ΠΛΗ31

ΕΝΟΤΗΤΑ 4: ΓΕΝΕΤΙΚΟΙ ΑΛΓΟΡΙΘΜΟΙ

Μάθημα 4.5:
Το Πρόβλημα της Ικανοποιησιμότητας - SAT

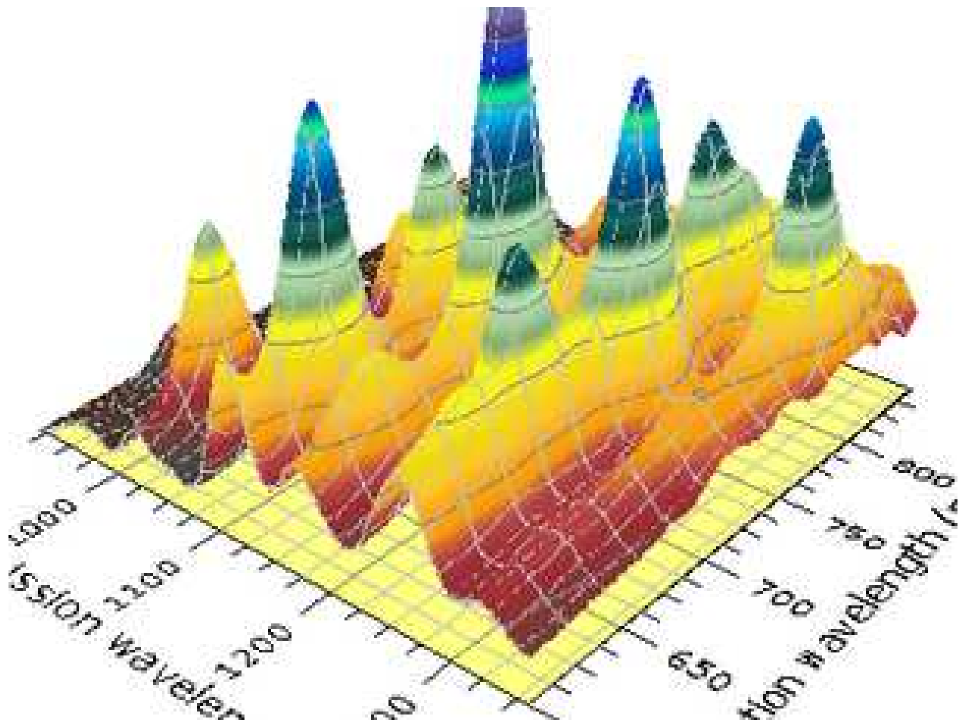
Δημήτρης Ψούνης

www.psounis.gr

ΠΕΡΙΕΧΟΜΕΝΑ

Α. Θεωρία
Προβλήματα Συνδυαστικής Βελτιστοποίησης

- Τα Προβλήματα Συνδυαστικής Βελτιστοποίησης είναι το κατ΄εξοχήν πεδίο εφαρμογής των Γενετικών Αλγορίθμων
  - Τέτοια προβλήματα είναι το TSP, το SAT κ.λπ.
- Οι αντικειμενικές συναρτήσεις αυτών των προβλημάτων είναι ιδιαίτερα περίπλοκες με αποτέλεσμα ο χώρος αναζήτησης να παρουσιάζει ιδιαίτερες αυξομειώσεις.
- Είναι πολύ εύκολο να γίνει εγκλωβισμός σε τοπικά μέγιστα
- Οι Γενετικοί Αλγόριθμοι έρχονται να προσπεράσουν αυτό το πρόβλημα!



Β. Θεωρία

1. Το πρόβλημα της Ικανοποιησιμότητας (SAT)

1. Διατύπωση του Προβλήματος

Το πρόβλημα της ικανοποιησιμότητας (Satisfiability - SAT):

- **Είσοδος:** Δίνεται φόρμουλα φ σε κάνονική συζευκτική μορφή.
- **Ερώτημα:** Είναι η φ ικανοποιήσιμη;

Υπενθύμιση:
Μια φόρμουλα προτασιακής λογικής φ είναι σε κανονική συζευκτική μορφή (Κ.Σ.Μ.) αν είναι στην μορφή:

$$C_1 \wedge C_2 \wedge \cdots \wedge C_m$$

δηλαδή αποτελείται από m προτάσεις που η κάθε πρόταση C_i είναι της μορφής:

$$z_{i_1} \vee z_{i_2} \vee \cdots \vee z_{i_n}$$

δηλαδή αποτελείται από i_n όρους με κάθε όρο να είναι είτε μία προτασιακη μεταβλητή x_j είτε άρνηση προτασιακής μεταβλητής $\neg x_j$.

Επίσης μία φόρμουλα λέμε ότι είναι ικανοποίησιμη, αν υπάρχει αποτίμηση των μεταβλητών (τιμές Α,Ψ) που να την κάνουν αληθή.

Β. Θεωρία

1. Το πρόβλημα της Ικανοποιησιμότητας (SAT)

2. Παραδείγματα Στιγμιοτύπων

Παράδειγμα 1:

Η φόρμουλα SAT:

$$\varphi_1 = (x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2 \vee x_3) \wedge (\neg x_2 \vee x_3)$$

Είναι ικανοποιήσιμη, για παράδειγμα με την αποτίμηση $x_1 = \mathrm{A}, x_2 = \mathrm{A}, x_3 = \mathrm{A}$

Παράδειγμα 2:

Η φόρμουλα SAT:

$$\varphi_2 = (\neg x_1 \vee x_2) \wedge (x_1 \vee x_2 \vee x_3) \wedge (\neg x_2) \wedge (x_1 \vee x_2 \vee \neg x_3)$$

Δεν είναι ικανοποιήσιμη.

- Για να αληθεύει η πρόταση $(\neg x_2)$ πρέπει $x_2 = \Psi$
- Συνεπώς η πρόταση $(\neg x_1 \vee x_2)$ είναι $(\neg x_1 \vee \Psi)$ πρέπει $x_1 = \Psi$
- Συνεπώς η πρόταση $(x_1 \vee x_2 \vee x_3)$ είναι $(\Psi \vee \Psi \vee x_3)$ πρέπει $x_3 = \mathrm{A}$
- Συνεπώς η πρόταση $(x_1 \vee x_2 \vee \neg x_3)$ είναι $(\Psi \vee \Psi \vee \Psi)$ άρα δεν είναι αληθής.

Άρα η φόρμουλα είναι ψευδής.

Α. Θεωρία

1. Το πρόβλημα της ικανοποιησιμότητας - SAT

3. Σχόλια για την μοντελοποίηση του προβλήματος

- Μία λύση θα αναπαρίστατα με μία δυαδική συμβολοσειρά μήκους n, όπου στον αριθμό 0 θα αντιστοιχίσουμε το λογικό Ψευδές και στο 1 θα αντιστοιχίσουμε το λογικό Αληθές
  - Η $1^{η}$ λύση θα αναπαρίσταται με το διάνυσμα: 111
  - Η $2^{η}$ λύση θα αναπαρίσταται με το διάνυσμα: 001
- Οι λύσεις που έχει ο χώρος αναζήτησης είναι εκθετικά πολλές.
  - Αποδεικνύεται ότι είναι 2^{n} (όσες και οι διαφορετικές αναθέσεις τιμών στις μεταβλητές).
- Θα τροποποιήσουμε τον γενετικό αλγόριθμο ώστε να επιλύει το πρόβλημα της λογικής ικανοποιησιμότητας

Α. Θεωρία

2. Γενετικός Αλγόριθμος για το πρόβλημα SAT

- Θα μελετήσουμε πως μπορούμε να χρησιμοποιήσουμε τον γενετικό αλγόριθμο για να υπολογίσουμε μια ικανοποιητική λύση για το πρόβλημα SAT.

1. **Αρχικοποίηση** του πληθυσμού (Initialization)
2. Επανέλαβε:
    1. **Αξιολόγηση** κάθε στοιχείου του πληθυσμού
    2. **Επιλογή** ενός νέου πληθυσμού (τελεστής επιλογής)
    3. **Διασταύρωση** στοιχείων του πληθυσμού (τελεστής διασταύρωσης)
    4. **Μετάλλαξη** στοιχείων του πληθυσμού (τελεστής μετάλλαξης)

    Εως ότου να ικανοποιηθεί το **κριτήριο τερματισμού** του ΓΑ

Α. Θεωρία

3. Γενετικός Αλγόριθμος για το πρόβλημα SAT

-1. Διατύπωση της αντικειμενικής συνάρτησης

Η αντικειμενική συνάρτηση για το πρόβλημα SAT ορίζεται ως το πλήθος των προτάσεων που ικανοποιούνται από μία αποτίμηση.

Η συνάρτηση αυτή είναι μια συνάρτηση **μεγιστοποίησης**.

Παράδειγμα

Έστω ότι θέλουμε να ικανοποιήσουμε την λογική συνάρτηση:

$$(x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2 \vee x_4) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_2 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee x_2)$$

Η ελάχιστη τιμή της αντικειμενικής συνάρτησης είναι 0 και η μέγιστη τιμή είναι 5 (αφού η φόρμουλα έχει 5 προτάσεις)

Α. Θεωρία

3. Γενετικός Αλγόριθμος για το πρόβλημα SAT

0. Κωδικοποίηση των λύσεων

Η κωδικοποίηση της λύσης θα γίνει με ένα δυαδικό διάνυσμα.

Μία λύση θα αναπαρίσταται δυαδικό διάνυσμα μήκους n, όπου στην θέση i θα θα έχουμε

- 0 αν η τιμή της μεταβλητής x_i είναι Ψ
- 1 αν η τιμή της μεταβλητής x_i είναι Α

Παράδειγμα

Π.χ. το διάνυσμα ακεραίων 1110 αντιστοιχεί στην ανάθεση των τιμών στις μεταβλητές:

- $x_1 = \mathrm{A}$
- $x_2 = \mathrm{A}$
- $x_3 = \mathrm{A}$
- $x_4 = \Psi$

Α. Θεωρία

3. Γενετικός Αλγόριθμος για το πρόβλημα SAT

1. Αρχικοποίηση

Αρχικοποίηση

Στο βήμα της αρχικοποίησης δημιουργούμε έναν τυχαίο πληθυσμό από δυνατές λύσεις

Το πλήθος των τυχαίων λύσεων που παράγονται είναι **pop_size** (παράμετρος του προβλήματος)

Η αρχικοποίηση μπορεί να γίνει:

- Είτε μέσω των τυχαίων αριθμών να επιλέγουμε κάθε bit ισοπίθανα αν είναι 0 ή 1 (π.χ. θέτοντας 0 αν ο τυχαίος αριθμός είναι ≤0.50 και 1 αλλιώς)
- Είτε με κάποιον ευρετικό αλγόριθμο.

Αρχικοποίηση

Έστω ότι ο πληθυσμός έχει **pop_size=4**
Παράγουμε 4 διανύσματα ακεραίων με τυχαίο τρόπο

Α: 0101
Β: 1100
Γ: 0000
Δ: 0100

Α. Θεωρία

3. Γενετικός Αλγόριθμος για το πρόβλημα SAT

2.1. Αξιολόγηση

Αξιολόγηση: Η αξιολόγηση γίνεται με ευθύ τρόπο υπολογίζοντας για κάθε λύση την τιμή της στην αντικειμενική συνάρτηση.

Παράδειγμα: Αξιολογούμε κάθε μέλος του πληθυσμού με βάση την συνάρτηση αξιολόγησης:

$$(x_1 \vee \neg x_2) \wedge (\neg x_1 \vee x_2 \vee x_4) \wedge (\neg x_2 \vee x_3) \wedge (\neg x_2 \vee \neg x_3 \vee \neg x_4) \wedge (\neg x_1 \vee x_2)$$

Συνεπώς η αξιολόγηση των μελών του πληθυσμού είναι:

Α: 0101 αντιστοιχεί στην $x_1 = \Psi, x_2 = \text{A}, x_3 = \Psi, x_4 = \text{A}$ αξιολόγηση f(A)=3

Β: 1100 αντιστοιχεί στην $x_1 = \text{A}, x_2 = \text{A}, x_3 = \Psi, x_4 = \Psi$ αξιολόγηση f(B)=4

Γ: 0000 αντιστοιχεί στην $x_1 = \Psi, x_2 = \Psi, x_3 = \Psi, x_4 = \Psi$ αξιολόγηση f(Γ)=4

Δ: 0100 αντιστοιχεί στην $x_1 = \Psi, x_2 = \text{A}, x_3 = \Psi, x_4 = \Psi$ αξιολόγηση f(Δ)=3

Α. Θεωρία

3. Γενετικός Αλγόριθμος για το πρόβλημα SAT

2.2. Επιλογή

Επιλογή

Η επιλογή γίνεται με την μέθοδο της εξαναγκασμένης ρουλέτας όπως την μελετήσαμε σε προηγούμενο μάθημα.

Το αποτέλεσμα της εκτέλεσης της εξαναγκασμένης ρουλέτας είναι η παραγωγή ενός προσωρινου πληθυσμού μεγέθους pop_size

Άσκηση

Κατασκευάστε την εξαναγκασμένη ρουλέτα για να παράγετε τον προσωρινό πληθυσμό.

Χρησιμοποιήστε την ακολουθία τυχαίων αριθμών 0.112, 0.974, 0.412, 0.229

Α. Θεωρία

3. Γενετικός Αλγόριθμος για το πρόβλημα SAT

2.3. Διασταύρωση

Διασταύρωση: Γίνεται μέσω του τελεστή διασταύρωσης μονού σημείου

Η διασταύρωση γίνεται με διασταύρωση μονού σημείου αφού η αναπαράσταση της λύσης γίνεται με δυαδική συμβολοσειρά μεγέθους n.

Άσκηση

Εφαρμόστε τελεστή διασταύρωσης μονού σημείου θεωρώντας ισοπίθανη την επιλογή του σημείου διαχωρισμού για κάθε ζεύγος το οποίο κατασκευάστηκε στο προηγούμενο βήμα.
Χρησιμοποιήστε τους τυχαίους αριθμούς: 0.442 και 0.793

Α. Θεωρία

3. Γενετικός Αλγόριθμος για το πρόβλημα SAT

2.4. Μετάλλαξη

Μετάλλαξη

Η μετάλλαξη γίνεται αντιστρέφοντας την τιμή του bit ενός γονιδίου ανάλογα με το αν ο τυχαίος αριθμός που επιλέγεται έχει τιμή ≤ από την πιθανότητα μετάλλαξης p_m

Άσκηση

Εφαρμόστε τον τελεστή της μετάλλαξης στον πληθυσμό της τρέχουσας γενιάς, θεωρώντας την πιθανότητα μετάλλαξης ίση με 0.2. Μπορείτε να χρησιμοποιήσετε την παρακάτω ακολουθία τυχαίων αριθμών: 0.34, 0.30, 0.85, 0.56, 0.63, 0.47, 0.19, 0.80, 0.98, 0.58, 0.03, 0.57. Αν χρειαστείτε παραπάνω τυχαίους αριθμούς επαναλάβετε από την αρχή.

Β. Ασκήσεις
Εφαρμογή 1

Έστω ότι θέλουμε να βρούμε μια ανάθεση τιμών στις δυαδικές μεταβλητές a, b, c, d, e και f, ώστε να ικανοποιείται η παρακάτω σχέση (είναι γνωστό σαν πρόβλημα λογικής ικανοποιησιμότητας – SAT):

$$G = (\neg a \vee c) \wedge (\neg a \vee c \vee \neg e) \wedge (\neg b \vee c \vee d \vee \neg e) \wedge (a \vee \neg b \vee c) \wedge (\neg e \vee f)$$

Θέλουμε, λοιπόν, να χρησιμοποιήσουμε Γενετικό Αλγόριθμο για να επιλύσουμε το παραπάνω πρόβλημα SAT, 6 μεταβλητών.

Υπόδειξη 1: Χρησιμοποιείστε τους τυχαίους αριθμούς για να δημιουργήσετε τον αρχικό πληθυσμό, για να επιλέξετε ζευγάρια χρωμοσωμάτων, τα σημεία διασταύρωσης και τις ενδεχόμενες μεταλλάξεις. Σε περίπτωση που χρειαστείτε επιπλέον τυχαίους αριθμούς, χρησιμοποιείστε τον πίνακα ξανά από την αρχή.

Θεωρείστε ότι έχετε γεννήτρια τυχαίων αριθμών η οποία σας δίνει (με τη σειρά) την παρακάτω ακολουθία τυχαίων αριθμών από το 0 ως το 1.

0.5653, 0.7850, 0.3352, 0.4554, 0.2919, 0.5357, 0.2466, 0.5077, 0.4815, 0.6790, 0.4668, 0.6764, 0.4161,

0.7796, 0.5559, 0.1280, 0.7301, 0.1737, 0.2309, 0.7655, 0.3338, 0.1255, 0.5173, 0.3148, 0.2881, 0.6349,

0.8326, 0.3914, 0.7681, 0.5750, 0.0540, 0.6870, 0.6314, 0.6923, 0.2917, 0.9627, 0.4428, 0.4976, 0.0262,

0.0744, 0.2175, 0.7504, 0.8668, 0.6196, 0.0340, 0.3349, 0.2569, 0.6596, 0.8477, 0.3751, 0.9119, 0.4655,

0.3057, 0.1837, 0.7605, 0.8132, 0.2156, 0.3142, 0.5552, 0.8473, 0.4889, 0.0474, 0.6617, 0.1524, 0.3824,

0.2644, 0.3426, 0.1142, 0.3901, 0.1443, 0.7898, 0.5873

(α) (2/25) Ποιο είναι το μέγεθος του χώρου αναζήτησης;

(β) (2/25) Να επιλέξετε ένα τρόπο αναπαράστασης και κωδικοποίησης για τα άτομα (χρωμοσώματα) του πληθυσμού. Να εξηγήσετε το λόγο για την επιλογή σας και να δώσετε ένα παράδειγμα ατόμου (χρωμοσώματος).

(γ) (2/25) Να δημιουργήσετε έναν αρχικό πληθυσμό τεσσάρων χρωμοσωμάτων με τυχαίο τρόπο.

(δ) (4/25) Η G προφανώς, δεν μπορεί να χρησιμοποιηθεί σαν συνάρτηση καταλληλότητας, γιατί για κάθε συμβολοσειρά που δεν είναι λύση, δίνει την τιμή 0 (False), (αντίστοιχα, για κάθε συμβολοσειρά που είναι λύση δίνει την τιμή 1 (True)), άρα όλα τα άτομα που δεν είναι λύσεις δίνουν την ίδια καταλληλότητα (το ίδιο συμβαίνει και για τα άτομα που είναι λύσεις). Να ορίσετε μια συνάρτηση καταλληλότητας, που δεν έχει το συγκεκριμένο πρόβλημα.

(ε) (10/25) Να επιλύσετε το πρόβλημα εφαρμόζοντας τις τεχνικές της διασταύρωσης και της μετάλλαξης. Θεωρείστε πιθανότητα διασταύρωσης ίση με 1 και μετάλλαξης ίση με 0.001.

(στ) (5/25) Στο ΓΑ που σχεδιάσατε στα παραπάνω ερωτήματα, μετά από μερικές γενιές, τι πρόβλημα δημιουργείται εάν τα άτομα του πληθυσμού γίνουν όμοια;