Journal of Discrete Algorithms 35 (2015) 1–8

Contents lists available at ScienceDirect

Journal of Discrete Algorithms

www.elsevier.com/locate/jda

On the complexity of role colouring planar graphs, trees and cographs

Christopher Purcell, Puck Rombach [*,1]

Department of Mathematics, University of California Los Angeles, 520 Portola Plaza, CA 90095, USA

ARTICLE INFO

Article history:
Received 12 June 2014
Received in revised form 3 July 2015
Accepted 6 August 2015
Available online 10 August 2015

Keywords:
Role colouring
Regular equivalence
Locally surjective homomorphism
Complexity
Planar
Tree
Cograph

ABSTRACT

We prove several results about the complexity of the role colouring problem. A *role colouring* of a graph G is an assignment of colours to the vertices of G such that two vertices of the same colour have identical sets of colours in their neighbourhoods. We show that the problem of finding a role colouring with $1 < k < n$ colours is NP-hard for planar graphs. We show that restricting the problem to trees yields a polynomially solvable case, as long as k is either constant or has a constant difference with n, the number of vertices in the tree. Finally, we prove that cographs are always k-role-colourable for $1 < k \le n$ and construct such a colouring in polynomial time.

Published by Elsevier B.V.

1. Introduction

A *role colouring* of a graph G is an assignment of colours to the vertices of G such that two vertices of the same colour have identical sets of colours in their neighbourhoods. For example, suppose we colour the vertices of G red or blue. If this colouring is a role colouring then for all red vertices u and v we have that u has a blue neighbour if and only if v has a blue neighbour. The concept arises from the study of social networks. Network science is an increasingly important application of graph theory and role colourings are a natural formulation of roles played by nodes in a real-world network [16,17]. This structure was formalised by White and Reitz in terms of graph homomorphisms in [22], and developed extensively by Borgatti and Everett [2,1,7,8]. A fast, applicable algorithm for finding role colourings is proposed in [12,3]. A homomorphism h is said to be *locally surjective* if h is surjective when restricted to the neighbourhood set of any vertex. Locally surjective homomorphisms are equivalent to role colourings and they appear in the literature under many other names, *e.g.* role assignment [21], role equivalence [3], regular equivalence [2]. Throughout this paper we use the language of graph colourings and we refer to a role colouring using k colours as a k-role-colouring.

We consider the computational problem associated with role colourings whose input is a graph G and whose output is a partition of the vertices of G into k non-empty subsets satisfying the definition of a role colouring given above. We call this problem k-ROLE-COLOURABILITY, or k-ROLECOL for short. This problem differs from the more common COLOURABILITY problem in a few important ways. A k-role-colouring does not usually imply the existence of a $k+1$-role-colouring and one cannot necessarily combine role colourings given for each connected component of a graph. Additionally, every graph with

* Corresponding author.
E-mail address: rombach@math.ucla.edu (P. Rombach).
1 Tel.: +1 (310) 206 9974.

http://dx.doi.org/10.1016/j.jda.2015.08.001
1570-8667/Published by Elsevier B.V.

no isolated vertex has a 1-role-colouring obtained by giving each vertex the same colour. Any colouring obtained in this way or by giving each vertex its own colour is said to be a *trivial* role colouring.

Finding role colourings of a given size is known to be NP-complete in general [15,10]. For $k \geq 3$, the k-ROLECOL problem is NP-complete when restricted to chordal graphs [21]. However, 2-role-colouring can be solved in polynomial time for chordal graphs [19]. Not many other partial results on complexity of role colouring are known. In fact, interval graphs and trees are the only non-trivial classes in which a polynomial solution is known to exist, and only for a constant number of colours.

The rest of this paper is organised as follows. In Section 2, we prove that k-ROLECOL remains NP-complete even when restricted to planar graphs, a class that was suggested for examination in [21] and is one of the most extensively studied in the literature. In Section 3, we give an explicit algorithm that computes a k-role-colouring of a tree in polynomial time, as long as k is either constant or has a constant difference with n, the number of vertices in T. Finally, in Section 4, we show that every cograph (with at least k vertices) has a k-role-colouring, and hence that the decision version of the problem is solvable in polynomial time in this class. Our proof is constructive and gives an explicit algorithm to construct such a colouring.

2. Planar graphs

In order to prove that k-ROLECOL is NP-complete when restricted to planar graphs, we introduce the SATISFIABILITY PROBLEM, defined below. A boolean formula ϕ (in *conjunctive normal form*) is a set of clauses $C_1, C_2, \dots$, each of which is a set of variables $x_1, x_2, \dots$. The variables may take values TRUE or FALSE. For a given assignment of these values to the variables, a clause is said to be *satisfied* if at least one of its variables is assigned the value TRUE. A formula is satisfied if each of its clauses is satisfied. The SATISFIABILITY problem takes a boolean formula on n variables as its input and asks if there is an assignment of TRUE and FALSE to the variables that satisfies the formula. The general SATISFIABILITY problem was the first to be revealed to be NP-complete [5], and remains a central problem in theoretical computer science.

We will use a reduction from a certain restricted version of SATISFIABILITY. In order to describe this restricted problem, we define the following graph theoretic notion. The *formula graph* G_ϕ of a given formula ϕ is a bipartite graph whose vertices correspond to the clauses and variables of ϕ with an edge between C and x if the variable x appears in the clause C. Let k-SATISFIABILITY be the SATISFIABILITY problem with the restriction that each clause contains at most k variables. The 3-SATISFIABILITY problem is NP-complete even when restricted to formulas with planar formula graphs [13]. In [20], Tovey showed that the problem is NP-complete under the restriction that each clause has two or three variables and each variable appears at most three times. We call the corresponding problem $3*, 3*$-SATISFIABILITY. We now combine the restrictions imposed by Tovey and planarity to show that PLANAR $3*, 3*$-SATISFIABILITY, which is $3*, 3*$-SATISFIABILITY restricted to formulas with planar formula graphs, is also NP-complete. We list a couple of planarity preserving operations that we will need throughout the coming proofs, in an easy lemma. See also [11].

Lemma 1. *If G' is a graph created from a planar graph G by any of the following operations, then G' is planar.*

(a) *Adding a path $x, z_1, \dots, z_k, y$ where $x, y \in V(G)$, and x, y share a face in some planar drawing of G, and $z_1, \dots, z_k$ are new vertices.*
(b) *Replacing a vertex $x \in V(G)$ with $d_G(x) = k$ by a cycle $z_1, \dots, z_k$ with edges z_i, z_{i+1}, $1 \leq i \leq k-1$ and z_k, z_1, and edges z_i, y_i, $1 \leq i \leq k$, where $y_1, \dots, y_k$ are the neighbours in G of x appearing in clockwise order in a planar drawing of G.*
(c) *Attaching a new planar subgraph H to G, such that $V(G') = V(G) \cup V(H)$, $E(G') = E(G) \cup E(H) \cup xz$, where $x \in V(G)$, $z \in V(H)$.*

Proof.

(a) Replacing an edge by a multi-edge does not destroy planarity and replacing an edge xy by a path $x, z_1, \dots, z_k, y$ clearly does not destroy planarity either.
(b) Cycles are planar and since the neighbours $y_1, \dots, y_k$ and new vertices $z_1, \dots, z_k$ are in the same order clockwise, the edges z_i, y_i, $1 \leq i \leq k$ do not cross each other or any new cycle-edges.
(c) We take the disjoint union of G and H and draw G such that x is on the outer face and H such that z is on the outer face. Adding an edge between two vertices on the same face does not destroy planarity. □

Lemma 2. *The PLANAR $3*, 3*$-SATISFIABILITY problem is NP-complete.*

Proof. We follow the method [20] of reducing any 3-SATISFIABILITY problem to a $3*, 3*$-SATISFIABILITY problem. Let the formula ϕ be an instance of the PLANAR 3-SATISFIABILITY problem, with a given planar drawing of G_ϕ. We will obtain a formula ϕ' which is an instance of PLANAR $3*, 3*$-SATISFIABILITY from ϕ. By applying operations from Lemma 1 to G_ϕ we obtain a planar drawing of $G_{\phi'}$. Suppose that variable x appears in $k > 3$ clauses, labelled $C_{j_1}, C_{j_2}, \dots, C_{j_k}$ such that their respective edges meet x in clockwise order around x in our given drawing of G_ϕ. Create k new variables $x_1, \dots, x_k$ and in each clause C_{j_i} replace the variable x with the variable x_i. Then add new clauses $C_{x_i} = \{x_i \vee \bar{x}_{i+1}\}$ for $1 \leq i \leq k-1$ and