

ΧΡΩΜΑΤΙΣΜΟΣ ΓΡΑΦΗΜΑΤΩΝ

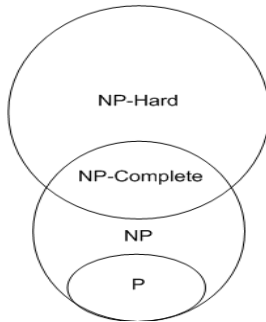
Ονοματεπώνυμο : Γιάννης Σαμαρτζής

ΠΕΡΙΛΗΨΗ :

Στόχος της εργασίας είναι η επίλυση του προβλήματος χρωματισμού γραφημάτων με τη χρήση δυο διαφορετικών αλγορίθμων (FIRST FIT, DSATUR)

1 ΕΙΣΑΓΩΓΗ

Αρχικά θα πρέπει να επισημάνουμε τι σημαίνουν ορισμένες έννοιες προβλημάτων.



Υπάρχουν υπολογιστικά προβλήματα που δεν μπορούν να επιλυθούν με αλγόριθμους ακόμη και σε απεριόριστο χρόνο.

NP-Complete προβλήματα είναι προβλήματα των οποίων η κατάσταση είναι άγνωστη. Κανένας αλγόριθμος πολυωνυμικού χρόνου δεν έχει ακόμη ανακαλυφθεί για οποιοδήποτε NP-Complete πρόβλημα, ούτε κανείς μπορεί να αποδείξει ότι δεν υπάρχει αλγόριθμος πολυωνυμικού χρόνου για κανένα από αυτά. Το ενδιαφέρον είναι ότι, εάν κάποιο από τα NP-Complete προβλήματα μπορεί να λυθεί σε πολυωνυμικό χρόνο, τότε όλα μπορούν να λυθούν.

Το P είναι ένα σύνολο προβλημάτων που μπορούν να λυθούν με μια ντετερμινιστική μηχανή Turing σε πολυωνυμικό χρόνο.

Το NP είναι ένα σύνολο προβλημάτων λήψης αποφάσεων που μπορούν να επιλυθούν με ένα μη-ντετερμινιστικό Turing Machine σε πολυωνυμικό χρόνο. Το P είναι υποσύνολο του NP (οποιοδήποτε πρόβλημα μπορεί να λυθεί με ντετερμινιστική μηχανή σε πολυωνυμικό χρόνο μπορεί επίσης να λυθεί με μη ντετερμινιστική μηχανή σε πολυωνυμικό χρόνο).

Ένα πρόβλημα είναι NP-hard αν ένας αλγόριθμος για την επίλυσή του μπορεί να μεταφραστεί σε έναν για την επίλυση οποιουδήποτε NP προβλήματος.

2 ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ ΧΡΩΜΑΤΙΣΜΟΥ ΓΡΑΦΗΜΑΤΩΝ

Το πρόβλημα χρωματισμού γραφήματος είναι η εκχώρηση χρωμάτων σε ορισμένα στοιχεία ενός γραφήματος. Ο χρωματισμός κορυφής είναι το πιο κοινό πρόβλημα χρωματισμού γραφημάτων. Το πρόβλημα είναι να βρούμε έναν τρόπο χρωματισμού των κορυφών ενός γραφήματος έτσι ώστε να μην χρωματίζονται δύο γειτονικές κορυφές με το ίδιο χρώμα. Το πρόβλημα εύρεσης χρωματικού αριθμού ενός δεδομένου γραφήματος είναι NP Complete.

3 ΠΡΟΣΕΓΓΙΣΕΙΣ ΕΠΙΛΥΣΗΣ

3.1 ΔΕΔΟΜΕΝΑ ΠΡΟΒΛΗΜΑΤΟΣ (Toronto datasets)

Πρόβλημα	Αρχείο Δεδομένων	Εξετάσεις	Φοιτητές	Εγγραφές
car-f-92	car-f-92.stu	543	18419	55522
car-s-91	car-s-91.stu	682	16925	56877
ear-f-83	ear-f-83.stu	190	1125	8109
hec-s-92	hec-s-92.stu	81	2823	10632
kfu-s-93	kfu-s-93.stu	461	5349	25113
lse-f-91	lse-f-91.stu	381	2726	10918
pur-s-93	pur-s-93.stu	2419	30029	120681
rye-s-93	rye-s-93.stu	486	11483	45051
sta-f-83	sta-f-83.stu	139	611	5751
tre-s-92	tre-s-92.stu	261	4360	14901
uta-s-92	uta-s-92.stu	622	21266	58979
ute-s-92	ute-s-92.stu	184	2749	11793
yor-f-83	yor-f-83.stu	181	941	6034

3.2 ΠΙΝΑΚΑΣ ΣΤΑΤΙΣΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ ΠΡΟΒΛΗΜΑΤΩΝ

Name	v	Density	Min	Med	Max	Mean	CV
"car_f_92"	"543"	"0.137478333874986"	"0"	"63"	"381"	"74.7882136279926"	"75.3453219477669"
"car_s_91"	"682"	"0.12801037341726"	"0"	"77"	"472"	"87.4310850439883"	"70.9100354514407"
"ear_f_83"	"190"	"0.264149903554698"	"4"	"45"	"134"	"50.4526315789474"	"56.1132202421981"
"hec_s_92"	"81"	"0.410418548629931"	"9"	"32"	"62"	"33.6543209876543"	"36.326432578737"
"kfu_s_93"	"461"	"0.0553380097848645"	"0"	"18"	"247"	"25.5661605206074"	"119.986786189461"
"lse_f_91"	"381"	"0.0622638138819035"	"0"	"16"	"134"	"23.7847769028871"	"93.1559778679734"
"pur_s_93"	"2419"	"0.0294708898902969"	"0"	"47"	"857"	"71.3195535345184"	"129.479558893176"
"rye_s_93"	"486"	"0.0749697906896173"	"0"	"24"	"274"	"36.5102880658436"	"111.760954254931"
"sta_f_83"	"139"	"0.141932168550874"	"7"	"16"	"61"	"19.8705035971223"	"67.3647701703798"
"tre_s_92"	"261"	"0.179316194320143"	"0"	"45"	"145"	"46.9808429118774"	"59.6186944164299"
"uta_s_92"	"622"	"0.125154191160911"	"1"	"65"	"303"	"77.9710610932476"	"73.6710414481887"
"ute_s_92"	"184"	"0.0840188014101058"	"2"	"13"	"58"	"15.5434782608696"	"69.1350613432102"
"yor_f_83"	"181"	"0.285714285714286"	"7"	"51"	"117"	"52"	"35.2268095548348"

V: Αριθμός κορυφών.

Density: Η πυκνότητα συγκρούσεων υπολογίζεται διαιρώντας τον αριθμό των στοιχείων του πίνακα συγκρούσεων που έχουν την τιμή 1 με το συνολικό πλήθος των στοιχείων του πίνακα.

Min Med Max Mean CV: Για τους βαθμούς (degrees) των κορυφών η ελάχιστη τιμή (min), η διάμεσος τιμή (median), η μέγιστη τιμή (max), η μέση τιμή (mean) καθώς και ο συντελεστής διακύμανσης (CV=coefficient of variation) που ορίζεται ως η τυπική απόκλιση προς τη μέση τιμή.

4 ΕΠΙΛΥΣΗ ΠΡΟΒΛΗΜΑΤΟΣ

4.1 FIRST FIT

Ο αλγόριθμος first fit είναι ένας άπληστος (greedy) αλγόριθμος που λαμβάνει κάθε κορυφή και την αναθέτει στο μικρότερο αριθμό χρώματος που δεν προκαλεί σύγκρουση, δημιουργώντας νέα χρώματα όταν χρειάζεται. Οι κορυφές μπορούν αρχικά να ταξινομηθούν σε φθίνουσα σειρά βαθμού, όπως έχει προταθεί στο [WP67] και ο χρωματισμός των κορυφών να γίνει από την κορυφή με τον υψηλότερο βαθμό προς την κορυφή με το χαμηλότερο βαθμό.

4.2 DSATUR

Ο αλγόριθμος DSATUR [Bré79] καθορίζει δυναμικά την επόμενη κορυφή που θα χρωματιστεί επιλέγοντας ανάμεσα στις κορυφές που δεν είναι χρωματισμένες εκείνη που κάθε φορά έχει το μεγαλύτερο αριθμό διαφορετικών χρωμάτων σε γειτονικές κορυφές. Η χειρότερη περίπτωση πολυπλοκότητας του DSatur είναι $O(n^2)$, ωστόσο στην πράξη ορισμένα επιπλέον έξοδα προκύπτουν από την ανάγκη συγκράτησης του βαθμού κορεσμού των άχρωμων κορυφών.

5 ΑΠΟΤΕΛΕΣΜΑΤΑ

Name	V	Density	Min	Med	Max	Mean	CV	FirstFit	Dsatur
"car_f_92"	"543"	"0.137478333874986"	"0"	"63"	"381"	"74.7882136279926"	"75.3453219477669"	"44"	"31"
"car_s_91"	"682"	"0.12801037341726"	"0"	"77"	"472"	"87.4310850439883"	"70.9100354514407"	"48"	"34"
"ear_f_83"	"190"	"0.264149903554698"	"4"	"45"	"134"	"50.4526315789474"	"56.1132202421981"	"29"	"25"
"hec_s_92"	"81"	"0.410418548629931"	"9"	"32"	"62"	"33.6543209876543"	"36.326432578737"	"22"	"20"
"kfu_s_93"	"461"	"0.0553380097848645"	"0"	"18"	"247"	"25.5661605206074"	"119.986786189461"	"25"	"20"
"lse_f_91"	"381"	"0.0622638138819035"	"0"	"16"	"134"	"23.7847769028871"	"93.1559778679734"	"22"	"19"
"pur_s_93"	"2419"	"0.0294708898902969"	"0"	"47"	"857"	"71.3195535345184"	"129.479558893176"	"54"	"37"
"rye_s_93"	"486"	"0.0749697906896173"	"0"	"24"	"274"	"36.5102880658436"	"111.760954254931"	"28"	"26"
"sta_f_83"	"139"	"0.141932168550874"	"7"	"16"	"61"	"19.8705035971223"	"67.3647701703798"	"13"	"13"
"tre_s_92"	"261"	"0.179316194320143"	"0"	"45"	"145"	"46.9808429118774"	"59.6186944164299"	"29"	"24"
"uta_s_92"	"622"	"0.125154191160911"	"1"	"65"	"303"	"77.9710610932476"	"73.6710414481887"	"43"	"36"
"ute_s_92"	"184"	"0.0840188014101058"	"2"	"13"	"58"	"15.5434782608696"	"69.1350613432102"	"13"	"11"
"yor_f_83"	"181"	"0.285714285714286"	"7"	"51"	"117"	"52"	"35.2268095548348"	"27"	"23"

6 ΣΥΜΠΕΡΑΣΜΑΤΑ

Συγκρίνοντας τους δυο αλγορίθμους παρατηρούμε πως ο Dsatur χρωματίζει με ίσο ή μικρότερο σύνολο χρωμάτων. Στα συγκεκριμένα datasets φαίνεται πως όσο μεγαλύτερο το σύνολο των εγγραφών τόσο καλύτερα αποτελέσματα θα έχει ο Dsatur, π.χ. στο dataset sta_f_83 υπάρχουν 5751 εγγραφές και οι δυο αλγόριθμοι παράγουν τον ίδιο σύνολο χρωμάτων 13, ενώ στο dataset pur_s_93 που έχει 120681 εγγραφές ο Dsatur χρησιμοποιεί 17 χρώματα λιγότερο 37 αντί 54 σε σχέση με τον FirstFit. Συμπεραίνουμε λοιπόν πως αν ο στόχος μας είναι να χρωματίσουμε ένα γράφημα με όσο το δυνατόν λιγότερα χρώματα ο Dsatur είναι καλύτερη επιλογή σε σχέση με τον FirstFit.

ΑΝΑΦΟΡΕΣ:

[1] <https://www.geeksforgeeks.org/np-completeness-set-1>

[1] <https://mathworld.wolfram.com/NP-HardProblem.html>

[2] <https://www.geeksforgeeks.org/graph-coloring-applications>

[4.1] [WP67] Dominic JA Welsh and Martin B Powell. An upper bound for the chromatic number of a graph and its application to timetabling problems. The Computer Journal, 10(1):85–86, 1967.

[4.2] [Bré79] Daniel Brélaz. New methods to color the vertices of a graph. Communications of the ACM, 22(4):251–256, 1979.

[4.2] <https://en.wikipedia.org/wiki/DSatur>