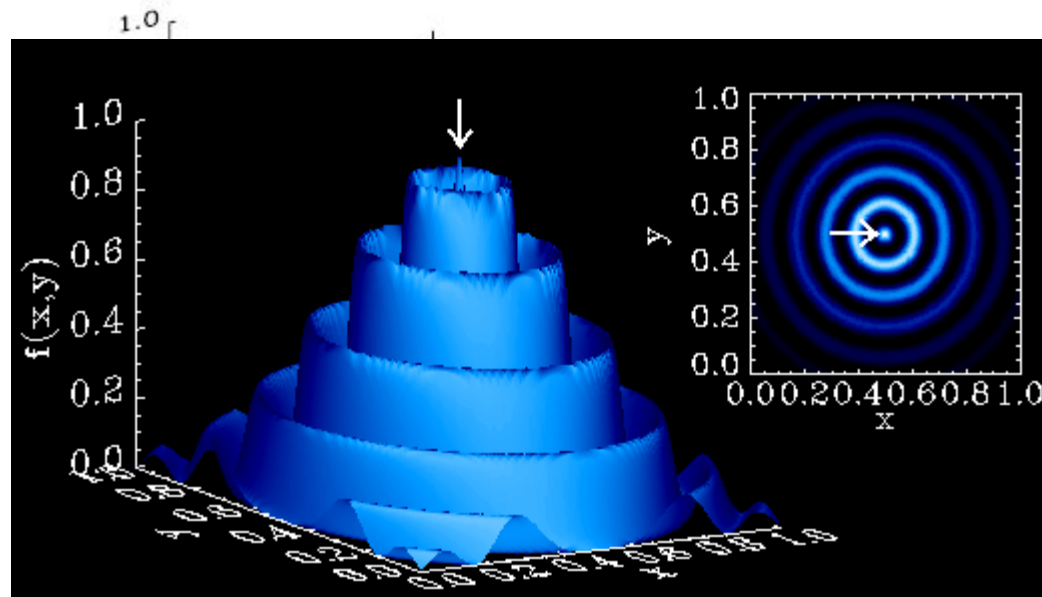
Local Optima -1



Local Optima -2

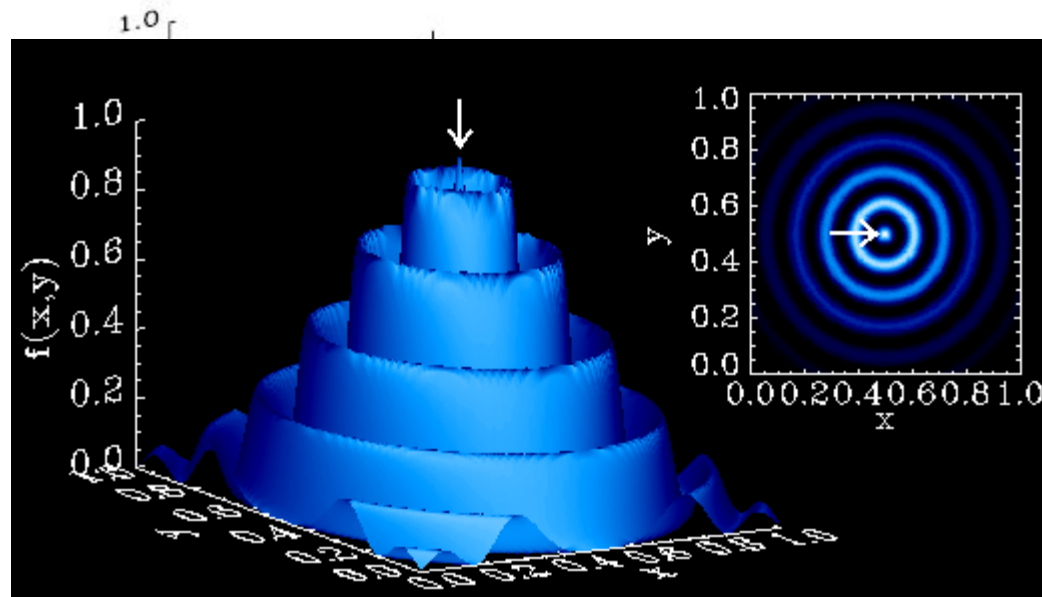


Αναρρίχηση λόφων (hill climbing)



- Αν η κεντρική κορυφή καλύπτει περίπου το 1% του χώρου αναζήτησης ➡ τι πιθανότητα έχει η μέθοδος να δουλέψει?

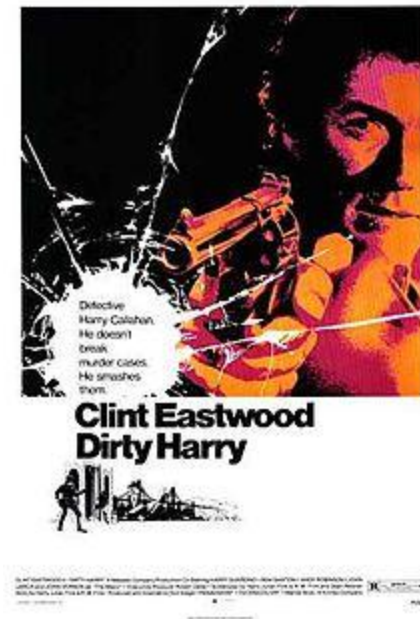
Αναρρίχηση λόφων (hill climbing)



- Αν η κεντρική κορυφή καλύπτει περίπου το 1% του χώρου αναζήτησης ➡ 1% πιθανότητα να πέσει σε σημείο ώστε να δουλέψει η μέθοδος hill climbing. Νοιώθετε τυχεροί?

First Rule of Global Optimization

- THE DIRTY HARRY RULE
 - “You should never feel lucky!”



Επαναληπτική Αναρρίχηση Λόφων- Iterated Hill Climbing



Επαναληπτική Αναρρίχηση Λόφων

- Κάλυψη του χώρου κατά 1% από την κεντρική κορυφή 
Αναμένεται να χρειαστούν κατά μέσο όρο 10^2 τρεξίματα πριν βρεθεί η κεντρική κορυφή.
- Σε προβλήματα βελτιστοποιήσεις υψηλών διαστάσεων ή σε πρόβληματα όπου το ολικό βέλτιστο καλύπτει έναν πολύ μικρό χώρο του συνολικού χώρου παραμέτρων ο αριθμός των απαιτούμενων επαναλήψεων αυξάνει ραγδαία...

Second Rule of Global Optimization

- THE NO FREE LUNCH RULE
- “If you really want the global optimum you will have to work for it”

Απόδοση μεθόδου βελτιστοποίησης

- Τρεις διαφορετικές όψεις της έννοιας της απόδοσης:
 1. Απόλυτη Απόδοση (Absolute Performance)
 2. Ολική Απόδοση (Global Performance)
 3. Σχετική Απόδοση (Relative Performance)

Απόδοση μεθόδου βελτιστοποίησης

- Οι περισσότερες μέθοδοι προσπαθούν να απαντήσουν ταυτόχρονα στο 1. Απόλυτη Απόδοση και στο 3. Σχετική Απόδοση
- Για να απαντήσουν και στο 2. Ολική Απόδοση απαιτείται μια καλή αρχική εκτίμηση.

Αν έχετε πάντα μια καλή εκτίμηση αρκεί μια απλή μέθοδος βελτιστοποίησης ...και δεν χρειάζεται να παρακολουθείτε άλλο ...

Απόδοση μεθόδου βελτιστοποίησης

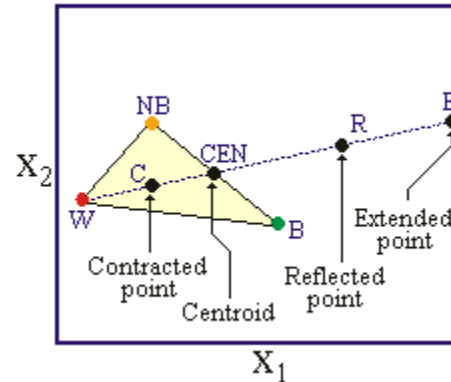
- Οι περισσότερες μέθοδοι προσπαθούν να απαντήσουν ταυτόχρονα στο 1. Απόλυτη Απόδοση και στο 3. Σχετική Απόδοση
- Για να απαντήσουν και στο 2. Ολική Απόδοση απαιτείται μια καλή αρχική εκτίμηση.

Αν έχετε πάντα μια καλή εκτίμηση αρκεί μια απλή μέθοδος βελτιστοποίησης ...και δεν χρειάζεται να παρακολουθείτε άλλο ...

But rest assured that Dirty Harry will catch up with you one of these days...

Μέθοδος Simplex

http://www.chem.uoa.gr/applets/AppletSimplex/Apl_Simplex1.html



Υπολογίζεται η θέση του κεντροειδούς (centroid) CEN στο ενδιάμεσο των σημείων B και NB.

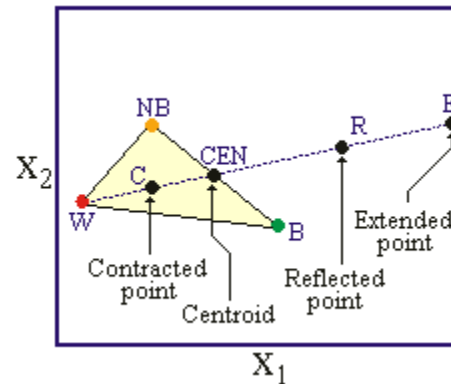
Πραγματοποιείται ανάκλαση (reflection) του χειρότερου σημείου W ως προς το CEN και μετρείται η απόκριση του νέου σημείου R.

Εάν το R βρίσκεται μέσα στην εξεταζόμενη περιοχή και η απόκρισή του RR είναι καλύτερη από την RW, τότε σχηματίζεται ένα νέο simplex με αντικατάσταση του W με το σημείο R. Έτσι έχουμε μια "κίνηση" του simplex. Η διαδικασία επαναλαμβάνεται από το βήμα <1> με το νέο simplex.



Μέθοδος Simplex

http://www.chem.uoa.gr/applets/AppletSimplex/Apl_Simplex1.html



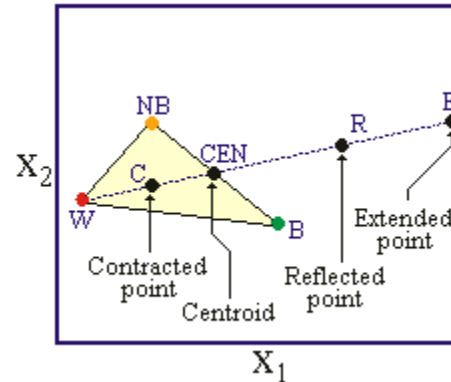
Εάν η απόκριση RR είναι ακόμη καλύτερη, δηλαδή καλύτερη και από την RB , αυτό αποτελεί ένδειξη ότι η κίνηση γίνεται προς τη σωστή κατεύθυνση και δοκιμάζεται επέκταση (extension) προς το σημείο E , που απέχει δυο φορές περισσότερο από το CEN απ' ό,τι το R .

Εάν το E βρίσκεται μέσα στην εξεταζόμενη περιοχή και η απόκρισή του RE είναι καλύτερη από εκείνη του RR , τότε το W αντικαθίσταται από το E , εάν όχι αντικαθίσταται από το R . Η διαδικασία επαναλαμβάνεται από το βήμα <1> με το νέο simplex.



Μέθοδος Simplex

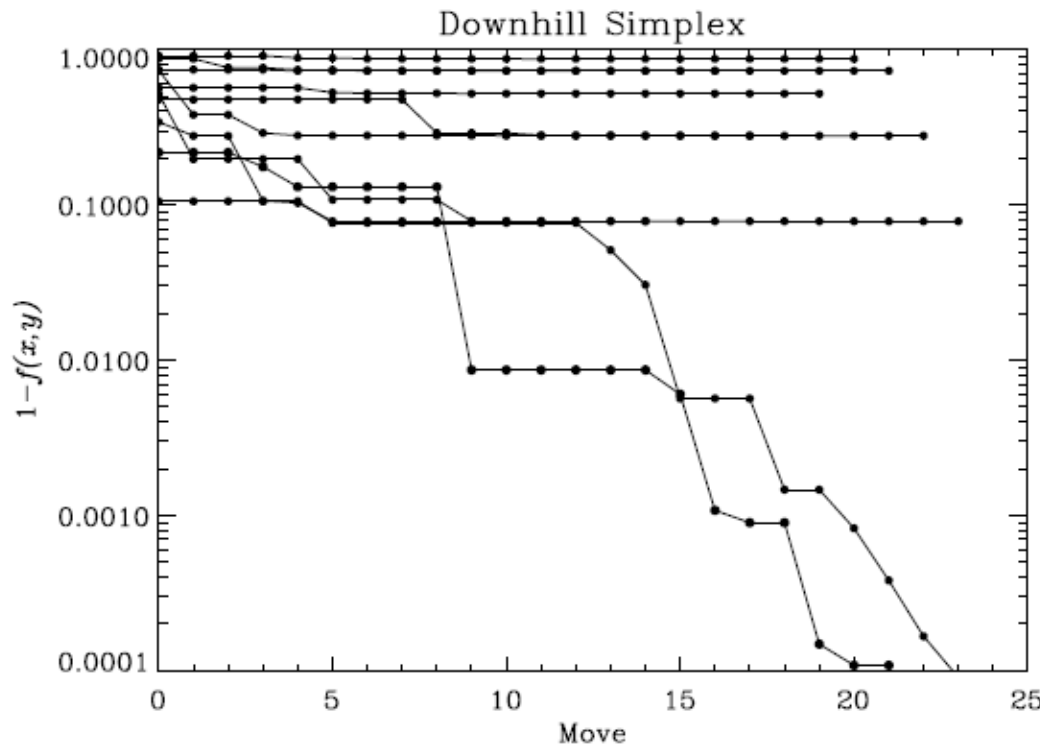
http://www.chem.uoa.gr/applets/AppletSimplex/Appl_Simplex1.html



Εάν η αρχική ανάκλαση αποτύχει, δηλ. αν το R βρίσκεται εκτός της εξεταζόμενης περιοχής ή η απόκρισή του RR είναι χειρότερη της RW , τότε πραγματοποιείται σύμπτυξη (contraction) προς το σημείο C. Το C βρίσκεται ενδιάμεσα των W και CEN. Η διαδικασία επαναλαμβάνεται από το βήμα <1> με το νέο simplex.



Απόδοση της Μεθόδου Simplex



Ολική Απόδοση (Global Performance)
μόνο 2%



Επαναληπτική Simplex



✓ Στις 500 επαναλήψεις το ολικό
βέλτιστο βρίσκεται τις 99.5% φορές!
✓ Όμως : Απαιτούνται περίπου 10^4
υπολογισμοί της συνάρτησης για κάθε
επανάληψη



Welcome back to the no free lunch rule...

Προβλήματα - P1

1.4.1 P1: maximizing a function of two variables [2 parameters]

Our first test problem (hereafter labeled “P1”) is our now familiar 2-D landscape of Figure 2. Mathematically, it is defined as

$$f(x, y) = \cos^2(n\pi r) \exp(-r^2/\sigma^2) , \quad (1a)$$

$$r^2 = (x - 0.5)^2 + (y - 0.5)^2, \quad x, y \in [0, 1] , \quad (1b)$$

where $n = 9$ and $\sigma^2 = 0.15$ are constants. The global maximum is located at $(x, y) = (0.5, 0.5)$, with $f(x, y) = 1.0$. This global maximum is surrounded by concentric rings of secondary maxima, centered on the global maximum at radial distances

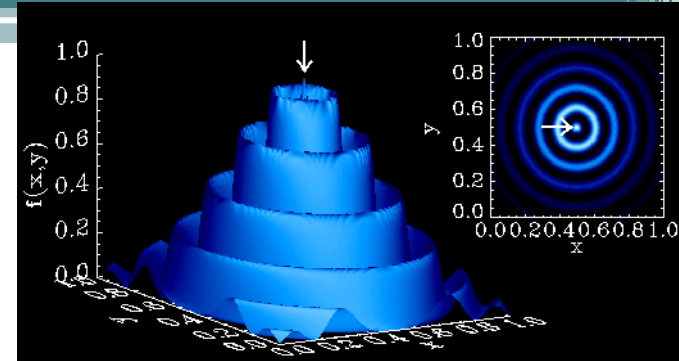
$$r_{\max} = \{0.110192, 0.220385, 0.330582, 0.440782, 0.550986\} . \quad (2)$$

Between these are located another series of concentric rings corresponding to minima:

$$r_{\min} = \frac{m - 1/2}{n}, \quad m = 1, \dots, 6 \quad (3)$$

The error ε associated with a given solution (x, y) can be defined as

$$\varepsilon = 1 - f(x, y) . \quad (4)$$



Προβλήματα - P2

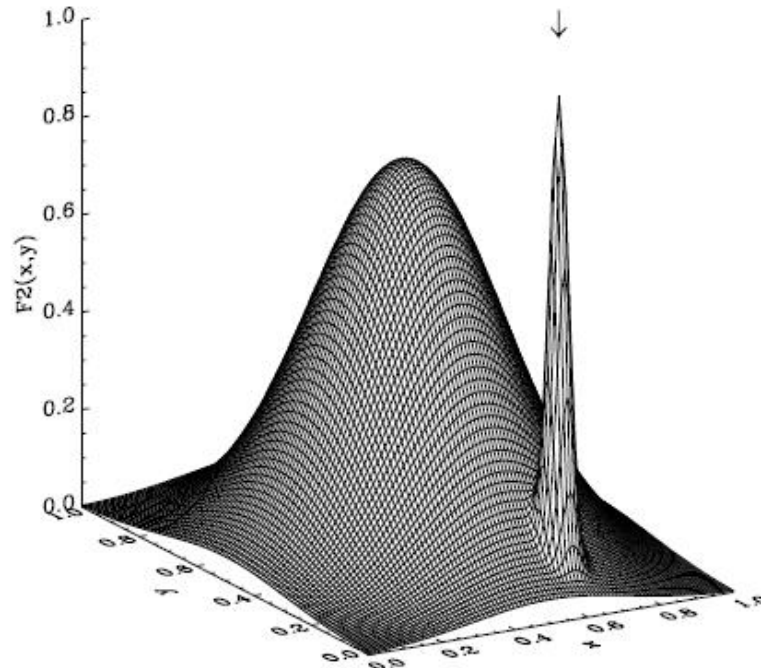


Figure 5: Test problem P2. The problem consist in maximizing a function of two variables, defined by two Gaussians (see eqs. [5]). The global maximum is $f(x, y) = 1$ at $(x, y) = (0.6, 0.1)$, and is indicated by the arrow.

Προβλήματα - P3

1.4.3 P3: maximizing a function of four variables [4 parameters]

Test problem P3 is a direct generalization of P1 to four independent variables (w, x, y, z) :

$$f(w, x, y, z) = \cos^2(n\pi r) \exp(-r^2/\sigma^2) , \quad (6a)$$

$$r^2 = (w - 0.5)^2 + (x - 0.5)^2 + (y - 0.5)^2 + (z - 0.5)^2 , \quad w, x, y, z \in [0, 1] , \quad (6b)$$

again with $n = 9$ and $\sigma^2 = 0.15$. Comparing performance on P1 and P3 will provide a measure of *scalability* of the method under consideration, namely how performance degrades as parameter space dimensionality is increased, everything else being equal. P3 is a *very hard* global optimization problem; the simplex method manages to find the global maximum only 6 times out of 10^5 trials.

Προβλήματα – P4

1.4.4 P4: Minimizing a least squares residual [6 parameters]

Our fourth and final test problem is defined as a “real” nonlinear least squares fitting problem. Consider a function of one variable (x) defined as the sum of two Gaussians:

$$y(x) = \sum_{j=1}^2 A_j \exp \left(-\frac{(x - x_j)^2}{\sigma_j^2} \right) . \quad (7)$$

Define now a “dataset” by evaluating this function for a set of K equidistant values of x_k in the interval $[0, 1]$, i.e., $y_k^* \equiv y(x_k)$, $x_{k+1} - x_k = \Delta x$, for some set values of A_1, x_1 , etc. Given that dataset and the functional form used to generate it (i.e., eq. [7]), the optimization problem is then to recover the parameter values for A_1, x_1 , etc originally used to produce the dataset. This is done by minimizing the square residual

$$R(A_1, x_1, \sigma_1, A_2, x_2, \sigma_2) = \sum_{k=1}^K [y^* - y(x_k; A_1, x_1, \sigma_1, A_2, x_2, \sigma_2)]^2 , \quad (8)$$

with respect to the 6 parameters defining the two Gaussians. If one is told *a priori* that two Gaussians are to be fit to the data, then this residual minimization problem is obviously equivalent to a 6-D function maximization problem for $1/R$ (say), which simply defines a function in 6-D space. Figure 6 shows the dataset generated using the parameter set

$$[A_1, x_1, \sigma_1, A_2, x_2, \sigma_2] = [0.9, 0.3, 0.1, 0.3, 0.8, 0.025] \quad (9)$$

Προβλήματα - P4

12

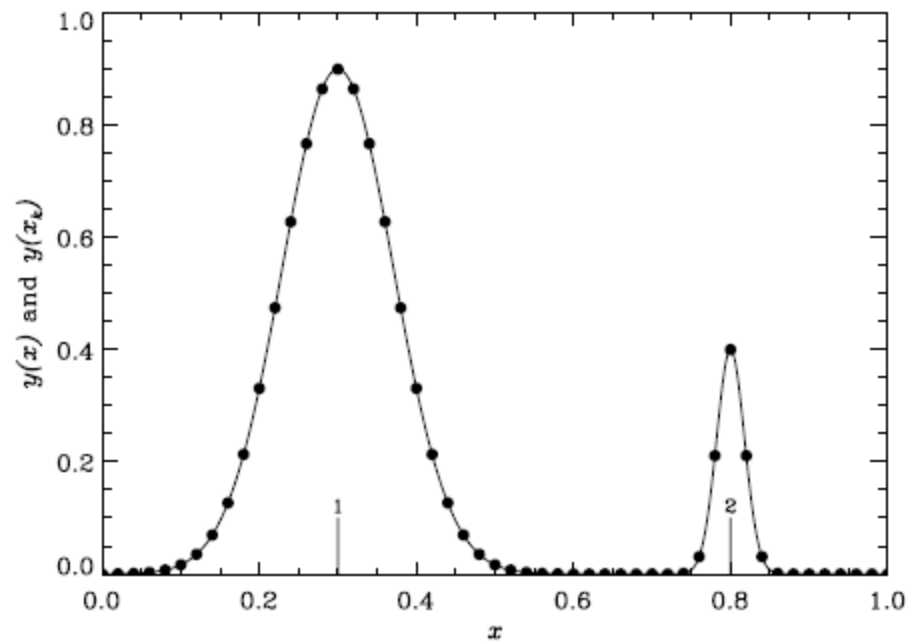


Figure 6: Test problem P4. This is a 6-parameter problem which consists in fitting two Gaussians to a “dataset” of 51 points. Note how the second Gaussian is poorly sampled by the discretization in x . The thin solid line is the underlying 2-Gaussian function defined by eq. (7).

Απόδοση της Simplex στα προβλήματα

Test Problem	Performance	Simplex	Iterated simplex
P1	$\langle 1 - f \rangle$	0.633	0.00793
	p_G	0.0213	0.898
	$\langle N_f \rangle$	37	3872
	N_t	1	100
P2	$\langle 1 - f \rangle$	0.194	0.0619
	p_G	0.0263	0.931
	$\langle N_f \rangle$	44	4412
	N_t	1	100
P3	$\langle 1 - f \rangle$	0.413	0.07713
	p_G	0.00006	0.016
	$\langle N_f \rangle$	70	35252
	N_t	1	500
P4	$\langle R \rangle$	0.332	0.0069
	p_G	0.123	0.941
	$\langle N_f \rangle$	753	37638
	N_t	1	20

Absolute performance:

Μέσος όρος της τιμής $1-f(x)$ για όλα τα τρεξίματα

Global Performance:

Ποσοστό για το οποίο βρέθηκε το πραγματικό ολικό μέγιστο (Για P1, P2, P3 $f \geq 0.95$ για P4 $R \leq 0.1$)

Relative Performance

Μέσος όρος αριθμού αποτιμήσεων της συνάρτησης/μοντέλου

Αριθμός επαναλήψεων της simplex

Εκ ορισμού 1 για την απλή simplex

Απόδοση της Simplex στα προβλήματα

- Κακή απόδοση της μεθόδου Simplex 
πράγματι δύσκολα προβλήματα
βελτιστοποίησης
- Επαναληπτική Simplex πολύ καλύτερη
απόδοση, αλλά και πολύ μεγαλύτερο
υπολογιστικό κόστος
- Ο λόγος που το υπολογιστικό κόστος είναι τόσο
μεγάλο είναι ότι η κάθε προσπάθεια εύρεσης
λύσης είναι ανεξάρτητη από τις άλλες και δεν
υπάρχει ανταλλαγή πληροφορίας  Γ. Α.

Η δύναμη της (cumulative)επιλογής

- JEG SNAKKER BARE LITT NORSK
- 27 χαρακτήρες (με τα κενά) από αλφάβητο 30 χαρακτήρων συμπεριλαμβανόμενου του κενού χαρακτήρα και των σκανδιναβικών χαρακτήρων Å, Ø, Æ

Η δύναμη της (cumulative)επιλογής

- Τυχαία πρόταση με 27 χαρακτήρες
- GE YTÅUMNBGH JHØA QMWCXNESØ
- Δεν μοιάζει πολύ με την αρχική, όμως 2 χαρακτήρες συμπίπτουν
- Ο συνολικός αριθμός των προτάσεων με 27 χαρακτήρες με αλφάβητο 30 χαρακτήρων: $30^{27} \approx 7.63 \cdot 10^{39}$  Η πιθανότητα να βρεθεί τυχαία την πρώτη φορά η πρόταση: $30^{-27} \approx 10^{-40}$

Αστρονομικά μικρή πιθανότητα!

Η δύναμη της (cumulative)επιλογής

- Αντί για τυχαία αναζήτηση αλγόριθμος:

Ουσιαστικά πρόκειται για μια μορφή αναρρίχησης λόφου!

- (1) Generate 10 sentences of 27 randomly chosen characters;
- (2) Select the sentence that has the most correct letters;
- (3) Duplicate this best sentence ten times;
- (4) For each such duplicate, randomly replace a few letters¹⁰;
- (5) Repeat steps (2) through (4) until the target sentence has been matched.

(Ντετερμινιστική) Φυσική Επιλογή

Κληρονομικότητα (Ακραία μορφή)

Στοχαστική Διαδικασία που επιτρέπει την
ποικιλομορφία (Variability)

- ✓ Ελάχιστη “fitness” πληροφορία: Γνωρίζουμε μόνο τον αριθμό των λάθος χαρακτήρων, όχι ποιοι χαρακτήρες είναι λάθος
- ✓ Δεν υπάρχει ανταλλαγή πληροφορίας ακόμη...

Η δύναμη της (cumulative)επιλογής

Iteration	Target	J E G	S N A K K E R	B A R E	L I T T	N O R S K	
$p_m=0.01$	1	Z E B Y E N A T U V P	Q Å O D E M I	F V G H D O O	23		
50		V E G Æ E N Æ R O E O	Q Å B D E M I	F V N Å D O K	19		
100		V E G Æ E N Æ K C E O	O P H Z E M I	F V N Å Ø O K	18		
150		V E G W X N Æ K C E O	N A H A D M I	C F N N E R O K	16		
200		J E G W X N P K K E O	B A H A	R I C E Å N E R O K	12		
250		J E G W V N R K K E	B A E A	R I Å E Å N	12		
300		J E G R N R K K E	B A E T	U I Ø Ø	N Q R K K	10	
350		J E G K N K K K E	B A R	U I Ø Ø	N Q R M K	9	
400		J E G K N V K K E	B A R	P I H Ø	N Q R S K	8	
450		J E G K N V K K E R	B A R	L I D Ø	N Å R S K	6	
500		J E G Ø N V K K E R	B A R	L I S Ø	N K R S K	6	
550		J E G Ø N F K K E R	B A R E	L I S I	N B R S K	5	
600		J E G S N F K K E R	B A R E	L I A W	N B R S K	4	
650		J E G S N A K K E R	B A R E	L I A W	N O R S K	2	
700		J E G S N A K K E R	B A R E	L I A T	N O R S K	1	
750		J E G S N A K K E R	B A R E	L I A T	N O R S K	1	
800		J E G S N A K K E R	B A R E	L I A T	N O R S K	1	
850		J E G S N A K K E R	B A R E	L I A T	N O R S K	1	
900		J E G S N A K K E R	B A R E	L I Y T	N O R S K	1	
950		J E G S N A K K E R	B A R E	L I T T	N O R S K	0	

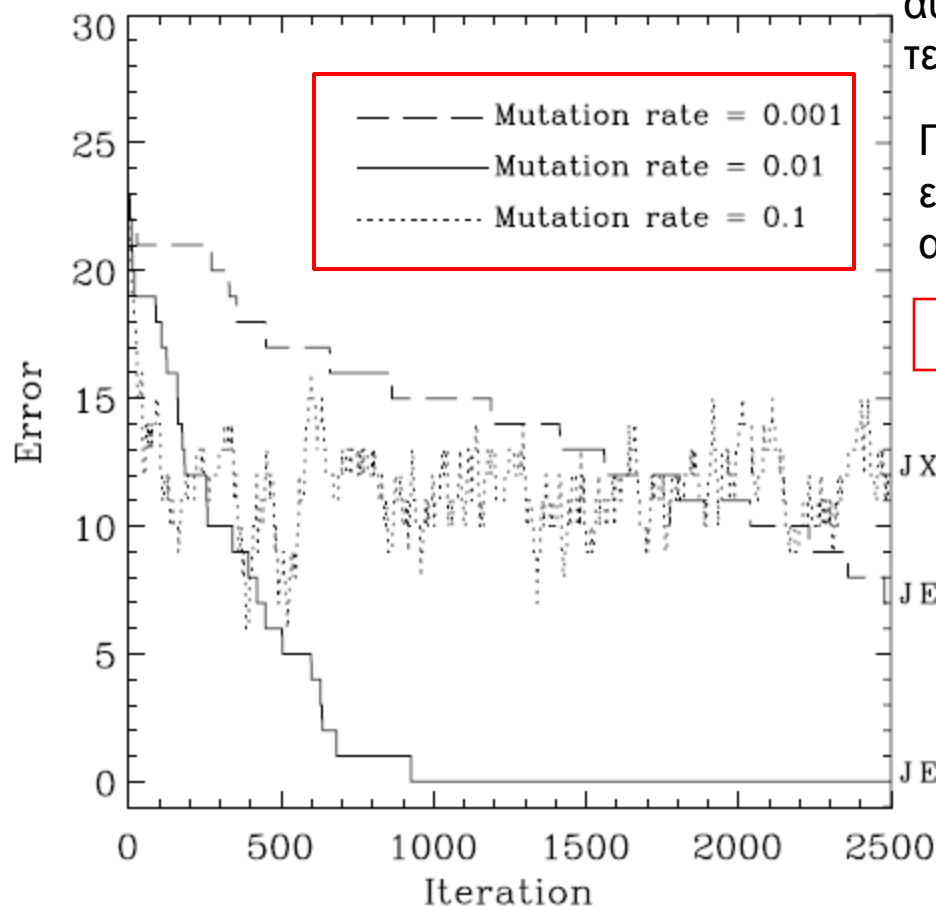
Λάθος: Αριθμός των λάθος χαρακτήρων

Λάθος πέφτει αρχικά ραγδαία μετά όμως πιο αργά αργότερα (Όσες επαναλήψεις χρειάζονται για να βρεθούν σωστά οι πρώτοι 20 χαρακτήρες απαιτούνται και για να βρεθεί τον τελευταίο χαρακτήρα)

Απαιτούνται 918 επαναλήψεις στις οποίες δημιουργούνται 9180 προτάσεις που συγκρίνονται με την ζητούμενη πρόταση.

Απείρως λιγότερες από τις $\approx 10^{-40}$ που απαιτούνται στην τυχαία αναζήτηση!

Η δύναμη της (cumulative)επιλογής



Πολύ μεγάλο p_m : τοποθετεί γρήγορα χαρακτήρες σε σειρά, αλλά αλλάζει και αυτούς που ήδη ήταν σε σωστή σειρά → τείνει στην τυχαία αναζήτηση

Πολύ μικρό p_m : δεν επιτρέπει την εξερεύνηση όλου του χώρου αναζήτησης

Η μετάλλαξη είναι δίκοπο μαχαίρι!

J X O N N A K K K R M B A R E R L I L H A N P L B K

J E G A S N N K K I R B A R E L I T A O N H S K

J E G S N A K K E R B A R E L I T T N O R S K

Γενετικός αλγόριθμος για τα προβλήματα βελτιστοποίησης

- (1) Randomly initialize population and evaluate fitness of its members;
- (2) Breed selected members of current population to produce offspring population (selection based on fitness);
- (3) Replace current population by offspring population;
- (4) Evaluate fitness of new population members;
- (5) Repeat steps (2) through (4) until the fittest member of the current population is deemed fit enough.

Αν δεν επαναλαμβανόταν για ολόκληρο πληθυσμό, αλλά για μια λύση θα θύμιζε πολύ....?!

Γενετικός αλγόριθμος για τα προβλήματα βελτιστοποίησης

- (1) Randomly initialize population and evaluate fitness of its members;
- (2) Breed selected members of current population to produce offspring population (selection based on fitness);
- (3) Replace current population by offspring population;
- (4) Evaluate fitness of new population members;
- (5) Repeat steps (2) through (4) until the fittest member of the current population is deemed fit enough.

Αν δεν επαναλαμβανόταν για ολόκληρο πληθυσμό, αλλά για μια λύση θα θύμιζε πολύ την επαναληπτική αναρρίχηση λόφων με εξαίρεση το βήμα 2 της αναπαραγωγής!

Γενετικός αλγόριθμος για τα προβλήματα βελτιστοποίησης

Encoding:	P(P1)	x=0.14429628	y=0.72317247	[01]
	P(P2)	x=0.71281369	y=0.83459991	[02]
		↓	↓	
		71281369	83459991	[03]
	S(P2)	7128136983459991		[04]
Breeding:	S(P1)	1442962872317247		[05]
	S(P2)	7128136983459991		[06]
(a) Crossover (gene=4):				
		144	2962872317247	[07]
			↓↑	
		712	8136983459991	[08]
		144	8136983459991	[09]
		712	2962872317247	[10]
	S(O1)	1448136983459991		[11]
	S(O2)	7122962872317247		[12]
(b) Mutation (Offspring=02, gene=10):				
	S(O2)	7122962872317247		[13]
		712296287	2 317247	[14]
		712296287	8 317247	[15]
	S(O2)	7122962878317247		[16]
Decoding:	S(O2)	7122962878317247		[17]
		71229628	78317247	[18]
		↓	↓	
	P(O2)	x=0.71229628	y=0.78317247	[19]
	P(O1)	x=0.14481369	y=0.83459991	[20]

Παρατηρήσεις

- Οι απόγονοι περιέχουν κομμάτια λύσης και από τους δυο γονείς ➡ κληρονομικότητα και ανταλλαγή πληροφοριών
- Η διασταύρωση και η μετάλλαξη περιέχουν όμως και εντελώς στοχαστικές διαδικασίες όπως το σημείο διασταύρωσης, το ψηφίο μετάλλαξης, η νέα τιμή στο μεταλλαγμένο ψηφίο ➡ ποικιλομορφία που απαιτείται
- Ο τρόπος κωδικοποίησης/αποκωδικοποίησης είναι ενδεικτικός, υπάρχει πολύ διαφορετικοί τρόποι (πχ δυαδική κωδικοποίηση) ο καθένας με τα καλά και τα κακά του.

Παρατηρήσεις

- Ο τρόπος που γίνεται η κωδικοποίηση, η διασταύρωση και η μετάλλαξη επιτρέπουν την μη παραβίαση του συνόλου ορισμού των παραμέτρων. Π.χ αν οι τιμές για τις παραμέτρους είναι στο διάστημα $[0.0, 1.0]$  και οι απόγονοι θα βρίσκονται σε αυτό το διάστημα

Πολύ απλός τρόπος για να διασφαλίσει κανείς έναν περιορισμό για θετικές τιμές!

- Εφόσον η μετάλλαξη ενεργεί στην κωδικοποιημένη λύση, ανάλογα με το ποιο ψηφίο θα αλλάξει, ο απόγονος μπορεί να διαφέρει πολύ ή λίγο από τους γονείς.  Μπορεί παράλληλα να γίνει και πλατιά (wide) εξερεύνηση του χώρου και “fine tuning”

Παρατηρήσεις

- Απαιτούνται δυο γονείς για την αναπαραγωγή ενός απογόνου. Θα μπορούσε κανείς να φανταστεί και «όργια» όπου συμμετέχουν περισσότεροι γονείς ➡ αλλά πρακτικά σπάνια βελτιώνονται τα αποτελέσματα.
- Απαιτείται να είναι δυνατή η αποτίμηση της συνάρτησης fitness , αλλά δεν είναι απαραίτητο να είναι διαφορίσιμη ως προς της παραμέτρους εισόδου ➡ πρακτικά μπορεί να είναι πολύ χρήσιμο.
- Στο επόμενο μάθημα θα δούμε στη θεωρία Σχημάτων (Shema Theory) πώς η Επιλογή σε συνδυασμό με την Διαστάυρωση επιτρέπουν την δημιουργία και επιβίωση περισσότερων ατόμων με απόδοση πάνω από τον μέσο όρο της συνάρτησης fitness.

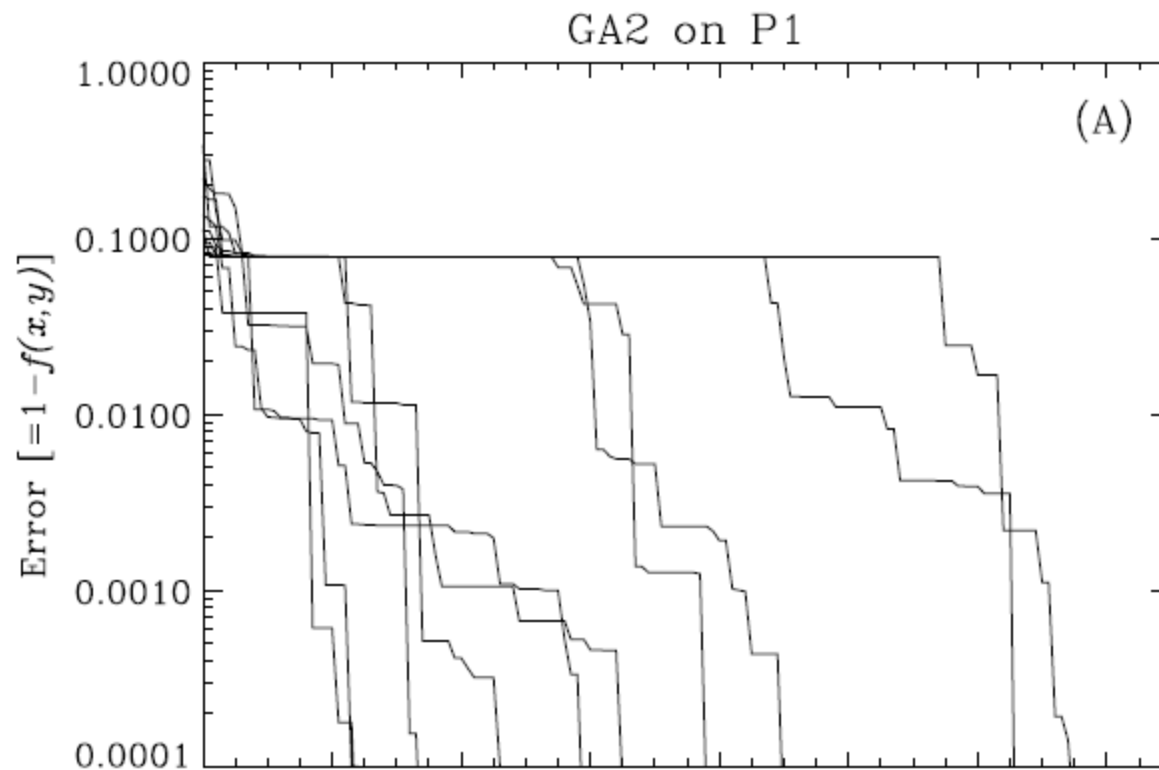
Εφαρμογή Γενετικού Αλγόριθμου στα Προβλήματα Βελτιστοποίησης

- 1. GA1: Απλός Γενετικός με παράμετρους: Roulette Wheel Selection, Πιθανότητα διασταύρωση 0,85, πιθανότητα μετάλλαξης 0.005
- 2. GA2: Ελιτισμός και δυναμικός τρόπος αλλαγής του ρυθμού μετάλλαξης!

Όταν η διαφορά Δf των τιμών της συνάρτησης καταλληλότητας του καταλληλότερου ατόμου και του μεσαίου ατόμου γίνεται μικρή
Αύξηση του ρυθμού μετάλλαξης.

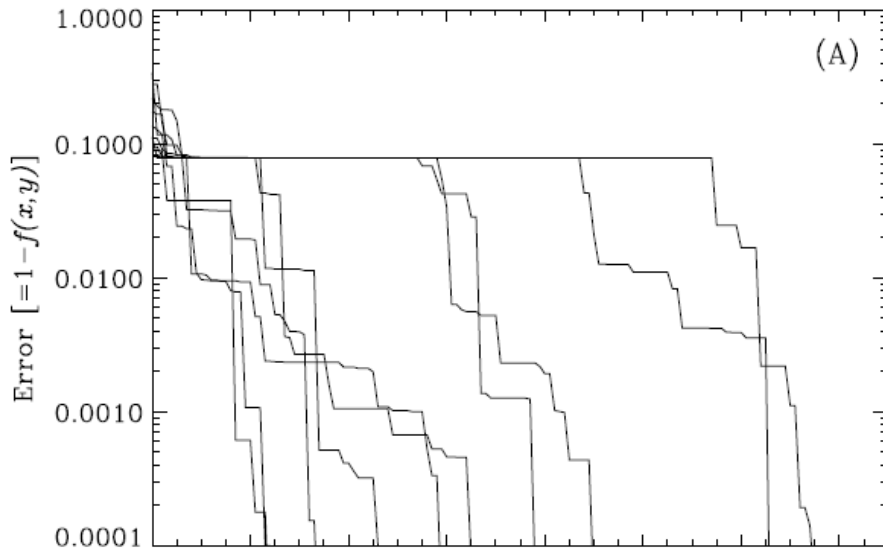


Εφαρμογή Γενετικού Αλγόριθμου στο P1

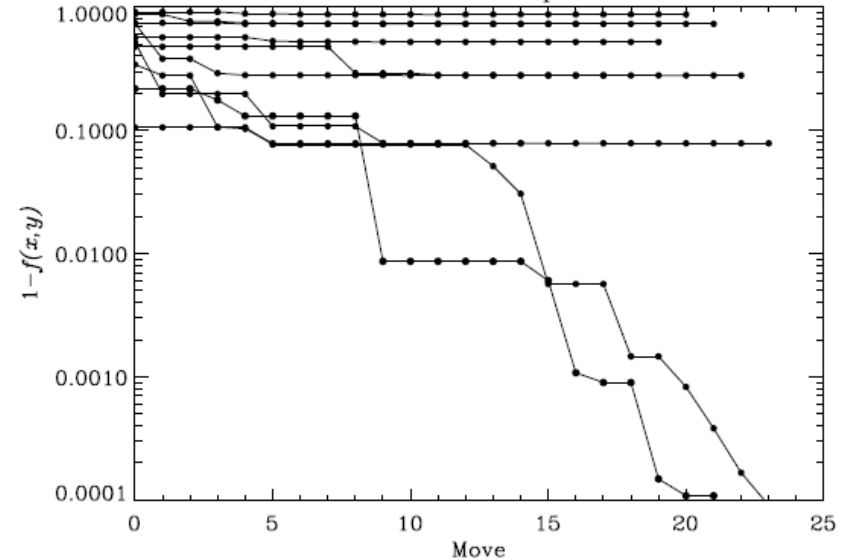


Εφαρμογή Γενετικού Αλγόριθμου στο P1 - Σύγκριση με Simplex

GA2 on P1

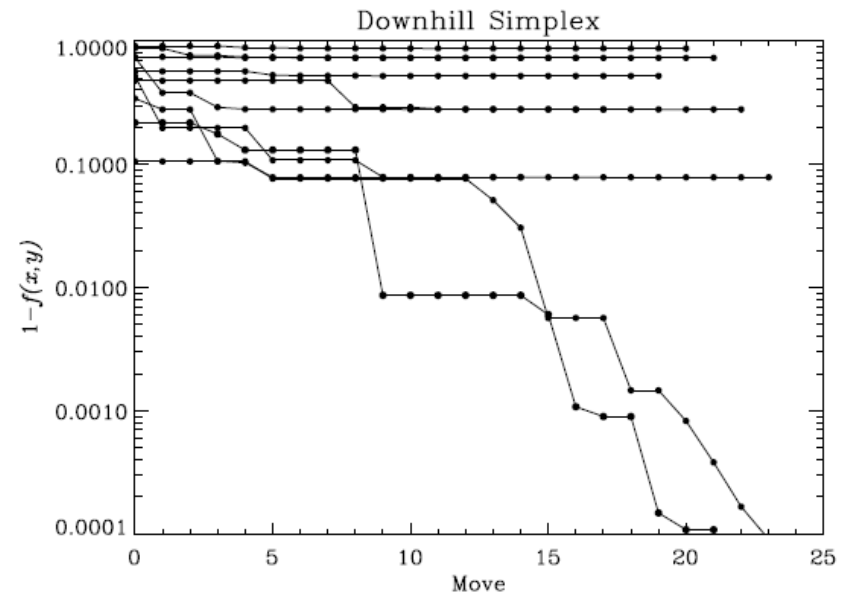
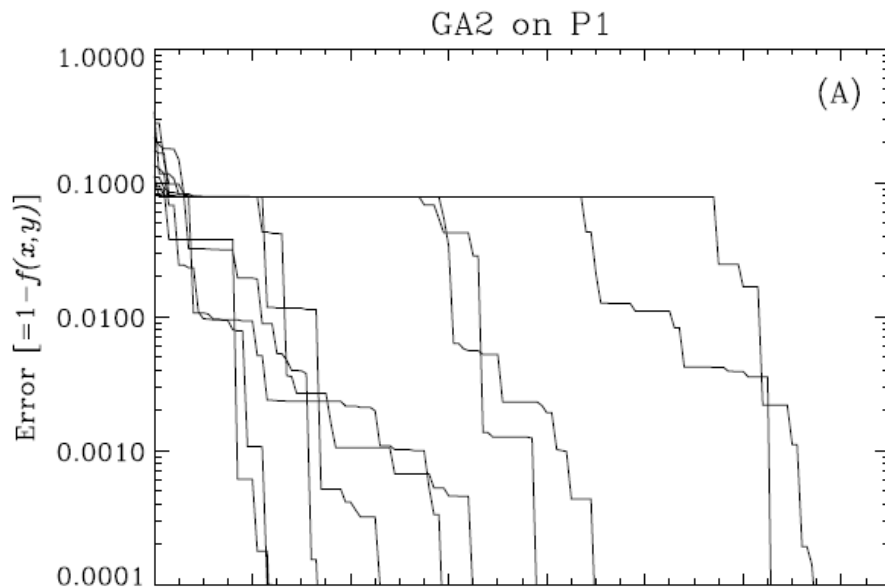


Downhill Simplex



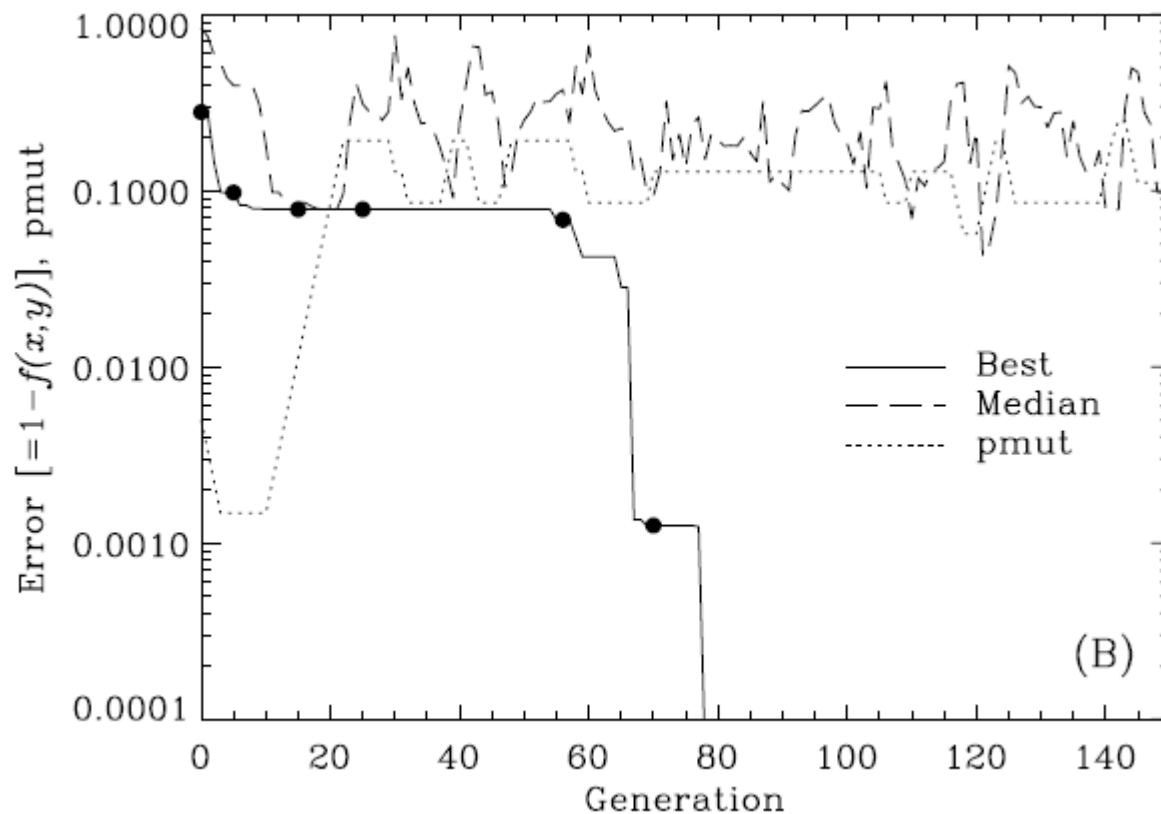
Τι παρατηρούμε?

Εφαρμογή Γενετικού Αλγόριθμου στο P1 - Σύγκριση με Simplex

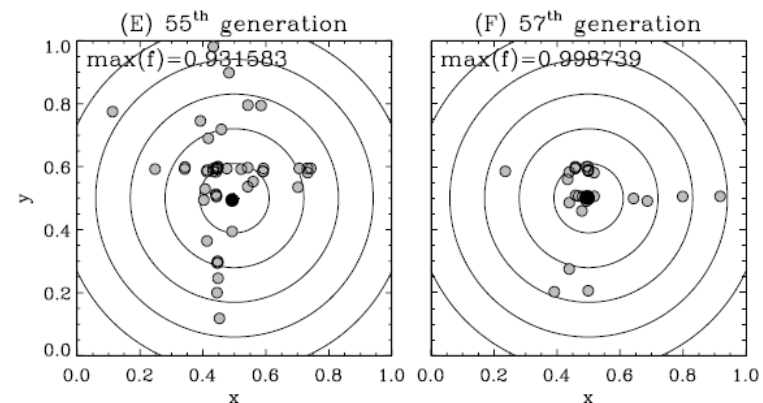
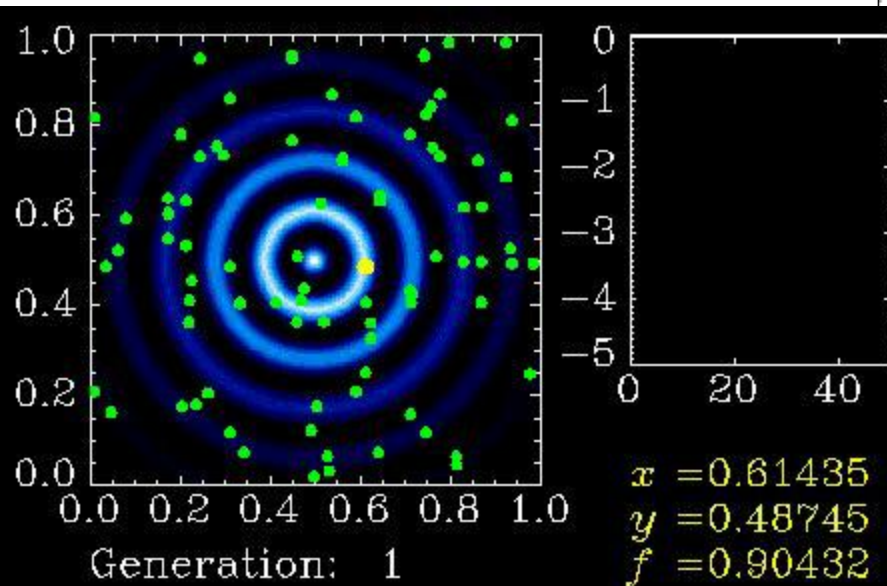
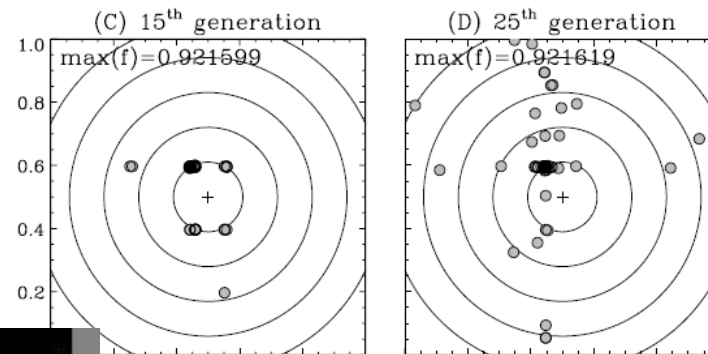
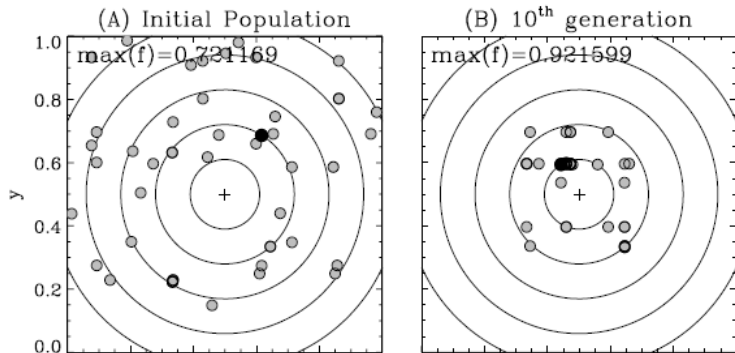
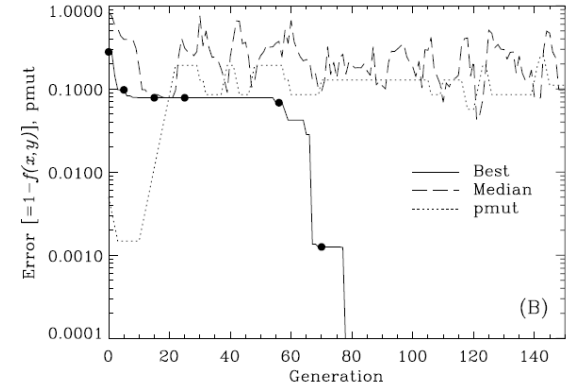


Παρατηρούμε μια παρόμοια μορφή, αλλά εδώ ο γενετικός καταφέρνει να ξεφύγει από τα τοπικά ελάχιστα (κυρίως λόγω της μετάλλαξης, αλλά και της διασταύρωσης)

Εφαρμογή Γενετικού Αλγόριθμου στο P1



Εφαρμογή Γενετικού Αλγόριθμου στο P1



Αλγόριθμος GA3 - Creep Mutation

- ✓ Έστω ότι απαιτείται το:21000.....; το οποίο αποκωδικοποιείται στο 2.1000
- ✓ Έστω ότι νωρίς εμφανίζεται το:19123..... το οποίο έχει μεγάλη τιμή fitness,
- ✓ Οπότε πιθανώς να διαδοθεί και να εμφανιστεί το:19994.....

Αν και πολύ κοντά για να πάει προς το 21000 απαιτείται **ταυτόχρονη** μετάλλαξη :
 $1 \longrightarrow 2$ και το πρώτο $9 \longrightarrow 0$.

Θα αργήσει πολύ να συμβεί....



Είτε διαφορετική κωδικοποίηση είτε τροποποίηση του τελεστή μετάλλαξης



Κώδικας Gray



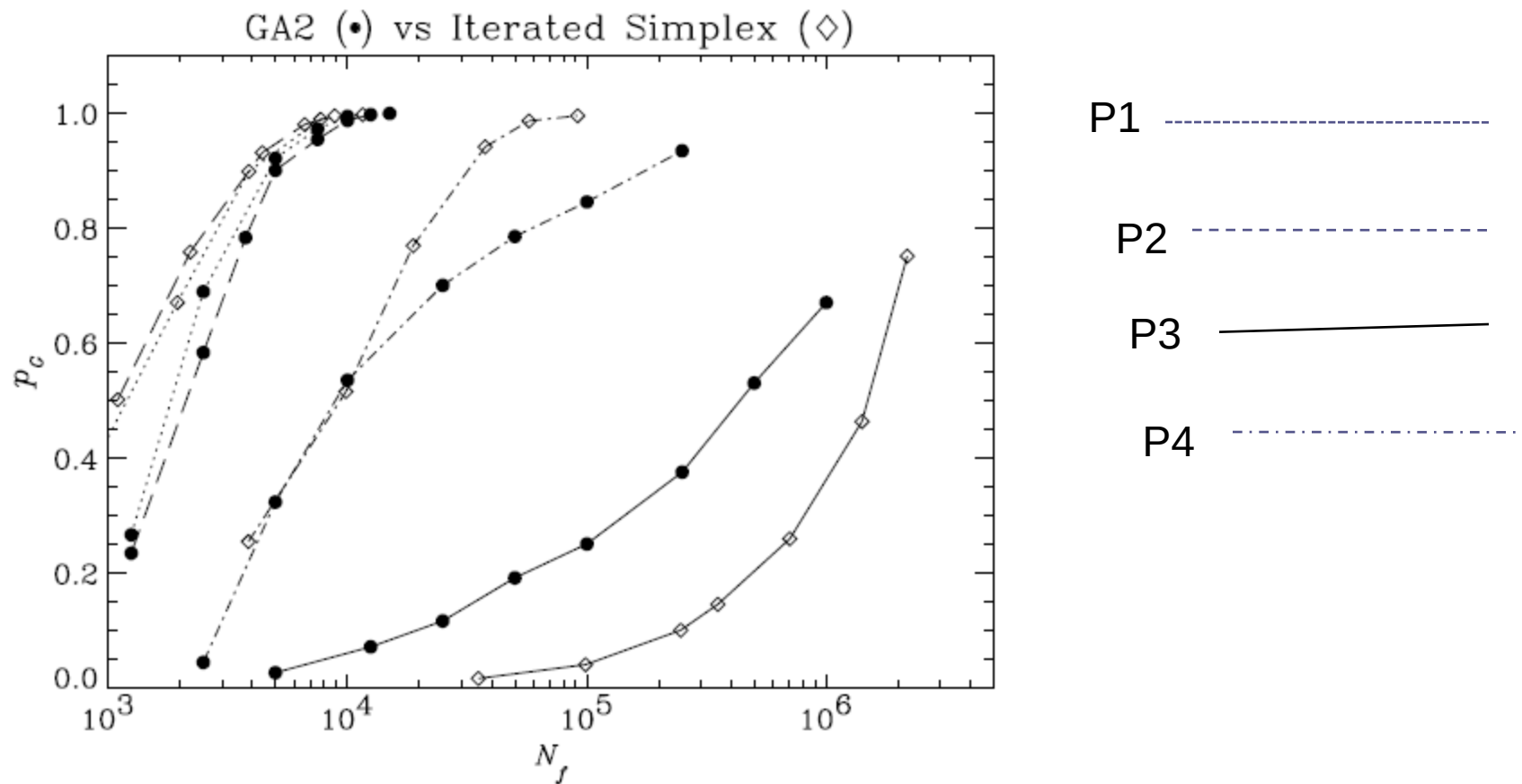
Creep mutation

Πχ μετάλλαξη στο $\longrightarrow$ 20094.....
 μεσαίο 9 με +1



Δεν επιτρέπει όμως μεγάλα άλματα, οπότε συνήθως συνδυασμός

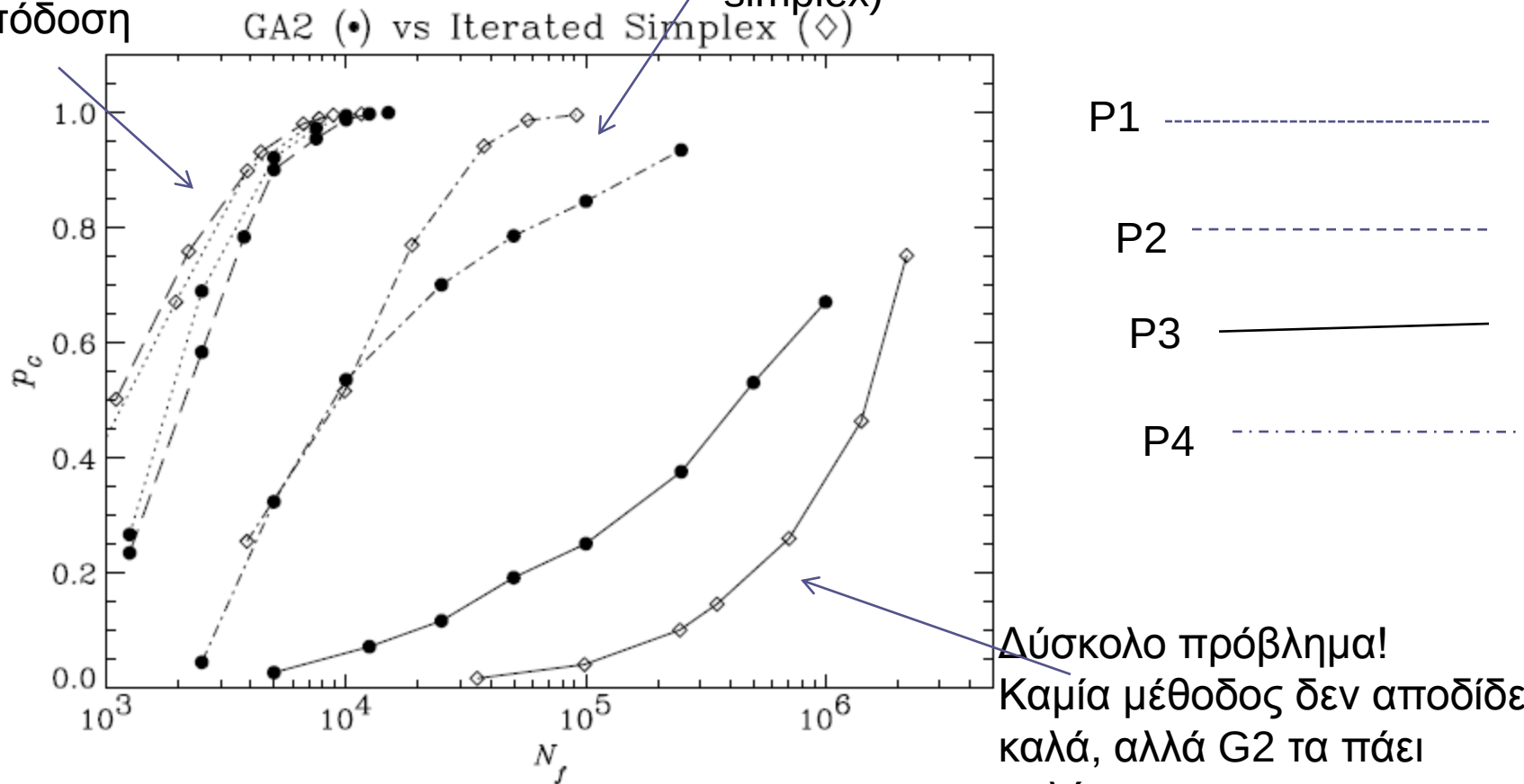
Απόδοση GA2 - Επαναληπτικής Simplex



Απόδοση GA2 - Επαναληπτικής Simplex

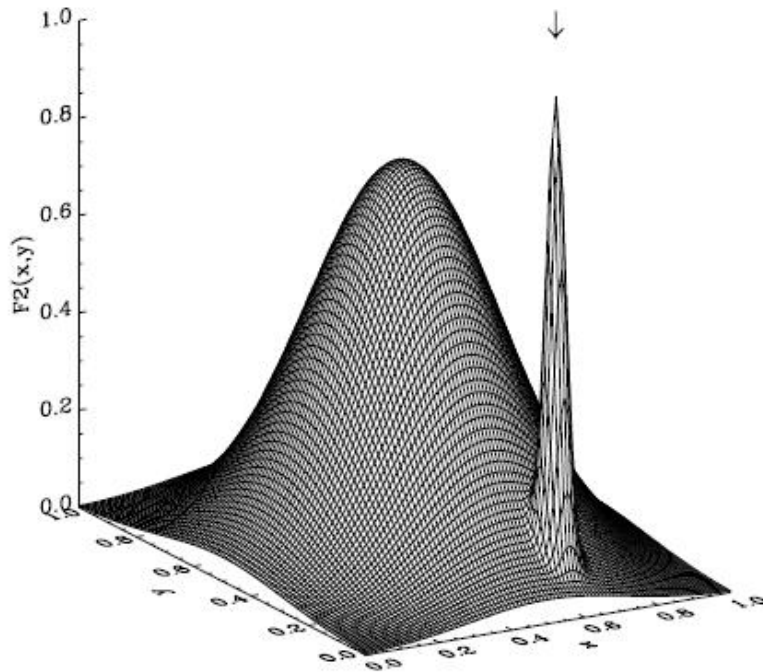
Και οι δύο μέθοδοι έχουν περίπου ίδια απόδοση

Περίπου ίδια απόδοση μέχρι το $p_g=0,5$ μετά υστερεί η GA2 (πρόβλημα που βολεύει την simplex)



Δύσκολο πρόβλημα!
Καμία μέθοδος δεν αποδίδει καλά, αλλά G2 τα πάει καλύτερα

Ο ρόλος της κωδικοποίησης



Για να ξεφύγει από το τοπικό μέγιστο στο
 $(x, y) \simeq (0.5, 0.5)$
 και να βρεθεί στο $(x, y) = (0.6, 0.1)$
 απαιτούνται 2 ταυτόχρονες μεταλλάξεις!

Αν το ολικό μέγιστο ήταν στο
 $(x, y) = (0.5, 0.9)$
 θα αρκούσε μια!

Figure 5: Test problem P2. The problem consist in maximizing a function of two variables, defined by two Gaussians (see eqs. [5]). The global maximum is $f(x, y) = 1$ at $(x, y) = (0.6, 0.1)$, and is indicated by the arrow.

Σε αυτό το τροποποιημένο πρόβλημα η απόδοση του GA2 είναι σαφώς καλύτερη από της επαναληπτικής Simplex! $p_g=0,987$ $N_f=2500$

Απόδοση

Test Problem	Performance	Iter.	Simplex	GA1	GA2	GA3
P1	$\langle 1 - f \rangle$	0.05471		0.0753	0.0071	0.0024
	p_G	0.302		0.119	0.914	0.971
	$\langle N_f \rangle$	671		5000	5000	5000
	N_t or N_g	17		100	100	100
P2	$\langle 1 - f \rangle$	0.11583		0.1961	0.0212	0.0289
	p_G	0.353		0.018	0.883	0.840
	$\langle N_f \rangle$	745		5000	5000	5000
	N_t or N_g	17		100	100	100
P3	$\langle 1 - f \rangle$	0.15547		0.166	0.0634	0.0614
	p_G	<0.0001		0.001	0.191	0.230
	$\langle N_f \rangle$	727		25000	50000	50000
	N_t or N_g	10		500	1000	1000
P4	$\langle R \rangle$	0.0619		0.199	0.039	0.076
	p_G	0.619		0.149	0.845	0.698
	$\langle N_f \rangle$	5593		50000	50000	50000
	N_t or N_g	7		1000	1000	1000

Θέτοντας

$$N_t = N_p / (n + 1)$$

Παρατηρήσεις

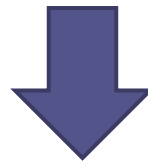
- GA2 και GA3 καλύτερη απόδοση από GA1
- GA2 και GA3 καλύτερη απόδοση από Simplex για δεδομένο επίπεδο δειγματοληψίας του χώρου αναζήτησης (Ανταλλαγή πληροφορίας μεταξύ των λύσεων)
- Όφελος της μετάλλαξης ερπυσμού (creep mutation) όταν υπάρχουν Hamming walls, διαφορετικά δεν επιτρέπονται μεγάλα πηδήματα και μπορεί να χειροτερεύσει την απόδοση : πχ στο πχ P_3 όπου απαιτείται μεγάλο βήμα από το $(x, y) \simeq (0.5, 0.5)$ στο $(x, y) = (0.6, 0.1)$

Παρατηρήσεις

- Οι τυχαία αρχικοποίηση του πληθυσμού στους ΓΑ δεν εισάγει κάποιο bias
- Για την αριθμητική βελτιστοποίηση ο ελιτισμός και ένας δυναμικός ρυθμός μετάλλαξης συνήθως είναι σημαντικές προσθήκες
- Επαναληπτική αναρρίχηση λόφων χρήση της μεθόδου simplex αποτελεί καλή μέθοδο βελτιστοποίησης – ιδίως στις χαμηλές διαστάσεις!
- Οι μετρικές απόδοσης οποιασδήποτε μεθόδου βελτιστοποίησης εξαρτώνται από το πρόβλημα!

Υβριδικές Μέθοδοι

- Αναρρίχηση Λόφων είναι γρήγορη, αλλά τοπική μέθοδος
- Οι Γενετικοί Αλγόριθμοι είναι global μέθοδος, αλλά αργή



Συνδυασμός!

Γενετικός για $N_g/3$ γενιές και αρχικοποίηση της simplex (ή οποιασδήποτε άλλης μέθοδο) με το καταλληλότερο άτομο

Πότε χρησιμοποιούμε Γενετικούς Αλγορίθμους?

- Προβλήματα multimodal όπου δύσκολα μπορεί να έχει κανείς αξιόπιστη αρχική εκτίμηση για το που βρίσκεται το ολικό μέγιστο
- Προβλήματα όπου είναι δύσκολο να υπολογιστούν οι παράγωγοι ως προς τις παραμέτρους. Αν υπάρχει αξιόπιστη εκτίμηση για το ολικό μέγιστο η simplex είναι ισχυρός αντίπαλος, διαφορετικά στρεφόμαστε στους Γενετικούς Αλγορίθμους
- Ill-conditioned προβλήματα μοντελοποίησης ιδίως αυτά που περιγράφονται από εξισώσεις με ολοκληρώματα
- Προβλήματα με περιορισμούς θετικότητας και μονοτονίας που μπορούν να ενσωματωθούν στην κωδικοποίηση του ΓΑ.

Πρακτικά όμως...

- If you already have something that works well enough for you don't mess with it!
- Use what works!

Και μην ξεχνάτε

THE NO-GHOST-IN-THE-MACHINE RULE

“Evolution is cleverer than you are”