Theoretical Computer Science 475 (2013) 13–20

Contents lists available at SciVerse ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Partitioning the arcs of a digraph into a star forest of the underlying graph with prescribed orientation properties☆

Jørgen Bang-Jensen [a,*], Daniel Gonçalves [b], Anders Yeo [c]

[a] Department of Mathematics and Computer Science, University of Southern Denmark, Odense DK-5230, Denmark

[b] LIRMM, CNRS and Université Montpellier 2, France

[c] Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom

ARTICLE INFO

Article history:
Received 4 June 2010
Received in revised form 23 October 2012
Accepted 4 December 2012
Communicated by T. Erlebach

Keywords:
2-SAT
NP-completeness proof
Star arboricity

ABSTRACT

A star in an undirected graph is a tree in which at most one vertex has degree larger than one. A star forest is a collection of vertex disjoint stars. An out-star (in-star) in a digraph D is a star in the underlying undirected graph of D such that all edges are directed out of (into) the center. The problem of partitioning the edges of the underlying graph of a digraph D into two star forests $\mathcal{F}_0$ and $\mathcal{F}_1$ is known to be NP-complete. On the other hand, with the additional requirement for $\mathcal{F}_0$ and $\mathcal{F}_1$ to be forests of out-stars the problem becomes polynomial (via an easy reduction to 2-SAT). In this article we settle the complexity of problems lying in between these two problems. Namely, we study the complexity of the related problems where we require each $\mathcal{F}_i$ to be a forest of stars in the underlying sense and require (in different problems) that in D, $\mathcal{F}_i$ is either a forest of out-stars, in-stars, out- or in-stars or just stars in the underlying sense.

1. Introduction

Notation on digraphs which is not given here is consistent with [2]. Generally a digraph is denoted by $D = (V, A)$ where V is the set of vertices in D and A is the set of arcs. We also use $V(D)$ $(A(D))$ to denote these two sets. If there is an arc from a vertex x to a vertex y in D, then we say that x **dominates** y and use the notation $xy \in A$ to denote this. An (x, y)-path is a directed path from x to y.

We let $N^-(x)$ (respectively, $N^+(x)$) denote the set of vertices dominating (respectively, dominated by) x in D, and let $d^-(x) = |N^-(x)|$, $d^+(x) = |N^+(x)|$ and $d(x) = d^+(x) + d^-(x)$.

The **underlying graph** $UG(D)$ of a digraph $D = (V, A)$ is the graph with vertex set V and edge set $E = \{xy|xy \in A$ or $yx \in A\}$. A digraph D is **connected** if $UG(D)$ is a connected graph and the connected components of D are the connected components of $UG(D)$.

The work in this paper was motivated by work in [3–6]. The purpose of these papers was to investigate what happens to the complexity of problems that are mixed versions of problems which are well-defined both for directed and for undirected graphs (this is made precise below).

- In their seminal paper [7] Fortune et al. characterized (among others) the complexity of linking problems where we are seeking a collection of paths (or cycles) $P_1, P_2, \ldots, P_k$ where P_i is an (s_i, t_i)-path or a cycle (when $s_i = t_i$) and no vertex on

☆ This work was done while the first author was on sabbatical at LIRMM Université Montpellier 2 whose hospitality is gratefully acknowledged. Financial support from the Danish National Science research council (FNU) (under grant no. 09-066741) is gratefully acknowledged.

* Corresponding author. Tel.: +45 65502335; fax: +45 65502325.
E-mail addresses: jbj@imada.sdu.dk (J. Bang-Jensen), goncalves@lirmm.fr (D. Gonçalves), anders@cs.rhul.ac.uk (A. Yeo).

0304-3975/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2012.12.007

any P_i occurs on any other path $P_j, j \neq i$ except possibly that s_i or t_i is the initial or terminal vertex of P_j. They showed that these problems are always NP-complete unless we have $s_1 = \cdots = s_k$ or $t_1 = \cdots = t_k$, in which case the problems are all polynomially solvable. In [3] a mixed version of these linking problems was considered, namely we are looking for paths or cycles $P_1, P_2, \ldots, P_k$ as above but with the difference that now some prescribed set of these just have to be paths or cycles in the underlying digraph $UG(D)$ of the input digraph D and hence do not have to respect the orientation of the arcs of D. In the case when all of the paths and cycles may ignore the directions on arcs in D we have the k-linkage problem for undirected graphs which is polynomially solvable by the famous disjoint path algorithm of Robertson and Seymour [12]. Hence, given these extremes (always polynomial for undirected and almost always NP-complete for directed graphs) it is natural to ask whether there is any case of the problem which involves both undirected (ignoring the direction of arcs in D) and directed (must follow the arcs of D) paths which is polynomial. The answer is that there are none, they are all NP-complete [3].

- Deciding whether a (di)graph has a pair of vertex-disjoint cycles is polynomially solvable both in the undirected and the directed case [10,11]. In the directed case this is highly non-trivial to prove. In [4] it was shown that one can decide in polynomial time whether the underlying graph $UG(D)$ of a strongly connected digraph D has vertex-disjoint cycles C, C' such that C' is also a directed cycle in D (but C may not be). However, when D is not strongly connected, the problem becomes NP-complete [5].
- Deciding whether a graph has two edge-disjoint spanning trees as well as whether a digraph has two arc-disjoint out-branchings[1] with prescribed roots are both polynomial problems, see e.g. [2, Chapter 9]. The mixed version of these problems is the following. Given a digraph D, does D contain an out-branching from some root s such that deleting the arcs of this branching leaves a connected digraph? This problem turns out to be NP-complete [6].

All of the results discussed above indicate that one can obtain an interesting insight into the structure of (problems concerning the structure of) digraphs by studying problems of the mixed type above where only part of the structure that is sought has to obey the direction of the arcs in D.

An **out-star** (resp. **in-star**) is a directed graph whose underlying graph is a star, and such that all the arcs are oriented from the center to the leaves (resp. from the leaves to the center). Similarly, an **out-star forest** (resp. an **in-star forest**) is a digraph whose connected components are out-stars (resp. in-stars). A **dir-star forest** is a forest of out-stars and in-stars. A **gen-star forest** is a directed graph whose underlying graph is a star forest. The problem of partitioning the edges of an undirected graph into two star forests (a.k.a. 2-STAR ARBORICITY) is NP-complete even when restricted to the class of bipartite graphs [8,9]. Conversely, the problem of partitioning the edges of a directed graph D into two forests of out-stars (a.k.a. 2-DIRECTED-STAR ARBORICITY) is polynomial [1] (see also the proof in Section 2 below).

In this article we consider the (X,Y)-STAR ARBORICITY problems where we have X,Y $\in$ {OUT,IN,DIR,GEN} for directed graphs. Here the problem is to decide whether a directed graph admits a partition of its arc set into an X-star forest $\mathcal{F}_0$ and a Y-star forest $\mathcal{F}_1$. The motivation for studying these problems comes from the fact that they may be seen as variants (restrictions) of the 2-STAR ARBORICITY problem. Namely (GEN,GEN)-STAR ARBORICITY is the original 2-STAR ARBORICITY problem on the underlying digraph $UG(D)$ of a digraph D and in the other cases we only allow a subset of all possible stars in each of the two forests in $UG(D)$ and the restrictions on these come for the orientations of the edges in D.

Clearly the complexity of (X,Y)-STAR ARBORICITY is the same as that of (Y,X)-STAR ARBORICITY. We shall use this without further mention below.

We have mentioned that (OUT,OUT)-STAR ARBORICITY is polynomial and it is clear (e.g. by reversing all arcs) that it is also the case for (IN,IN)-STAR ARBORICITY. We show in Section 2 that it is also the case of (IN,OUT)-STAR ARBORICITY.

Theorem 1.1. *The* (X,Y)-STAR ARBORICITY *problem is polynomial for every choice of* X,Y $\in$ {IN,OUT}.

An $(U \to W)$**-digraph** is a bipartite digraph D with independent sets U and W, such that every arc is oriented from U to W. Since 2-STAR ARBORICITY is NP-complete for the class of bipartite graphs, it is clear that (DIR,DIR)-STAR ARBORICITY, (DIR,GEN)-STAR ARBORICITY and (GEN,GEN)-STAR ARBORICITY are NP-complete problems, even when reduced to the class of $(U \to W)$-digraphs (as by orienting all edges in a bipartite graph such that an $(U \to W)$-digraph is obtained we see that the problems are equivalent).

Thus the complexity status of all the (X,Y)-STAR ARBORICITY problems are known except for (DIR,OUT)-STAR ARBORICITY (which is equivalent to (DIR,IN)-STAR ARBORICITY, by reversing all arcs) and (GEN,OUT)-STAR ARBORICITY (which is equivalent to (GEN,IN)-STAR ARBORICITY). In Sections 3 and 4 we give the complete answer to these questions.

Theorem 1.2. *The* (DIR,OUT)-STAR ARBORICITY *problem is polynomially solvable.*

Theorem 1.3. *The* (GEN,OUT)-STAR ARBORICITY *problem is NP-complete for general digraphs and polynomially solvable for* $(U \to W)$*-digraphs.*

[1] An out-branching in a digraph is a spanning tree in the underlying graph which is oriented such that every vertex except one, the root, has in-degree one.

2. Proof of Theorem 1.1: decomposing a digraph into two forests of out-stars or a forest of in-stars and a forest of out-stars

In this short section we show how to reduce the (OUT,OUT)-STAR ARBORICITY and (OUT,IN)-STAR ARBORICITY problems to 2-SAT in polynomial time. It is well-known that 2-SAT is solvable in linear time (see e.g. [2, Section 17.5]).

First consider an instance $D = (V, A)$ of (OUT,OUT)-STAR ARBORICITY and form the following instance $\mathcal{I}_D$ of 2-SAT. There is a variable x_{uv} for each arc $uv \in A$ and the following clauses:

- For each directed path uvw in D, $\mathcal{I}_D$ contains the clauses $(x_{uv} \vee x_{vw})$, $(\bar{x}_{uv} \vee \bar{x}_{vw})$ where $\bar{x}$ is the negation of variable x.
- For each pair of arcs uv, wv in A, $\mathcal{I}_D$ contains the clauses $(x_{uv} \vee x_{wv})$, $(\bar{x}_{uv} \vee \bar{x}_{wv})$.

Note that if D contains an out-star that is a connected component in D, then $\mathcal{I}_D$ contains variables that do not appear in any clause, and may therefore be assigned any truth value. It is easy to check that if we interpret a true variable x_{uv} as putting the arc uv in $\mathcal{F}_0$ and putting all other arcs (corresponding to false variables) in $\mathcal{F}_1$ then every truth assignment which satisfies $\mathcal{I}_D$ corresponds to a partition of A into two out-star forests $\mathcal{F}_0$, $\mathcal{F}_1$ (as no directed path of length two and no vertex of in-degree two appears in any of the classes $\mathcal{F}_0$, $\mathcal{F}_1$) and conversely, given such a partition, we can make a satisfying truth assignment by setting x_{uv} true if and only if the arc uv is in $\mathcal{F}_0$. Clearly the reduction can be done in polynomial time.

Now consider an instance $H = (V, A)$ of the (OUT,IN)-STAR ARBORICITY problem and form the following instance $\mathcal{I}_H$ of 2-SAT. There is a variable x_{uv} for each arc $uv \in A$ and the following clauses:

- For each directed path uvw in H, $\mathcal{I}_H$ contains the clauses $(x_{uv} \vee x_{vw})$ and $(\bar{x}_{uv} \vee \bar{x}_{vw})$.
- For each pair of arcs $uv, uw \in A$, $\mathcal{I}_H$ contains the clause $(\bar{x}_{uv} \vee \bar{x}_{uw})$.
- For each pair of arcs $uv, wv \in A$, $\mathcal{I}_H$ contains the clause $(x_{uv} \vee x_{wv})$.

Now it is easy to see that the 2-SAT formula $\mathcal{I}_H$ is satisfiable if and only if the desired partition into an out-star forest $\mathcal{F}_0$ and an in-star forest $\mathcal{F}_1$ exists. Note that here we interpret an arc being true as equivalent to the arc belonging to $\mathcal{F}_1$. Again the reduction is easily performed in polynomial time.

3. Proof of Theorem 1.2: decomposing a digraph into a forest of out-stars and a forest of in- and out-stars

Our next goal is to prove Theorem 1.2. For this we again apply a polynomial reduction to the problem 2-SAT. In order to perform this reduction, which is less straightforward than the one used above, we first derive some properties of a positive solution of our problem.

Definition 3.1. Let D be a digraph. Define an arc, $uv \in A(D)$, to be an ***i*-arc** (“*i*” stands for “in”) if one of the following three conditions holds.

(i) $d^-(v) \geq 3$.
(ii) $d^-(v) = 2$ and $d^+(v) \geq 1$.
(iii) $d^-(v) = 2$ and $d^+(v) = 0$ and $N^-(v) = \{u, u'\}$ and $d^-(u) = d^+(u) = d^-(u') = d^+(u') = 1$.

Definition 3.2. An *IOO*-**partition** of the arcs of a digraph D is a partition of $A(D)$ into three disjoint sets I_0, O_0, O_1 such that each of O_0 and O_1 induces a collection of out-stars and I_0 induces a collection of in-stars and all stars in $O_0 \cup I_0$ are vertex-disjoint and all stars in O_1 are vertex-disjoint.

Furthermore we say that an arc has **color** 0 if it belongs to $I_0 \cup O_0$ and it has color 1 if it belongs to O_1. We will often denote the color of an arc e by $c(e)$.

Note that if (I_0, O_0, O_1) is an *IOO*-partition then $(I_0 \cup O_0, O_1)$ is an (DIR,OUT)-partition.

Lemma 3.3. *If there exists an IOO-partition, (I_0, O_0, O_1), of $A(D)$, then there exists such a partition where all arcs in I_0 are i-arcs and no arc in O_0 is an i-arc.*

Proof. Let D be a digraph and let (I_0, O_0, O_1) be an *IOO*-partition of $A(D)$ with the maximum number of arcs in O_1. Furthermore assume that any star in $I_0 \cup O_0$ which only consists of one arc belongs to I_0 if it is an *i*-arc and otherwise it belongs to O_0 (we may do this as each such arc can be placed arbitrarily in one of O_0, I_0).

Consider some star in I_0 with center vertex x and leaves $y_1, y_2, \ldots, y_l$. If $l \geq 3$ then y_jx is an *i*-arc for all $j = 1, 2, \ldots, l$, by part (i) of Definition 3.1. If $l = 1$ then by our assumption on the *IOO*-partition (I_0, O_0, O_1) we note that y_1x is an *i*-arc. So assume that $l = 2$. If $d^+(x) \geq 1$, then by part (ii) of Definition 3.1 y_1x and y_2x are *i*-arcs, so assume that $d^+(x) = 0$.

If $d^+(y_1) \geq 2$, then let $w \in N^+(y_1) \setminus \{x\}$ and note that $y_1w \in A(O_1)$ and we may move the arc y_1x from I_0 to O_1 (where it gets added to the same star as y_1w), contradicting the maximality of $|A(O_1)|$. Therefore $d^+(y_1) = 1$ and analogously $d^+(y_2) = 1$.

If $d^-(y_1) = 0$, then as above we may move the arc y_1x from I_0 to O_1 (where it becomes a star with one arc). So $d^-(y_1) \geq 1$. If $d^-(y_1) \geq 2$, then we get a contradiction as the arcs into y_1 must be of color 1, but then O_1 is not a vertex-disjoint collection of out-stars. Therefore $d^-(y_1) = 1$ and analogously $d^-(y_2) = 1$. So by part (iii) of Definition 3.1 we note that y_1x and y_2x are *i*-arcs. We have now shown that all arcs in I_0 are *i*-arcs.

We will now prove that no arc in O_0 is an i-arc. For the sake of contradiction assume that uv is an i-arc in O_0. We consider the three reasons why uv is an i-arc in Definition 3.1 separately. First assume that $d^-(v) \geq 3$. As the stars in $I_0 \cup O_0$ are vertex-disjoint, we note that all arcs into v, except uv belong to O_1, a contradiction to O_1 being a collection of out-stars. So now assume that $d^-(v) = 2$ and $d^+(v) \geq 1$. Clearly the arc out of v must be of color 1, which means that both arcs into v have color 0, a contradiction against the stars in $I_0 \cup O_0$ being vertex-disjoint. So we now consider part (iii) of Definition 3.1 and assume that $d^-(v) = 2$ and $d^+(v) = 0$ and $N^-(v) = \{u, u'\}$ and $d^-(u) = d^+(u) = d^-(u') = d^+(u') = 1$. Let $N^-(u) = \{w\}$ and note that $c(wu) = 1$. Clearly $c(u'v) = 1$, as uv is an arc in O_0, which implies that uv is a star in O_0 containing only one arc. Therefore it would have been added to I_0 instead of O_0, a contradiction. Therefore no arc in O_0 is an i-arc. □

Proof of Theorem 1.2. By the remark after Definition 3.2, it suffices to show that we can decide in polynomial time if a digraph D has an *IOO*-partition of $A(D)$. Let D be any digraph. We will now build an instance of 2-SAT as follows. For each arc, e, in D let x_e be a variable, which will be true if e will belong to O_0 and false otherwise. We now add the following clauses.

(a) For all directed paths, uvw, of length 2 add the clauses $(x_{uv} \vee x_{vw})$ and $(\bar{x}_{uv} \vee \bar{x}_{vw})$.
(b) If uv and $u'v$ are distinct arcs into the same vertex, v, then add the clause $(\bar{x}_{uv} \vee \bar{x}_{u'v})$.
(c) If uv and $u'v$ are distinct arcs into the same vertex, v, and they are not both i-arcs, then add the clause $(x_{uv} \vee x_{u'v})$.
(d) If uv and uv' are distinct arcs out of the same vertex, u, and at least one of them is an i-arc, then add the clause $(x_{uv} \vee x_{uv'})$.

We will now show that the above instance of 2-SAT is satisfiable if and only if D contains an *IOO*-partition (I_0, O_0, O_1) of $A(D)$. First assume that it is satisfiable. Assign color 1 to an arc, e, if and only if x_e is true. Consider D_1 which is the digraph induced by the arcs of color 1 and note that there is no path of length two in D_1 by (a) above. Furthermore no vertex has in-degree more than one in D_1 by (b) above. This implies that D_1 is a set of vertex-disjoint out-stars. Now consider D_0 which is the digraph induced by the arcs of color 0.

If $uv \in A(D_0)$ is an i-arc then $d^+_{D_0}(u) = 1$ by (d) and $d^+_{D_0}(v) = 0$ by (a). If $d^-_{D_0}(v) = 1$ then uv is an in-star in D_0, so now assume that $d^-_{D_0}(v) \geq 2$ and $wv \in A(D_0)$ is different from uv. By (c) we note that wv is an i-arc and therefore analogously to above $d^+_{D_0}(w) = 1$. Continuing this for all arcs into v in D_0 we note that we get an in-star consisting of i-arcs and which is vertex-disjoint from all other arcs in D_0.

If $uv \in A(D_0)$ is not an i-arc then $d^-_{D_0}(v) = 1$ by (c) and $d^-_{D_0}(u) = 0$ by (a). If $d^+_{D_0}(u) = 1$ then uv is an out-star in D_0, so now assume that $d^+_{D_0}(u) \geq 2$ and $uz \in A(D_0)$ is different from uv. By (d) we note that uz is not an i-arc and therefore analogously to above $d^-_{D_0}(z) = 1$. Continuing this for all arcs out of u in D_0 we note that we get an out-star containing no i-arcs and which is vertex-disjoint from all other arcs in D_0.

Therefore we do have a *IOO*-partition of $A(D)$ when the instance of 2-SAT is satisfiable.

Now assume that there exists a *IOO*-partition (I_0, O_0, O_1) of $A(D)$ and let x_e be true if and only if $c(e) = 1$ for all arcs e in D. By Lemma 3.3 we may assume that all stars in I_0 consist of i-arcs and there is no i-arc in $A(O_0)$. Clearly the constraints in (a) are satisfied as there is no path of length two in an in-star or out-star. As no vertex has in-degree greater than one in O_1 we note that the constraints in (b) are satisfied. If a constraint, $(x_{uv} \vee x_{u'v})$, in (c) was not satisfied, then $c(uv) = c(u'v) = 0$ which implies that uv and $u'v$ belong to I_0, contradicting the fact that one of the arcs is not an i-arc. So the constraints in (c) are satisfied. If a constraint, $(x_{uv} \vee x_{uv'})$, in (d) was not satisfied, then $c(uv) = c(uv') = 0$ which implies that uv and uv' belong to O_0, contradicting the fact that one of the arcs is an i-arc. So the constraints in (d) are satisfied. This implies that the instance of 2-SAT is satisfiable when we have an *IOO*-partition of $A(D)$. □

4. Decomposing a digraph into a forest of out-stars and a forest of stars

In this section we prove Theorem 1.3. First observe that if D is an $U \rightarrow W$-digraph then every star of a star forest in $UG(D)$ is either an in-star or an out-star and hence it follows from Theorem 1.2 that the (OUT,GEN)-STAR ARBORICITY is polynomial in this case. This proves the last assertion in Theorem 1.3. Clearly the (OUT,GEN)-STAR ARBORICITY problem belongs to the class NP so it remains to prove that it is NP-hard.

The proof of this below uses a number of small digraphs (so-called gadgets) and certain properties of these with respect to legal partitionings of their arcs.

Definition 4.1. Let D be any digraph and let c be any 2-coloring of $A(D)$ into $\{0, 1\}$. Let D_i denote the digraph with vertex set $V(D)$ and containing exactly the arcs of D with color i (for $i = 0, 1$). The 2-coloring is called **good** if D_1 is a vertex-disjoint collection of out-stars and D_0 is a vertex-disjoint collection of stars in $UG(D_0)$.

If a vertex x has at least two arcs incident with it of color i then we call it an ***i*-center**. Note that a vertex can be both a 0-center and a 1-center.

Definition 4.2. Let $D(x, y)$ denote the digraph with vertex set $\{p_0, p_1, p_2, p_3, p_4, w, x, y\}$ and arc set $\{p_0p_1, p_1p_2, p_2p_3, p_3p_4, p_1w, p_2x, p_3y\}$. See Fig. 1.

Lemma 4.3. *If c is a good 2-coloring of $D(x, y)$, then the following holds.*

(a) $c(p_2x) \neq c(p_3y)$.
(b) *p_2 is a $c(p_2x)$-center and p_3 is a $c(p_3y)$-center.*

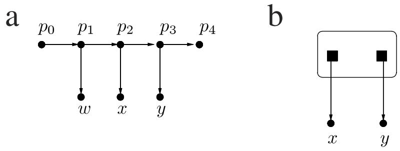

Fig. 1. Part (a) shows the Gadget $D(x, y)$ and (b) shows a symbolic representation with the two important vertices x, y. Here and in the figures below the black rectangles indicate vertices that are centers in any good 2-coloring.

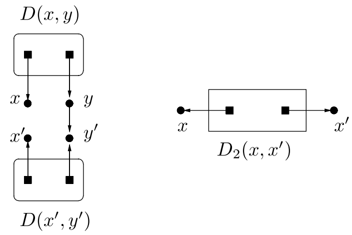

Fig. 2. The left part shows the gadget $D_2(x, x')$ and the right part is a symbolic representation of $D_2(x, x')$.

Furthermore there exists a good 2-coloring of $D(x, y)$, say c_0, where $c_0(p_2x) = 1$ and $c_0(p_3y) = 0$ and also a good 2-coloring, c_1, where $c_1(p_2x) = 0$ and $c_1(p_3y) = 1$.

Proof. Let c be a good 2-coloring of $D(x, y)$ and name the vertices of $D(x, y)$ as it is done in Definition 4.2. As $d(p_2) = 3$, we note that p_2 is either a 0-center or a 1-center. Analogously we note that p_3 is either a 0-center or a 1-center. We now consider the following four possibilities.

p_2 and p_3 are 0-centers: We must have $c(p_2p_3) = 1$ as otherwise we get a path of length three in $UG(D_0)$. As p_2 is a 0-center we must therefore have $c(p_1p_2) = c(p_2x) = 0$. However the arcs p_0p_1 and p_1w cannot both be colored 1 as D_1 is a vertex-disjoint collection of out-stars, so we obtain a path of length three in $UG(D_0)$, a contradiction. So p_2 and p_3 cannot both be 0-centers.

p_2 and p_3 are 1-centers: In this case all out-neighbors of p_2 and p_3 are colored 1 since these vertices have precisely two out-neighbors each (as color 1 is a collection of out-stars). However this implies that $c(p_2p_3) = c(p_3p_4) = 1$, a contradiction, so p_2 and p_3 cannot both be 1-centers.

p_2 is a 0-center and p_3 is an 1-center: As p_3 is an 1-center we note that $c(p_3p_4) = c(p_3y) = 1$. This implies that $c(p_2p_3) = 0$. If $c(p_2x) = 1$, then we must have $c(p_1p_2) = 0$ and as either p_0p_1 or p_1w is colored 0 we get a path of length three in $UG(D_0)$, a contradiction. So we must have $c(p_2x) = 0$. Therefore both (a) and (b) hold.

p_2 is a 1-center and p_3 is a 0-center: As p_2 is a 1-center we note that $c(p_2p_3) = c(p_2x) = 1$. As p_3 is a 0-center we note that $c(p_3p_4) = c(p_3y) = 0$. Therefore both (a) and (b) hold.

Finally let c_0 be the 2-coloring such that the arcs $\{p_2x, p_2p_3, p_0p_1\}$ have color 1 and the arcs $\{p_1p_2, p_1w, p_3p_4, p_3y\}$ have color 0 and note that c_0 is a good 2-coloring. Swap all colors of c_0 in order to obtain c_1 and note that c_1 is a good 2-coloring. □

Definition 4.4. Let $D_2(x, x')$ denote the digraph obtained from two disjoint copies of D, say $D(x, y)$ and $D(x', y')$, by adding the arc yy' (see Fig. 2).

Lemma 4.5. *Let c be a good 2-coloring of $D_2(x, x')$ and denote the vertices in $D(x', y')$ by $p'_0, p'_1, p'_2, p'_3, p'_4, x', y', w'$. The following now holds.*

(i) $c(p_2x) = c(p'_2x')$.
(ii) *p_2 and p'_2 are both $c(p_2x)$-centers.*

Furthermore there exists a good 2-coloring, c_0, where $c_0(p_2x) = 0$ and $c_0(p'_2x') = 0$ and a good 2-coloring, c_1, where $c_1(p_2x) = 1$ and $c_1(p'_2x') = 1$.

Proof. Note that by Part (b) of Lemma 4.3 we have $c(yy') \neq c(p'_3y')$ and $c(yy') \neq c(p_3y)$, which implies that $c(p'_3y') = c(p_3y)$. Part (a) of Lemma 4.3 now implies that $c(p_2x) \neq c(p_3y) = c(p'_3y') \neq c(p'_2x')$, which proves part (i). Part (ii) follows immediately from Part (b) of Lemma 4.3.

Using Lemma 4.3 it is not difficult to construct the colorings c_0 and c_1 mentioned in the statement of the Lemma (we just color yy' differently to p_3y and p'_3y'). □

Definition 4.6. Let $Q(x, a_1, a_2)$ denote the digraph with vertex set $\{x, v_0, v_1, v_2, a_1, a_2\}$ and arc set xv_0, v_0v_1, v_1v_2, v_2a_2, v_1a_1 (Fig. 3).

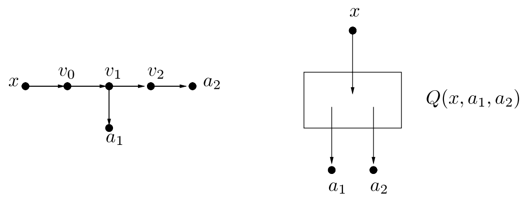

Fig. 3. The digraph $Q(x, a_1, a_2)$. The right part is the symbolic representation used below.

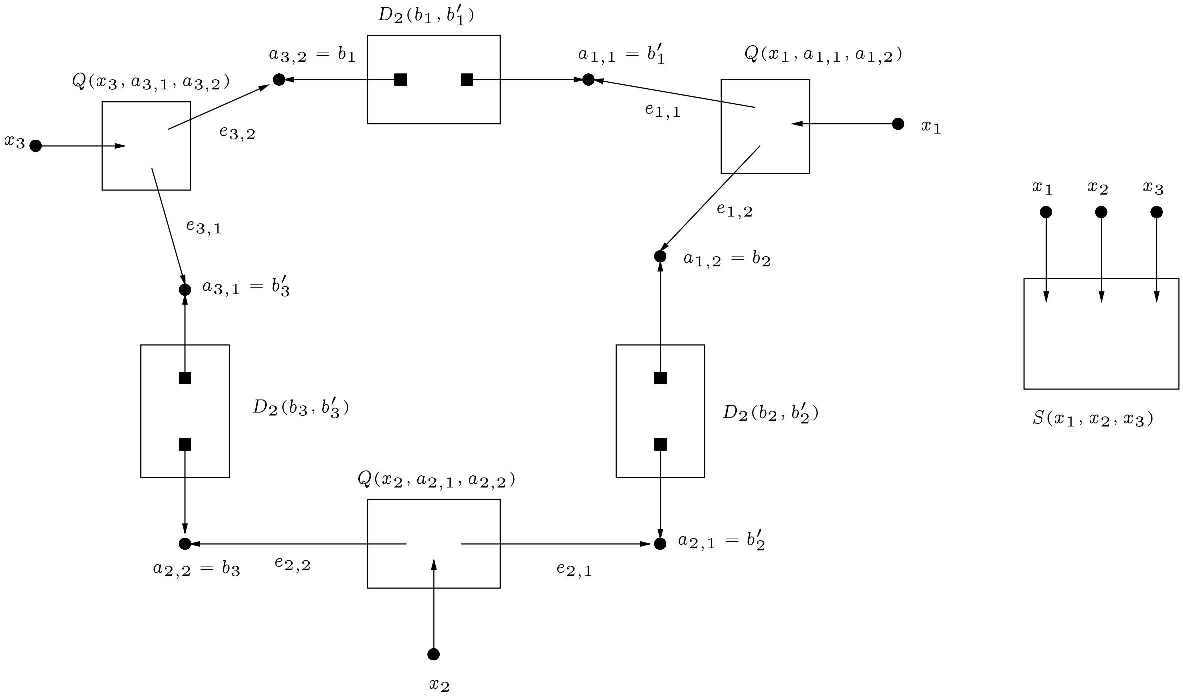

Fig. 4. The digraph $S(x_1, x_2, x_3)$ with a symbolic drawing to the right.

Lemma 4.7. *If c is a good 2-coloring of $Q(x, a_1, a_2)$, then the following holds.*

(1) $(c(xv_0), c(v_1a_1), c(v_2a_2)) \notin \{(0, 0, 0), (0, 1, 1)\}$

For every $q_0, q_1, q_2 \in \{0, 1\}$ with $(q_0, q_1, q_2) \notin \{(0, 0, 0), (0, 1, 1)\}$ there exists a good 2-coloring, c, with $c(xv_0) = q_0$, $c(v_1a_1) = q_1$ and $c(v_2a_2) = q_2$.

Proof. First assume that $(c(xv_0), c(v_1a_1), c(v_2a_2)) = (0, 0, 0)$. As either v_0v_1 or v_1v_2 must have color 0 we get a path of length three in $UG(D_0)$, a contradiction. So now assume that $(c(xv_0), c(v_1a_1), c(v_2a_2)) = (0, 1, 1)$. Clearly $c(v_0v_1) = 0$ (as $c(v_1a_1) = 1$) and $c(v_1v_2) = 0$ (as $c(v_2a_2) = 1$), which gives us a path of length three in $UG(D_0)$. Therefore (1) holds.

Let $q_0, q_1, q_2 \in \{0, 1\}$ with $(q_0, q_1, q_2) \notin \{(0, 0, 0), (0, 1, 1)\}$. Let $c(xv_0) = q_0$, $c(v_1a_1) = q_1$, $c(v_2a_2) = q_2$ and let $c(v_1v_2) = 1 - q_2$. Finally if $(q_0, q_1, q_2) = (0, 0, 1)$, then let $c(v_0v_1) = 1$ otherwise let $c(v_0v_1) = 0$. We note that c is a good 2-coloring with $c(xv_0) = q_0$, $c(v_1a_1) = q_1$ and $c(v_2a_2) = q_2$. □

Definition 4.8. Let $S(x_1, x_2, x_3)$ denote the digraph obtained from three copies of Q, say $Q_1(x_1, a_{1,1}, a_{1,2})$, $Q_2(x_2, a_{2,1}, a_{2,2})$ and $Q_3(x_3, a_{3,1}, a_{3,2})$ and three copies of D_2, say $D_2(b_1, b'_1)$, $D_2(b_2, b'_2)$ and $D_2(b_3, b'_3)$, by identifying the following pairs of vertices $(b'_1, a_{1,1})$, $(a_{1,2}, b_2)$, $(b'_2, a_{2,1})$, $(a_{2,2}, b_3)$, $(b'_3, a_{3,1})$ and $(a_{3,2}, b_1)$ (Fig. 4).

Lemma 4.9. *Let c be a good 2-coloring of $S(x_1, x_2, x_3)$ and let y_i be the unique out-neighbor of x_i for $i = 1, 2, 3$. Then the following holds.*

(I) *We do not have $c(x_1y_1) = c(x_2y_2) = c(x_3y_3) = 0$.*

For every $q_1, q_2, q_3 \in \{0, 1\}$ with $(q_1, q_2, q_3) \neq (0, 0, 0)$ there exists a good 2-coloring, c, with $c(x_1y_1) = q_1$, $c(x_2y_2) = q_2$ and $c(x_3y_3) = q_3$.

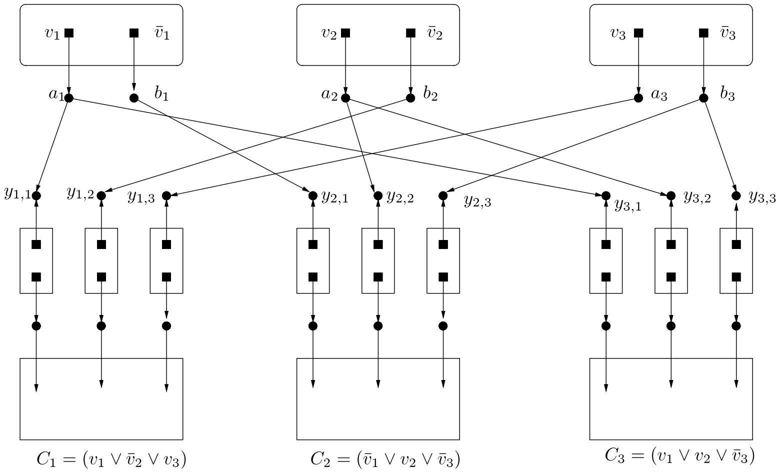

Fig. 5. The digraph $D_{\mathcal{I}}$ corresponding to the formula $\mathcal{I} = (v_1 \vee \bar{v}_2 \vee v_3)\ (\bar{v}_1 \vee v_2 \vee \bar{v}_3)\ (v_1 \vee v_2 \vee \bar{v}_3)$.

Proof. First assume that $c(x_1y_1) = c(x_2y_2) = c(x_3y_3) = 0$. Let the arc into $a_{i,j}$ in Q_i be denoted by $e_{i,j}$ for $i = 1, 2, 3$ and $j = 1, 2$ By Lemma 4.7 we note that $c(e_{i,1}) \neq c(e_{i,2})$, for $i = 1, 2, 3$. By Lemma 4.5 we note that $c(e_{1,1}) = c(e_{3,2})$ and $c(e_{2,1}) = c(e_{1,2})$ and $c(e_{3,1}) = c(e_{2,2})$. However at least two out of these three values (that is, $c(e_{1,1})$, $c(e_{2,1})$ and $c(e_{3,1})$) have the same value, contradicting the fact that $c(e_{i,1}) \neq c(e_{i,2})$, for $i = 1, 2, 3$.

Let $q_1, q_2, q_3 \in \{0, 1\}$ be arbitrary with $(q_1, q_2, q_3) \neq (0, 0, 0)$. By symmetry we may without loss of generality assume that $q_1 = 1$. Let $c(e_{1,1}) = c(e_{1,2}) = c(e_{3,2}) = c(e_{2,1}) = 0$ and $c(e_{3,1}) = c(e_{2,2}) = 1$ and note that by Lemmas 4.5 and 4.7 we can color the rest of the arcs in $S(x_1, x_2, x_3)$ in order to get a good 2-coloring, c, with $c(x_1y_1) = q_1$, $c(x_2y_2) = q_2$ and $c(x_3y_3) = q_3$. □

Theorem 4.10. *It is NP-hard to decide if a digraph has a good* 2*-coloring.*

Proof. We will reduce 3-SAT to the problem of deciding if a digraph has a good 2-coloring. Let $\mathcal{I}$ be an instance of 3-SAT with n variables $v_1, v_2, \ldots, v_n$ and m clauses $C_1, C_2, \ldots, C_m$. For each clause C_i we take a copy of the gadget S, say $S_i(x_{i,1}, x_{i,2}, x_{i,3})$, and three copies of the gadget D_2, say $D^1_{i,2}(x'_{i,1}, y_{i,1})$, $D^2_{i,2}(x'_{i,2}, y_{i,2})$ and $D^3_{i,2}(x'_{i,3}, y_{i,3})$. For each variable, v_i, we take a copy of $D(x, y)$, say $D_i(a_i, b_i)$, and we connect these subgraphs in the following way (Fig. 5):

Identify $x_{i,j}$ with $x'_{i,j}$ for all $i \in [m]$ and $j \in [3]$.

Add an arc from a_i to $y_{r,s}$ if and only if v_i is the s'th literal in C_r (for all $i \in [n]$, $r \in [m]$ and $s \in [3]$).

Add an arc from b_i to $y_{r,s}$ if and only if $\bar{v}_i$ is the s'th literal in C_r (for all $i \in [n]$, $r \in [m]$ and $s \in [3]$).

Call the resulting digraph $D_{\mathcal{I}}$. Clearly we can construct $D_{\mathcal{I}}$ in polynomial time given $\mathcal{I}$. We will now show that $D_{\mathcal{I}}$ has a good 2-coloring if and only if I is satisfiable. So first assume that I is satisfiable, and fix a satisfying assignment of truth values. Let $q_i = 0$ if v_i is true and let $q_i = 1$ if v_i is false. Build a good 2-coloring c as follows. Let the arc into a_i in $D_i(a_i, b_i)$ have color q_i and let the arc into b_i have color $1 - q_i$. Let all arcs of the form $a_iy_{r,s}$ have color $1 - q_i$ and all arcs of the form $b_iy_{r,s}$ have color q_i. For each $D^s_{r,2}(x'_{r,s}, y_{r,s})$ let the arc into $y_{r,s}$ within $D^s_{r,2}(x'_{r,s}, y_{r,s})$ have the opposite color of the arc into $y_{r,s}$ from an a_i or b_i. Let the arc into $x'_{r,s}$ in $D^s_{r,2}(x'_{r,s}, y_{r,s})$ have the same color as the arc into $y_{r,s}$ within $D^s_{r,2}(x'_{r,s}, y_{r,s})$. Finally let the arc out of $x_{r,s}$ in $S_r(x_{r,1}, x_{r,2}, x_{r,3})$ have color opposite to the arc into $y_{r,s}$ within $D^s_{r,2}(x'_{r,s}, y_{r,s})$. As we started with a satisfying assignment we note that no gadget $S_r(x_{r,1}, x_{r,2}, x_{r,3})$ has color 0 on all arcs out of $x_{r,1}$, $x_{r,2}$ and $x_{r,3}$ within $S_r(x_{r,1}, x_{r,2}, x_{r,3})$. Therefore the above lemmas imply that we can complete this coloring to a 2-good coloring of $D_{\mathcal{I}}$.

So now assume that we have a 2-good coloring, c, of $D_{\mathcal{I}}$. Let q_i denote the color of the arc into a_i within $D_i(a_i, b_i)$ and let v_i be true if and only if $q_i = 0$. By Lemma 4.9 we note that for all $r \in [m]$ there is some $s \in [3]$, such that the color on the arc out of $x_{r,s}$ in $S_r(x_{r,1}, x_{r,2}, x_{r,3})$ is 1. By Part (ii) of Lemma 4.5 we note that the arc into $x'_{r,s}$ in $D^s_{r,2}(x'_{r,s}, y_{r,s})$ must be colored 0. By Part (i) of Lemma 4.5 we note that the arc into $y_{r,s}$ in $D^s_{r,2}(x'_{r,s}, y_{r,s})$ must be colored 0. By Part (ii) of Lemma 4.5 we note that if the s'th literal in C_r is v_i then the arc $a_iy_{r,s}$ must be of color 1, which implies that $q_i = 0$, by Lemma 4.3, and so v_i is true. Furthermore if the s'th literal in C_r is $\bar{v}_i$ then the arc $b_iy_{r,s}$ must be of color 1, which implies that $q_i = 1$, by Lemma 4.3, and so v_i is false. So in both cases the clause C_r is satisfied and as r was arbitrary we note that $\mathcal{I}$ is satisfiable. □

As Theorem 4.10 proves the first part of Theorem 1.3 and the remark at the beginning of this section proves the second part of Theorem 1.3 we note that Theorem 1.3 has now been proved.

5. Remarks

The digraph $D_{\mathcal{I}}$ is not bipartite so our proof above does not show that (GEN,OUT)-STAR ARBORICITY is NP-complete already for bipartite graphs. However, the following remarks show that this is indeed the case.

Definition 5.1. Let $D^*(x, x')$ denote the digraph obtained from two disjoint copies of D, say $D(x, y)$ and $D(x', y')$, by identifying y and y'. Furthermore denote the vertices in $D(x, y)$ as in Definition 4.2 and the vertices in $D(x', y')$ by $p'_0, p'_1, p'_2, p'_3, p'_4, x', y', w'$ in the natural way.

Lemma 5.2. *If c is a good 2-coloring of $D^*(x, x')$ (see Definition 5.1), then the following holds.*

(a) $c(p_2x) \neq c(p'_2x')$.
(b) *p_2 is a $c(p_2x)$-center and p'_2 is a $c(p'_2x')$-center.*

Furthermore there exists a good 2-coloring of $D^*(x, x')$, say c_0, where $c_0(p_2x) = 1$ and $c_0(p'_2x') = 0$ and also a good 2-coloring, c_1, where $c_1(p_2x) = 0$ and $c_1(p'_2x') = 1$.

Proof. By Lemma 4.3 we note that $c(p_2x) = 1 - c(p_3y) = c(p'_3y') = 1 - c(p'_2x')$, so (a) holds.
By Lemma 4.3 it is not difficult to show that c_0 and c_1 exist as we always set the color of p_3y to be different from p'_3y'. □

We note that in the proof of Theorem 4.10 we can use a copy of $D^*(x, x')$ instead of $D(x, y)$ for each variable in the instance of 3-SAT. In this case $D_{\mathcal{I}}$ in the proof of Theorem 4.10 becomes bipartite by the following easy observations.

Observation 1. *The following holds.*

- *$D(x, y)$ defined in Definition 4.2 is bipartite and x and y belong to different partite sets.*
- *$D^*(x, x')$ defined in Definition 5.1 is bipartite and x and x' belong to the same partite sets.*
- *$D_2(x, x')$ defined in Definition 4.4 is bipartite and x and x' belong to different partite sets.*
- *$Q(x, a_1, a_2)$ defined in Definition 4.6 is bipartite and x and a_2 belong to the same partite set, which is different to the partite set containing a_1.*
- *$S(x_1, x_2, x_3)$ defined in Definition 4.8 is bipartite and x_1, x_2 and x_3 belong to the same partite sets.*

References

[1] O. Amini, F. Havet, F. Huc, S. Thomassé, WDM and directed star arboricity, Combinatorics, Probability and Computing 19 (2010) 161–182.
[2] J. Bang-Jensen, G. Gutin, Digraphs: Theory, Algorithms and Applications, 2nd ed., Springer Verlag, London, 2009, 812 pp+xxii.
[3] J. Bang-Jensen, M. Kriesell, Disjoint directed and undirected paths and cycles in digraphs, Theoret. Comput. Sci 410 (2009) 5138–5144.
[4] J. Bang-Jensen, M. Kriesell, On the problem of finding disjoint cycles and dicycles in a digraph, Combinatorica 31 (2011) 639–668.
[5] J. Bang-Jensen, M. Kriesell, A. Maddaloni, S. Simonsen, Vertex-disjoint directed and undirected cycles in general digraphs (submitted for publication).
[6] J. Bang-Jensen, A. Yeo, Arc-disjoint spanning sub(di)graphs in digraphs, Theoret. Comput. Sci 438 (2012) 48–54.
[7] S. Fortune, J. Hopcroft, J. Wyllie, The directed subgraph homeomorphism problem, Theoret. Comput. Sci 10 (1980) 111–121.
[8] D. Gonçalves, P. Ochem, On star and caterpillar arboricity, Discrete Math. 309 (11) (2009) 3694–3702.
[9] S.L. Hakimi, J. Mitchem, E. Schmeichel, Star arboricity of graphs, Discrete Math. 149 (1–3) (1996) 93–98.
[10] L. Lovász, On graphs not containing independent circuits, Math. Lapok 16 (1965) 289–299.
[11] W. McCuaig, Intercyclic digraphs, in: Graph structure theory (Seattle WA, 1991), in: Comtemp. Math, vol. 147, Amer. Math. Soc, Providence, RI, 1993, pp. 203–245.
[12] N. Robertson, P.D. Seymour, Graph minors. XIII: the disjoint paths problem, J. Combin. Theory Ser. B 63 (1995) 65–110.