ELSEVIER



# **Discrete Applied Mathematics**



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

# Interval graphs, adjusted interval digraphs, and reflexive list homomorphisms

Tomás Feder<sup>a</sup>, Pavol Hell<sup>b,\*</sup>, Jing Huang<sup>c,1</sup>, Arash Rafiey<sup>d</sup>

<sup>a</sup> 68 Waverley Street, Palo Alto, CA 94301, United States

<sup>b</sup> School of Computing Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6

<sup>c</sup> Department of Mathematics and Statistics, University of Victoria, Victoria, B.C., Canada V8W 3R4

<sup>d</sup> Istituto Dalle Molle di Studi sull'Intelligenza Artificiale, 6928 Manno-Lugano, Switzerland

#### ARTICLE INFO

Article history: Received 15 April 2010 Received in revised form 11 April 2011 Accepted 23 April 2011 Available online 2 June 2011

Keywords: Interval graphs Interval digraphs Adjusted interval digraphs Forbidden structure characterizations Polynomial algorithms List homomorphism problems Dichotomy

# ABSTRACT

Interval graphs admit linear-time recognition algorithms and have several elegant forbidden structure characterizations. Interval digraphs can also be recognized in polynomial time, and they admit a characterization in terms of incidence matrices. Nevertheless, they do not have a known forbidden structure characterization or low-degree polynomial-time recognition algorithm.

We introduce a new class of 'adjusted interval digraphs'. By contrast, for these digraphs we exhibit a natural forbidden structure characterization, in terms of a novel structure which we call an 'invertible pair'. Our characterization also yields a low-degree polynomial-time recognition algorithm of adjusted interval digraphs.

It turns out that invertible pairs are also useful for undirected interval graphs, and our result yields a new forbidden structure characterization of interval graphs. In fact, it can be shown to be a natural link proving the equivalence of some known characterizations of interval graphs—the theorems of Lekkerkerker and Boland, and of Fulkerson and Gross.

In addition, adjusted interval digraphs naturally arise in the context of list homomorphism problems. If H is a reflexive undirected graph, the list homomorphism problem LHOM(H) is polynomial if H is an interval graph, and NP-complete otherwise. If H is a reflexive digraph, LHOM(H) is polynomial if H is an adjusted interval graph, and we conjecture that it is also NP-complete otherwise. We show that our results imply the conjecture in two important cases.

© 2011 Elsevier B.V. All rights reserved.

# 1. Introduction

This is the full journal version of the conference note [10]. We include all proofs and provide additional connections and applications.

An interval graph [13] is a graph H which admits an interval representation, i.e., a family of intervals  $I_v$ ,  $v \in V(H)$ , such that  $uv \in E(H)$  if and only if  $I_u$  and  $I_v$  intersect. A digraph analogue has been defined in [25]—an interval digraph is a digraph H which admits an interval pair representation, i.e., a family of pairs of intervals  $I_v$ ,  $v \in V(H)$ , such that  $uv \in E(H)$  if and only if  $I_u$  intersects  $J_v$ . Interval graphs admit elegant characterizations [22,12], see [13], and linear-time recognition algorithms [1,15,4]. By contrast, the class of interval digraphs so far lacks comparable simple forbidden structure

E-mail addresses: tomas@theory.Stanford.edu (T. Feder), pavol@cs.sfu.ca (P. Hell), jing@math.uvic.ca (J. Huang), arash@idsia.ch (A. Rafiey).

<sup>\*</sup> Corresponding author. Tel.: +1 778 782 3391; fax: +1 778 782 3045.

<sup>&</sup>lt;sup>1</sup> Tel.: +1 250 721 7447; fax: +1 250 721 8962.

 $<sup>0166\</sup>text{-}218X/\$$  – see front matter s 2011 Elsevier B.V. All rights reserved. doi:10.1016/j.dam.2011.04.016

characterizations, and the best algorithm for their recognition to date is a dynamic programming algorithm of complexity  $O(nm^6(n+m) \log n)$  [23]. Motivated by the study of list homomorphisms (as explained below), we introduce a new digraph analogue of interval graphs, and argue that it has much nicer properties than the usual interval digraphs. Indeed, we will prove a simple forbidden structure characterization, which yields a low-degree polynomial-time recognition algorithm.

An *adjusted interval digraph* is an interval digraph *H* that admits an interval pair representation  $I_v$ ,  $J_v$ ,  $v \in V(H)$ , in which the intervals  $I_v$  and  $J_v$  have the same left endpoint. Note that the definition of an interval graph implies that an interval graph is *reflexive* (each vertex has a loop). Interval digraphs in the classical sense may lack loops. (If the intervals  $I_v$ ,  $J_v$  are disjoint there is no loop at v.) However, an adjusted interval digraph must again be reflexive. In [5], we studied the special case of adjusted interval digraphs *H* representable by intervals  $I_v$ ,  $J_v$ ,  $v \in V(H)$ , in which each interval  $J_v$  is just one point. These are called *chronological interval digraphs* [5], and we have shown that they can be characterized by the absence of certain special forbidden structures. In [24], a related class of *interval catch digraphs* has been characterized by the absence of certain other forbidden structures.

Here, we provide a forbidden structure characterization of adjusted interval digraphs. The forbidden structure is described in terms of a novel mechanism of "invertible pairs". Although invertible pairs may appear technical at first, we demonstrate they are a natural technique for describing obstructions to interval graphs and digraphs. In particular, we derive a characterization of undirected interval graphs in terms of invertible pairs, and exhibit its equivalence with other well-known characterizations of interval graphs, in terms of induced cycles and asteroidal triples [22], or in terms of a consecutive clique enumerations [12].

The presence of invertible pairs can be detected by an obvious simple algorithm implied by the definition. Thus our characterization directly implies a simple polynomial-time recognition algorithm for the class of adjusted interval digraphs.

Each digraph *H* is associated with two related undirected graphs. We denote by U(H) the *underlying graph* of *H*, which has an edge uv whenever  $u \neq v$  and  $uv \in E(H)$  or  $vu \in E(H)$ , and by S(H) the symmetric graph of *H*, which has an edge uv whenever  $u \neq v$  and  $uv \in E(H)$  and  $vu \in E(H)$ . Note that the loops of *H*, if any, are removed from both U(H) and S(H).

Adjusted interval digraphs are also motivated by the study of list homomorphisms. A homomorphism f of a digraph G to a digraph H is a mapping  $f : V(G) \rightarrow V(H)$  in which  $f(u)f(v) \in E(H)$  whenever  $uv \in E(G)$  [18]. If  $L(v), v \in V(G)$ , are lists (subsets of V(H)), then a list homomorphism of G to H (with respect to the lists L) is a homomorphism satisfying  $f(v) \in L(v)$  for all  $v \in V(G)$ . The list homomorphism problem LHOM(H) asks whether or not an input digraph G equipped with lists L admits a list homomorphism  $f : G \rightarrow H$  with respect to L. The complexity of the list homomorphism problem LHOM(H) for undirected graphs H has been classified in [6,8,9].

Of particular interest for this paper is the classification in the special case of reflexive graphs.

#### **Theorem 1.1** ([6]). Let H be a reflexive graph.

*If H is an interval graph, then problem LHOM (H) is polynomial-time solvable. Otherwise, problem LHOM(H) is NP-complete.* 

The complexity of LHOM(H) for general relational structures (including digraphs) H has been classified in [2]. In the special case of digraphs, a new forbidden structure characterization is given in [20]. The forbidden structure is called a *digraph asteroidal triple*, or DAT; see Theorem 5.1. This result also yields a simplified useful form of the characterization from [2] restricted to digraphs; see Theorem 5.2.

For reflexive digraphs H, we believe that LHOM(H) is polynomial precisely when H is an adjusted interval digraph. Specifically, we observe that each adjusted interval digraph H has polynomial-time solvable LHOM(H), and conjecture that, for any other reflexive digraph H, problem LHOM(H) is NP-complete. (This is an equivalent form of a conjecture from [11,17].) Note that both Theorems 5.1 and 5.2 classify reflexive digraphs H with polynomial LHOM(H), but the conjectured characterization is significantly more elegant, and does not follow from either of these theorems.

We observe that it suffices to verify the conjecture for digraphs whose underlying graphs are interval graphs. Then we proceed to verify it for digraphs whose underlying graphs are complete graphs and trees; these graphs can be viewed as the building blocks of interval graphs.

Thus it appears that, in the context of list homomorphisms, adjusted interval digraphs H play the same role for reflexive digraphs as interval graphs H play for reflexive graphs—namely, they exactly identify the tractable cases of LHOM(H).

# 2. Invertible pairs

Assume that u, v form an edge of the digraph H, i.e., that  $uv \in E(H)$  or  $vu \in E(H)$ . We say that uv is a forward edge if  $uv \in E(H)$ ; a backward edge if  $vu \in E(H)$ ; and a double edge if it is both a forward edge and a backward edge. We also say that a forward edge which is not double is a single forward edge, and similarly for a single backward edge. Since a loop is both a forward edge and a backward edge, we consider it a double edge. If  $uv \in E(H)$ , we say that u dominates v (and that v is dominated by u) in H, regardless of whether the forward edge uv is single or double.

We define two walks,  $P = x_0, x_1, ..., x_n$  and  $Q = y_0, y_1, ..., y_n$  in H, to be *congruent* if they follow the same pattern of forward and backward edges, i.e., if  $x_i x_{i+1}$  is a forward edge if and only if  $y_i y_{i+1}$  is a forward edge. If P and Q as above are congruent walks, we say that P avoids Q if there is no edge  $x_i y_{i+1}$  in the same direction (forward or backward) as  $x_i x_{i+1}$ .

An *invertible pair* in H is a pair of distinct vertices u, v such that

• there exist congruent walks *P* from *u* to *v* and *Q* from *v* to *u* such that *P* avoids *Q*,

• there exist congruent walks P' from v to u and Q' from u to v such that P' avoids Q'.

Note that it is possible that P' is the inverse of P and Q' is the inverse of Q, as long as both P avoids Q and Q avoids P. (The *inverse* of a walk  $P = x_0, x_1, \ldots, x_n$  is the walk  $x_n, x_{n-1}, \ldots, x_0$ .)

It will turn out to be useful to reformulate these definitions in terms of an auxiliary digraph. The *pair digraph*  $H^+$  associated with H has vertices  $V(H^+) = \{(u, v) : u \neq v\}$ , and edges (u, v)(u', v'), where

 $uu', vv' \in E(H)$  and  $uv' \notin E(H)$ , or  $u'u, v'v \in E(H)$  and  $v'u \notin E(H)$ .

We note that a directed walk in  $H^+$  from (u, v) to (v, u) yields two congruent walks: *P*, from *u* to *v*, and *Q*, from *v* to *u*, such that *P* avoids *Q*; and conversely, such walks *P* and *Q* yield a directed walk from (u, v) to (v, u) in  $H^+$ .

**Lemma 2.1.** Suppose that u, v is an invertible pair in H. Then (u, v) and (v, u) belong to the same strong component C of the pair digraph  $H^+$ . Moreover, for any (x, y) in C, the reversed pair (y, x) also belongs to C, and thus each pair (x, y) in C is invertible.

If *H* has no invertible pair, then, for each strong component *C* of *H*<sup>+</sup>, there exists a reversed strong component  $C' \neq C$  such that  $(x, y) \in C$  if and only if  $(y, x) \in C'$ .

**Proof.** These properties follow from the definition of a strong component and the observation that  $(u, v)(u', v') \in E(H^+)$  implies that  $(v', u')(v, u) \in E(H^+)$ . For instance, if  $(u, v), (v, u), (x, y) \in C$ , then a directed closed walk containing (u, v), (x, y) yields by reversal a directed closed walk containing (v, u), (y, x), and, by concatenation with the directed closed walk containing (u, v), (v, u), (v, x).  $\Box$ 

We now illustrate the concept of invertible pairs. Consider first the directed four-cycle 01, 12, 23, 30. Here, 0, 2 is an invertible pair, as we have congruent walks 012 and 230 which avoid each other. They correspond to the closed directed walk (0, 2), (1, 3), (2, 0), (3, 1), (0, 2) in  $H^+$ .

For a more complex example, consider the reflexive tree  $T_2$  from Fig. 1. (We denote the middle vertex *c*, so the edges of  $T_2$  are *aa*, *a'a'*, *bb*, *b'b'*, *cc*, *ac*, *a'c*, *cb*, *cb'*.) The pair *a*, *a'* is an invertible pair. Indeed, in  $H^+$ , (*a*, *a'*) dominates (*c*, *a'*), which in turn dominates (*b*, *a'*), which dominates (*b*, *c*), which finally dominates (*b*, *b'*). By the same token, according to the definition of  $H^+$ , we also have that (*b*, *b'*) dominates (*c*, *b'*), since *cb*, *b'b'* are edges of *H* but *b'b* is not. Similarly, (*c*, *b'*) dominates (*a'*, *b'*), which dominates (*a'*, *c*), which dominates (*a'*, *a*). We have obtained a directed walk (*a*, *a'*), (*b*, *a'*), (*b*, *c*), (*b*, *b'*), (*c*, *b'*), (*a'*, *b'*), (*a'*, *c*), (*a'*, *a*) in  $H^+$ , which corresponds to the two congruent walks *a*, *c*, *b*, *b*, *b*, *c*, *a'*, *a'* and *a'*, *a'*, *c*, *b'*, *b'*, *c*, *a* in *H*, where the first walk avoids the second one. By symmetry, we also have the walk (*a'*, *a*), (*c*, *a*), (*b'*, *a*), (*b'*, *a*), (*b'*, *c*), (*b'*, *b*), (*c*, *b*), (*a*, *b*), (*a*, *c*), (*a*, *a'*) in  $H^+$ .

As a last example, consider the undirected reflexive four-cycle 0, 1, 2, 3. We view an undirected graph as a symmetric digraph, with each undirected edge *xy*,  $x \neq y$ , replaced by the double edge *xy*, *yx*. Thus the reflexive four-cycle has the edges 00, 11, 22, 33, 01, 10, 12, 21, 23, 32, 30, 03. In this example, the pair 0, 2 is again invertible, but the closed walk (0, 2), (1, 3), (2, 0), (3, 1), (0, 2) used for the directed version does not qualify, since now, for example,  $03 \in E(H)$ . Nevertheless,  $H^+$  contains the closed walk (0, 2), (1, 3), (2, 3), (2, 0), (3, 1), (0, 2).

A min ordering of H is a linear ordering < of the vertices of H that satisfies the following property: if  $uv \in E(H)$  and  $u'v' \in E(H)$ , then min(u, u') min $(v, v') \in E(H)$ . (A min ordering has also been called an *X*-underbar enumeration [14,18].) In the case of reflexive digraphs, there is an equivalent simpler definition of a min ordering.

**Lemma 2.2.** Let *H* be a reflexive digraph. Then a linear ordering < of *V*(*H*) is a min ordering if and only if, for any three vertices i < j < k, we have

- $ik \in E(H)$  implies  $ij \in E(H)$ , and
- $ki \in E(H)$  implies  $ji \in E(H)$ .

**Proof.** The necessity of the two properties follows by taking the edge *ik* (respectively *ki*) and the loop at *j*. To see the sufficiency, consider edges xy, x'y' of *H* and assume without loss of generality that x < x', y' < y; thus  $\min(x, x') \min(y, y') = xy'$ . If x = y', then xy' is an edge, since *H* is reflexive. If x < y', then xy' is an edge, because of the triple x < y' < y. If y' < x, then xy' is an edge, because of the triple y' < x < x'.

**Corollary 2.3.** Let *H* be a reflexive digraph. A linear ordering of the vertices of *H* is a min ordering if and only if, for each vertex v, the vertices that follow v in the ordering consist of

- 1. first, the vertices that are adjacent to v by double edges,
- 2. second, the vertices that are adjacent to v by single edges, either all forward or all backward, and
- 3. last, the vertices that have no edges to or from v.

Of course, any of the three groups could be empty. Note that, in particular, in a min ordering of H it cannot be the case that a vertex v has both single forward and single backward edges towards vertices that follow it in the ordering. The following result relates min orderings to adjusted interval digraphs.

**Theorem 2.4.** A reflexive digraph is an adjusted interval digraph if and only if it admits a min ordering.

It is interesting to note that a reflexive undirected graph *H* has a min ordering if and only if it is an interval graph [6]. Thus Theorem 2.4 provides additional motivation in favour of adjusted interval digraphs.

**Proof.** Given a min ordering of a reflexive digraph H, we can arrange the common starting points of  $I_v$ ,  $J_v$  in the same order as the vertices v of H appear in the min ordering, and define intervals  $I_v$  and  $J_v$  as follows. The interval  $I_v$  ends at the point corresponding to the last vertex w such that vw is a forward edge, and the interval  $J_v$  ends at the point corresponding to the last vertex such that vw is a backward edge (i.e., wv is an edge of H). It is clear that H is the interval digraph corresponding to the adjusted interval representation  $I_v$ ,  $J_v$ ,  $v \in V(H)$ . Conversely, given an adjusted interval pair representation  $I_v$ ,  $J_v$ ,  $v \in V(H)$ , we obtain a min ordering of H according to the left to right order of the common left endpoints of the intervals.  $\Box$ 

According to Corollary 2.3, if v has no single forward edges towards later vertices, the interval  $I_v$  ends at the last vertex w such that vw is a double edge, and the interval  $J_v$  ends at the last vertex w such that vw is a backward edge. (Similarly, if v has no backward edges towards later vertices.)

Min orderings also play an important role for list homomorphism problems; see [14,18].

**Theorem 2.5.** If *H* admits a min ordering, then problem LHOM(*H*) is polynomial-time solvable.

Finally, we observe that an invertible pair is an obstruction to the existence of a min ordering.

Lemma 2.6. If H has an invertible pair, then H does not admit a min ordering.

**Proof.** Suppose that (u, v)(u', v') is an edge of the pair digraph  $H^+$ . Suppose that < is a min ordering of H, and suppose that u < v. Then we must also have u' < v'. Following the directed closed walk in  $H^+$  which contains (u, v) and (v, u), we obtain a contradiction.  $\Box$ 

### 3. Adjusted interval digraphs

In this section, we give our forbidden structure characterization of adjusted interval digraphs. This is the main result of our paper.

**Theorem 3.1.** A reflexive digraph H is an adjusted interval digraph if and only if it has no invertible pair.

In fact, we shall prove the following stronger result.

**Theorem 3.2.** The following statements are equivalent for a reflexive digraph H.

- 1. *H* is an adjusted interval digraph.
- 2. H has a min ordering.
- 3. *H* has no invertible pairs.
- 4. The vertices of  $H^+$  can be partitioned into sets D, D' such that
  - $(x, y) \in D$  if and only if  $(y, x) \in D'$ ,
  - $(x, y) \in D$  and (x, y) dominates (x', y') in  $H^+$  implies that  $(x', y') \in D$ ,
  - $(x, y), (y, z) \in D$  implies that  $(x, z) \in D$ .

**Proof.** It suffices to assume that *H* is weakly connected.

The equivalence of 1 and 2 is proved in Theorem 2.4. Furthermore, Lemma 2.6 shows that 2 implies 3. It is also quite straightforward to see that 4 implies 2; it suffices to define x < y if  $(x, y) \in D$ . Thus it remains to show that 3 implies 4.

Therefore, we assume that H has no invertible pair. Note that we may assume that H is weakly connected, otherwise we can order each weak component separately. Recall that, for each strong component C of  $H^+$ , there is a corresponding reversed strong component C' whose pairs are precisely the reversed pairs of the pairs in C; we shall say that C, C' are *coupled* strong components. Note that a strong component C may be coupled with itself: Lemma 2.1 implies that invertible pairs lie in *self-coupled* components.

The partition of  $V(H^+)$  into D, D' will correspond to separating each pair of coupled strong components C, C' of  $H^+$ . The vertices of one strong components will be placed in the set D, and their reversed pairs will go to D'. We wish to make these choices in such a way as to avoid creating a *circular chain* in D, i.e., a sequence of pairs  $(x_0, x_1), (x_1, x_2), \ldots, (x_n, x_0) \in D$ .

We shall proceed as follows. Initially, the sets *D* and *D'* are empty. We say that a strong component *C* of  $H^+$  is *ripe* when it has no edge *to* another strong component in  $H^+ - D$ . In the general step, we shall take a ripe component *C* and place it in *D*, and simultaneously place *C'* in *D'*. (Note that *C'* need not be ripe, but it has no edge *from* another strong component.)

We will show that there is always at least one ripe strong component which can be added to *D* without creating a circular chain.

The sets D, D' will always have the following properties (which are true initially). There is no circular chain in D; each strong component of  $H^+$  belongs entirely to D, D', or to  $V(H^+) - D - D'$ ; the pairs in D' are precisely the reversed pairs of the pairs in D; there is no edge of  $H^+$  from D to a vertex outside of D; and there is no edge of  $H^+$  from a vertex outside D' to a vertex in D'. At the end of the algorithm, each pair (x, y) with  $x \neq y$  will belong either to D or to D', and hence the final D will have no circular chain, and hence satisfy the transitivity property of 4.

We now prove that the algorithm maintains these properties.

Suppose, for a contradiction, that the current *D* has no circular chain, but that the addition of *C* to *D* creates a circular chain in  $C \cup D$ . Suppose that  $(x_0, x_1), (x_1, x_2), \ldots, (x_n, x_0)$  is a circular chain that has occurred for the first time during the execution of the algorithm, and also suppose that at that time no shorter circular chain has occurred. Since there are no invertible pairs, and since we never place both a pair and its reverse in *D*, we must have  $n \ge 2$ . We may assume without loss of generality that  $(x_n, x_0) \in C$ ; note that other pairs of the circular chain could also be in *C*.

*Case* 1. Assume that, in *H*, there is at least one edge between the vertices  $x_0, x_1, \ldots, x_n$ , say an edge  $x_a x_b$ .

We claim that this implies that *H* is complete on  $x_0, x_1, \ldots, x_n$ . We make the following elementary observations, assuming that  $j \neq i$ .

1. If  $x_i$  dominates  $x_i$ , then  $x_{i-1}$  dominates  $x_i$  in H.

2. If  $x_i$  dominates  $x_i$ , then  $x_i$  dominates  $x_{i-1}$  in H.

To prove the first observation, we note that, if  $x_j$  dominates  $x_i$  but  $x_{j-1}$  does not dominate  $x_i$  in H, then  $(x_{j-1}, x_j)$  dominates  $(x_{j-1}, x_i)$  in  $H^+$ . Since  $(x_{j-1}, x_j)$  is in  $C \cup D$ , the pair  $(x_{j-1}, x_i)$  must belong to  $C \cup D$ , implying a shorter circular chain in  $C \cup D$ .

To prove the second observation, we similarly note that, if  $x_j$  dominates  $x_i$  but  $x_j$  does not dominate  $x_{i-1}$  in H, then  $(x_{i-1}, x_i)$  dominates  $(x_{i-1}, x_j)$  in  $H^+$ , also implying a shorter circular chain.

Consider now the fact that  $x_a$  dominates  $x_b$  in H. Property (1) implies that  $x_{a-1}, x_{a-2}, \ldots, x_{b+1}$  all dominate  $x_b$ . Since  $x_{b+1}$  dominates  $x_b$ , property (2) implies that  $x_{b+1}$  dominates  $x_{b-1}, x_{b-2}, \ldots, x_{b+2}$ , i.e., it dominates all other vertices. At this point we use (1) again to derive that  $x_b$  dominates  $x_{b-1}$ , and repeated application of (2) as before implies that  $x_b$  dominates all other vertices. Continuing in this way, we see that each  $x_j$  dominates all other vertices, i.e., the vertices  $x_0, x_1, \ldots, x_n$  induce a complete graph in H.

We conclude the proof of Case 1 by showing that *C* is a trivial component (with a single vertex). If *C* has more than one vertex, then so does its corresponding coupled component *C'*, which contains the vertex  $(x_0, x_n)$ . Hence we assume for contradiction that  $(x_0, x_n)$  dominates some (a, b) not in  $C \cup D$ .

Up to symmetry, we may assume that  $x_0$  dominates a in H,  $x_n$  dominates b in H, and  $x_0$  does not dominate b in H. Since (a, b) is not in  $C \cup D$ , the pair  $(x_0, x_1)$ , which is in  $C \cup D$ , cannot dominate (a, b), which implies that  $x_1$  does not dominate b in H. If  $x_2$  dominates b in H, then  $(x_1, x_2)$  dominates  $(x_0, b)$  which dominates (a, b) in  $H^+$ ; this is impossible, as this is a directed path starting in C and ending outside  $C \cup D$ , so some edge would exit from  $C \cup D$  against the rules we maintain. Therefore  $x_2$  does not dominate b in H; if  $x_3$  dominates b in H, then  $(x_2, x_3)$  dominates  $(x_1, b)$  which dominates  $(x_0, b)$  which dominates (a, b), yielding the same contradiction. Therefore,  $x_3$  does not dominate b in H, and continuing in this way we would derive that  $x_n$  does not dominate b, which is false.

Thus we have  $C = \{(x_n, x_0)\}, C' = \{(x_0, x_n)\}$ . The same proof also shows that C' is ripe, as no (a, b) dominated by  $(x_0, x_n)$  can exist outside  $C \cup D$ . It is now easy to see that, if both  $(x_n, x_0)$  and  $(x_0, x_n)$  complete a circular chain with D, then D already had a circular chain.

*Case* 2. Assume that vertices  $x_0, x_1, \ldots, x_n$  are independent in *H*.

# **Lemma 3.3.** Let $x_0, x_1, \ldots, x_n$ be independent in *H*.

Suppose that *p* is a vertex of *H*, distinct from  $x_0, x_1, ..., x_n$ , which dominates  $x_{i+1}$  but not  $x_i$  (or which is dominated by  $x_{i+1}$  but not by  $x_i$ ).

Then  $(x_0, x_1), \ldots, (x_i, p), (p, x_{i+2}), \ldots, (x_n, x_0)$  is also a circular chain created at the same time.

**Proof.** We conclude from the assumption that  $(x_i, x_{i+1})$  dominates  $(x_i, p)$  in  $H^+$ , and, since  $(x_i, x_{i+1})$  is in  $C \cup D$ , we must also have  $(x_i, p)$  in  $C \cup D$ . Furthermore, since  $x_{i+1}$  does not dominate or is dominated by  $x_{i+2}$  in H, we also have  $(x_{i+1}, x_{i+2})$  dominating  $(p, x_{i+2})$ , whence  $(p, x_{i+2})$  is in  $C \cup D$ . In conclusion, we see that any such vertex p can replace  $x_{i+1}$  in the circular chain  $(x_0, x_1), (x_1, x_2), \ldots, (x_n, x_0)$ .

**Lemma 3.4.** Let vertices  $x_0, x_1, \ldots, x_n$  be independent in *H*.

- If p is a vertex of H, distinct from  $x_0, x_1, \ldots, x_n$ , which dominates  $x_i$  and  $x_k$  with  $j \neq k$ , then p dominates each  $x_i$ .
- If p is a vertex of H, distinct from  $x_0, x_1, \ldots, x_n$ , which is dominated by  $x_i$  and  $x_k$  with  $j \neq k$ , then p is dominated by each  $x_i$ .
- If p, distinct from  $x_0, x_1, ..., x_n$ , dominates  $x_j$  and is dominated by  $x_k$  with  $j \neq k$ , then p both dominates and is dominated by each  $x_i, i \neq j, k$ .

**Proof.** If *p* dominates  $x_{i+1}$  but not  $x_i$ , then Lemma 3.3 implies that *p* can replace  $x_{i+1}$  in the circular chain; however, at least one of  $x_j$ ,  $x_k$  is not equal to  $x_{i+1}$ , whence the vertices of the chain are not independent, and we conclude by Case 1. The other items are proved similarly.  $\Box$ 

We now claim that the circular chain  $(x_0, x_1), (x_1, x_2), \ldots, (x_n, x_0)$  has at most one pair, say  $(x_n, x_0)$ , in *C* (with all other pairs in *D*). Otherwise, assume that some  $(x_i, x_{i+1}), i \neq n$  is also in the strong component *C*, and let *P* be a directed path in *C* from  $(x_n, x_0)$  to  $(x_i, x_{i+1})$ . Let the penultimate pair on this path be (p, q), and, without loss of generality, assume that  $px_i, qx_{i+1} \in E(H), px_{i+1} \notin E(H)$ . (In the case  $x_ip, x_{i+1}q \in E(H), x_{i+1}p \notin E(H)$ , the argument is symmetric.) By Lemma 3.4, *p* does not dominate any  $x_j$  with  $j \neq i$ . Next, we claim that *q* does not dominate  $x_i$ . Indeed, if *q* dominates  $x_i$ , then Lemma 3.4 implies that *q* dominates all  $x_j$ . This is a contradiction, since it would mean that (p, q) dominates  $(x_i, x_{i+2})$  in  $H^+$ , implying that  $(x_i, x_{i+2})$  is in  $C \cup D$  and thus there is a shorter circular chain in *H*. Therefore, *q* does not dominate  $x_i$ . By a double application of Lemma 3.3, we conclude that we can replace  $x_i$  and  $x_{i+1}$  by *p* and *q* in the circular chain in *H*. Continuing in this way, we replace (p, q) by the previous pair on the path *P*, until we obtain the pair (p', q'), which is the first pair after  $(x_n, x_0)$ . Since  $x_0$  is adjacent to q', we are back in Case 1.

Thus the circular chain  $(x_0, x_1), (x_1, x_2), \ldots, (x_n, x_0)$  has only the pair  $(x_n, x_0)$  in *C*, and any circular chain in  $C \cup D$  has exactly one pair in *C*. We now suppose, in addition to the previous assumptions, that our circular chain minimizes the sum of the lengths of all distances amongst the vertices  $x_0, x_1, \ldots, x_n$ , in the underlying graph of *H*.

The digraph *H* turns out to have a very special structure. We claim that in this situation there exists a non-empty set *K* of vertices of *H* such that  $H \setminus K$  has weak components  $C_1, C_2, \ldots, C_m$ , where  $x_i \in C_i$ ,  $i = 1, 2, \ldots, n$ , and such that if  $p \in K$  dominates (respectively is dominated by) a vertex in  $C_i$ , then *p* dominates (respectively is dominated by) all vertices in  $C_i$ ; moreover, if  $x'_0, x'_1, \ldots, x'_n$  are any vertices with  $x'_i \in C_i$ , then  $(x'_0, x'_1), (x'_1, x'_2), \ldots, (x'_n, x'_0)$  is also a circular chain.

Indeed, we let *K* consist of all vertices of *H* that dominate each  $x_i$ , or are dominated by each  $x_i$ . It is easy to see that *K* must be non-empty, as Lemma 3.4 implies that any *p* dominated by  $x_j$ ,  $x_k$ ,  $j \neq k$  belongs to *K*. Such a *p* must exist by our new minimality assumption, as otherwise we could replace  $x_i$  by its neighbour *p* on a path joining  $x_i$  to  $x_k$  by Lemma 3.3.

The same argument shows that two different  $x_j$ ,  $x_k$  cannot lie in the same weak component  $C_i$  of  $H \setminus K$ , as any path joining  $x_j$  to  $x_k$  was shown to contain a vertex of K. Therefore, we can number the components so that  $C_i$  contains  $x_i$  for i = 1, 2, ..., n. (There may be additional components  $C_i$  with i = n + 1, ..., m.) Now, Lemma 3.3 implies that each  $x_i$  can be replaced by any neighbour in  $C_i$ ; thus any vertex of  $C_i$  can be taken as  $x_i$ . Thus, each  $p \in K$  that dominates a vertex in  $C_i$  also dominates all vertices in  $C_i$ , and similarly for vertices p dominated by a vertex in  $C_i$ .

This creates a situation where any pair (y, y') in the strong component *C* of  $H^+$  containing  $(x_n, x_0)$  must satisfy  $y \in C_n$ ,  $y' \in C_0$ . This easily implies that the strong component *C* does not have any edges entering it from the outside, and hence the strong component *C'* coupled with *C* is also ripe. We claim that *C'* cannot complete a circular chain with *D*. Otherwise, the pair  $(x_0, x_n)$  would also complete a circular chain by the same argument. Thus, both  $(x_0, x_n)$  and  $(x_n, x_0)$  complete a circular chain with *D*, whence *D* must already contain a circular chain, a contradiction.

Of course, if the addition of C' does not create a circular chain, then we add C' to D and C to D'.  $\Box$ 

This gives us a polynomially verifiable forbidden subgraph characterization of adjusted interval digraphs. As noted above, checking for invertible pairs amounts to computing the strong components of  $H^+$  and checking for the existence of a pair (u, v), (v, u) within one strong component. Thus, the recognition of adjusted interval digraphs is polynomial: one can, for instance, explicitly construct the pair digraph  $H^+$  in time  $O(m^2 + n^2)$  and test it for invertible pairs in the same time. There may, however, be more efficient ways.

We ask the following questions.

- 1. Is there a linear-time recognition algorithm for adjusted interval digraphs?
- 2. Are there natural intractable digraph problems that can be solved in polynomial time on the class of adjusted interval digraphs?

In the undirected case of interval graphs, the answer to both questions is yes [13].

# 4. Interval graphs

As we noted above, an undirected graph can be viewed as a digraph in which each undirected edge uv is replaced by the directed edges uv, vu. Equivalently, the definitions of invertible pair, min ordering, etc., can be read as written above, but interpreting edges as undirected pairs. Congruent walks become walks of equal length. Note, however, that  $H^+$  remains a digraph.

Theorem 4.1. A reflexive graph is an interval graph if and only if it has no invertible pairs.

**Proof.** A reflexive graph has a min ordering if and only if it is an interval graph [6]. An invertible pair is an obstruction to having a min ordering, according to Lemma 2.6. For a graph H without invertible pairs, viewing H as a digraph, Theorem 3.2 implies that it has a min ordering, whence it must be an interval graph.  $\Box$ 

Theorem 4.1 offers a nice link between two of the best known classical characterizations of interval graphs. An *asteroidal triple* in a graph *H* is a triple of vertices, such that any two are joined by a path that is disjoint from the neighbourhood of the third vertex. An enumeration of the maximal cliques of *H* is called a *consecutive clique enumeration* if the cliques containing any particular vertex are consecutive in the enumeration.

- 1. *H* has no asteroidal triple or a chordless cycle  $C_k$ , k > 3.
- 2. *H* has a consecutive clique enumeration.
- 3. *H* has no invertible pair.

**Proof.** An elegant proof of the fact that 1 implies 2 is given in [16], and we do not reprove it here.

To see that 2 implies 3, consider a consecutive clique enumeration  $K^1, K^2, \ldots, K^m$  of H, and an invertible pair u, v in H, with its associated walks P, Q, P', Q' (from the definition of invertible pair). Say, P is the walk  $u = u_0, u_1, u_2, \ldots, u_k = v$ ; Q is the walk  $v = v_0, v_1, v_2, \ldots, v_k = u$ ; P' is the walk  $v = u_k, u_{k+1}, \ldots, u_{k+k'} = u$ ; and Q' is the walk  $u = v_k, v_{k+1}, \ldots, v_{k+k'} = v$ . Let s = s(x, y) denote the superscript of a maximal clique  $K^s$  which contains the edge xy. Suppose without loss of generality that  $s(u_0, u_1) < s(v_0, v_1)$ . Then we must have  $s(u_0, u_1) < s(v_1, v_2)$ , since otherwise  $s(v_1, v_2) < s(u_0, u_1) < s(v_0, v_1)$ , with  $v_1$  in the first clique and the third clique, but not in the second clique, which is impossible. (Recall that  $u_0v_1 \notin E(H)$ .) A similar argument implies that  $s(u_1, u_2) < s(v_1, v_2)$ , and, continuing in this vein along the paths P, Q, we obtain  $s(u_k, u_{k+1}) < s(v_k, v_{k+1})$ . Comparing  $s(u_0, u_1) < s(v_0, v_1)$  and  $s(u_k, u_{k+1}) < s(v_k, v_{k+1})$ , we observe that  $u = u_0 = v_k$  lies in the cliques with subscripts  $s(u_0, u_1), s(v_k, v_{k+1})$ , but not  $s(v_0, v_1)$ , implying that  $s(v_k, v_{k+1}) < s(v_0, v_1)$ . Thus we were able to derive  $s(v_k, v_{k+1}) < s(v_0, v_1)$  from  $s(u_0, u_1) < s(v_0, v_1)$  along P, Q. Similarly,  $s(u_k, u_{k+1}) < s(v_k, v_{k+1})$ , which is a contradiction; therefore, 2 implies 3.

To see that 3 implies 1, suppose that u, v, w is an asteroidal triple. Let, for  $\{x, y, z\} = \{u, v, w\}$ , P(x, y) denote a path joining x and y which does not contain a neighbour of z, and let  $\ell(x, y)$  be the length of P(x, y). We will show that u, v form an invertible pair. Indeed, let P be the walk consisting of  $u, u, \ldots, u$ , of length  $\ell(v, w)$ , concatenated with P(u, v), and concatenated with the walk  $v, v, \ldots, v$ , of length  $\ell(w, u)$ , and let Q be the path P(v, w), concatenated with the walk  $w, w, \ldots, w$  of length  $\ell(u, v)$ , and concatenated with the path P(w, u). It is easy to see that P avoids Q and Q avoids P; hence u, v form an invertible pair. Since an induced cycle of length six or more contains an asteroidal triple, it remains to consider only the cycles  $C_4$ ,  $C_5$ , in which case an invertible pair is easily constructed along the same lines.  $\Box$ 

Note that the fact that H is reflexive is relevant for testing for invertible pairs, but irrelevant for asteroidal triples, chordless cycles, or consecutive clique enumerations. Both the characterization of Lekkerkerker and Boland [22] and that of Fulkerson and Gross [12] can now be derived from Theorems 4.1 and 4.2. Note, however, that the proof uses the result of Halin [16].

# 5. An application to the list homomorphism problem

We now illustrate the usefulness of the new class (of adjusted interval digraphs) in the context of list homomorphisms. A digraph asteroidal triple (DAT) in a digraph *H* consists of three vertices u, v, w and three invertible pairs of vertices s(u), b(u), s(v), b(v), and s(w), b(w), so that for any permutation x, y, z of u, v, w there exist walks *P* from x to s(x), Q' from y to b(x), and Q'' from z to b(x), such that *P* avoids both Q' and Q''. Example DATs can be seen in the proof of Corollary 5.10.

**Theorem 5.1** ([20]). Let H be any digraph. If H contains a DAT, then LHOM(H) is NP-complete; otherwise, LHOM(H) is polynomial-time solvable.

The min orderings defined above are a particular case of the following general concept. Let k be a positive integer. The k-th power of H is the digraph  $H^k$  with vertex set  $V(H)^k$  in which  $(u_1, u_2, \ldots, u_k)(v_1, v_2, \ldots, v_k)$  is an edge just if each  $u_i v_i$  is an edge of H. A polymorphism of order k is a homomorphism of  $H^k$  to H. A polymorphism f is conservative if  $f(u_1, u_2, \ldots, u_k)$  always is one of  $u_1, u_2, \ldots, u_k$ . From now on, we shall use the word polymorphism to mean a conservative polymorphism.

A polymorphism f of order two is *commutative* if f(u, v) = f(v, u) for any u, v. If H admits a min ordering <, then clearly defining  $f(u, v) = \min(u, v)$  is a polymorphism, which is commutative.

A polymorphism  $f : H^3 \to H$  is called a *majority polymorphism* if f(u, u, v) = f(u, v, u) = f(v, u, u) = u for any u, v. A ternary polymorphism  $f : H^3 \to H$  is *majority over a*, *b* if f(a, a, b) = f(a, b, a) = f(b, a, a) = a, f(b, b, a) = f(b, a, b) = f(a, b, b) = b.

In proving Theorem 5.1 in [20], we have also obtained the following classification of LHOM(H) (a simplification of the result from [2]). Recall that by our definition each polymorphism is conservative.

# **Theorem 5.2** ([20]). Let H be any digraph.

*If, for every pair of vertices a, b of H, there exists a polymorphism of H which either is ternary and majority over a, b, or is binary and commutative over a, b, then LHOM(H) is polynomial-time solvable.* 

Otherwise, if some pair of vertices a, b does not admit either of these polymorphisms, then problem LHOM(H) is NP-complete.

The following fact follows directly from Theorems 2.5 (or Theorem 5.2) and 2.4.

**Theorem 5.3.** If *H* is an adjusted interval digraph, then LHOM(*H*) is polynomial-time solvable.

We conjecture that the converse also holds. (This is an equivalent form of a conjecture from [11,17].)

## **Conjecture 5.4.** If a reflexive digraph H is not an adjusted interval digraph, then LHOM(H) is NP-complete.

We note again that both Theorems 5.1 and 5.2 give criteria for the NP-completeness of LHOM(H). However, the conjecture proposes a much cleaner criterion. In particular, note that Theorem 5.1 says that the existence of a DAT in H implies that LHOM(H) is NP-complete. A DAT consists of some invertible pairs, together with much additional and complex structure. The conjecture claims that the simple existence of an invertible pair in H suffices to imply the NP-completeness of LHOM(H).

We also had a similar conjecture for irreflexive digraphs [11,17]. However, that conjecture has turned out to be false [19,3], and we shall discuss the case of irreflexive digraphs in [19]. We are currently working, with Carvalho, on a possible approach to Conjecture 5.4, [3].

In the remainder of the paper, we shall verify Conjecture 5.4 in two important cases, namely for semi-complete digraphs, and for trees. We begin by deriving a useful tool.

**Theorem 5.5.** Let *H* be a reflexive digraph. If *H* satisfies any of the conditions below, then LHOM(*H*) is NP-complete.

- *H* contains a directed reflexive three-cycle  $C_3$ .
- *U*(*H*) is not an interval graph.
- *S*(*H*) is not an interval graph.

These results can be deduced from Theorem 5.1. In particular, it is explicitly stated in [20] that, in a reflexive digraph H, an (undirected) asteroidal triple in U(H) yields a DAT in H. We also cite [7,21] for the case when U(H) contains a chordless cycle of length greater than three. This proves the second fact. In all other cases, DATs are not difficult to find. (The first fact was also shown in [11], and the last fact in [6].)

Thus we may restrict our attention to reflexive digraphs H for which both S(H) and U(H) are interval graphs, and moreover such that H does not contain an induced reflexive three-cycle. The most basic interval graphs are the complete graphs and certain trees (that is, caterpillars). These are the two classes of reflexive digraphs for which we shall verify Conjecture 5.4.

A digraph is *semi-complete* if its underlying graph is complete. A digraph is a *tree* if its underlying graph is a tree in the usual sense. We first verify Conjecture 5.4 for semi-complete digraphs.

**Theorem 5.6.** Suppose that H is a reflexive semi-complete digraph. If H contains an invertible pair, then LHOM (H) is NP-complete.

**Proof.** We will show that, if there exist invertible pairs in *H*, then some invertible pair *a*, *b* admits no polymorphism, as prescribed by Theorem 5.2.

It turns out that some structures in *H* limit our choices of polymorphisms from the theorem. The first such structure is an invertible pair; this is easy to see from the definition of an invertible pair, and we state it without proof.

**Lemma 5.7.** No binary polymorphism of *H* can be commutative over an invertible pair.

Let *R* be the reflexive digraph  $V(R) = \{a, b, c\}$  and  $E(R) = \{aa, bb, cc, ab, bc, ac, ca\}$ .

Lemma 5.8. There is no polymorphism g on the digraph R that is a majority over a, b.

**Proof.** Suppose that g is a polymorphism of R which is a majority over a, b, i.e., g(a, a, b) = g(a, b, a) = g(b, a, a) = a, and g(a, b, b) = g(b, a, b) = g(b, b, a) = b. We claim that g must also be a majority over b, c. Note that  $g(c, c, b)g(a, a, b) = g(c, c, b)a \in E(R)$ . Hence, g(c, c, b) = c, as b does not dominate a in R. Similarly, g(c, b, c) = g(b, c, c) = c. Also,  $g(b, b, c)g(b, b, a) = g(b, b, c)b \in E(R)$ ; thus g(b, b, c) = b, and similarly g(b, c, b) = g(c, b, b) = b. Now, we can conclude that g is also a majority over a, c, using the fact that  $g(a, a, c)g(b, b, c) \in E(R)$  and  $g(b, b, c)g(c, c, a) \in E(R)$ .

Now we note that we have  $g(a, b, c)g(b, b, c) = g(a, b, c)b \in E(R)$ , which implies that  $g(a, b, c) \in \{a, b\}$  (since *c* does not dominate *b* in *R*); we have  $g(a, b, b)g(a, b, c) \in E(R)$ , which similarly implies that  $g(a, b, c) \in \{b, c\}$ ; and we have  $g(c, a, c)g(a, b, c) \in E(R)$ , which similarly implies that  $g(a, b, c) \in \{b, c\}$ ; and we have

We now proceed with the proof of Theorem 5.6. Assume that *H* has an invertible pair. According to Theorem 5.5, we may assume that *H* does not contain an induced reflexive three-cycle  $\vec{C}_3$ , and both S(H) and U(H) are interval graphs; in particular, S(H) does not contain an induced four-cycle. Finally, we may assume that, in any copy of *R* induced in *H*, the pair corresponding to *a*, *b* is not invertible. This directly follows from Lemmas 5.7 and 5.8, and Theorem 5.2.

Since *H* has invertible pairs, the pair digraph  $H^+$  has a self-coupled strong component *C*. According to Lemma 2.1, all pairs in *C* are invertible. We first note that, if (a, b) dominates (c, d) in *C*, then (a, b) also dominates (a, d), and (a, d) dominates (c, d); thus, we also have  $(a, d) \in C$ , and a, d is an invertible pair as well.

Consider a closed directed walk  $W = (x_0, y_0), (x_1, y_1), \dots, (x_n, y_n), (x_0, y_0)$  in *C* that contains both (a, b), (b, a) for some  $a, b \in V(H)$ . All pairs  $x_i, y_i$  are invertible, and so are all pairs  $x_i, y_{i+1}$  (addition modulo n). In fact, the above argument shows that we may assume each  $(x_i, y_{i+1})$  to belong to W. Recall that, for each i, the edge from  $(x_i, y_i)$  to  $(x_{i+1}, y_{i+1})$  in  $H^+$  is due to the edges  $x_i x_{i+1}, y_i y_{i+1}$  in H, which could be forward or backward.

We first assume that for some *i* we have  $x_ix_{i+1}$ ,  $y_iy_{i+1}$  forward and  $x_{i+1}x_{i+2}$ ,  $y_{i+1}y_{i+2}$  backward. Without loss of generality, let us assume that i = 0, i.e., that  $x_0x_1$ ,  $x_2x_1$ ,  $y_0y_1$ ,  $y_2y_1 \in E(H)$  and  $x_0y_1$ ,  $y_2x_1 \notin E(H)$ . Since *H* is semi-complete, we must have  $y_1x_0$ ,  $x_1y_2 \in E(H)$ . If  $y_1x_1 \notin E(H)$ , then  $y_1$ ,  $x_0$ ,  $x_1$  are all distinct and  $y_1x_1 \in E(H)$ , whence either  $x_1x_0 \notin E(H)$ , yielding an induced  $C_3$  on  $y_1$ ,  $x_0$ ,  $x_1$ , or  $x_1x_0 \in E(H)$ , yielding an induced copy of *R* on the same vertices, with invertible pair  $x_1$ ,  $y_1$ ; both contradict our assumptions. Thus,  $y_1x_1 \in E(H)$ , and, by a symmetric argument focused on  $x_1$ ,  $y_1$ ,  $y_2$ , we also deduce that  $x_1y_1 \in E(H)$ . At this point, the absence of *R* on  $x_0$ ,  $x_1$ ,  $y_1$  implies that  $x_0x_1$  is a double edge, and similarly on  $x_1$ ,  $y_1$ ,  $y_2$  we conclude that  $y_1y_2$  is also a double edge. Consider now the pair  $x_0$ ,  $y_2$ : if  $x_0y_2$  is not an edge then we find an induced *R* on  $x_1$ ,  $y_2$ ,  $x_0$ . Therefore,  $x_0y_2$  is also a double edge; this means that *S*(*H*) contains an induced four-cycle  $x_0x_1y_1y_2x_0$ , which contradicts our assumptions.

It remains to consider the case when all edges  $x_i x_{i+1}$ ,  $y_i y_{i+1}$  are forward (or all backward). In this situation, we claim that all pairs  $x_i$ ,  $y_i$  form a double edge  $x_i y_i$ . (Indeed, if  $y_i x_i \notin E(H)$ , then  $x_i y_i \in E(H)$ , and we derive an induced  $\vec{C}_3$  or R on  $x_i$ ,  $y_i$ ,  $y_{i+1}$ ; and if  $x_i y_i \notin E(H)$ , then we consider instead  $x_j y_j$ , where  $x_j = y_i$ ,  $y_j = x_i$ , and proceed similarly.) This is impossible, as we have shown that the pairs  $(x_i, y_{i+1})$  may be assumed to be on the cycle, and they cannot form double edges, by assumption.  $\Box$ 

Thus the conjecture holds for semi-complete digraphs.

We now turn to trees. In this case, we will provide a direct proof, as we are able to describe exactly which trees H yield tractable problems LHOM(H).

It is well known [13] that a tree is an interval graph if and only if it is a *caterpillar*, i.e., the removal all leaves yields a path. Thus we want to decide which orientations of caterpillars yield adjusted interval digraphs. Let S(x) denote the set of leaves of H adjacent to the vertex  $x \in P$ . As usual, we refer to H as a tree, or star, etc., to mean that U(H) (without the loops) is a tree, or star, etc., respectively.

If *H* is a star, we shall define *H* to be a *good caterpillar* if it does not contain, as an induced subgraph, the tree  $T_2$  depicted below. If *H* is not a star, we define it to be a *good caterpillar* if it has a longest path  $P = v_0, v_1, \ldots, v_k, v_{k+1}$  satisfying the following conditions for all *i*. (Note that  $v_1, v_2, \ldots, v_k$  is the path *P*, and that  $v_0 \in S(v_1), v_{k+1} \in S(v_k)$ .)

1. If  $v_i v_{i+1} \in E(H)$ , then  $v_i v \in E(H)$ , for all  $v \in S(v_i) - v_{i-1}$ .

2. If 
$$v_{i+1}v_i \in E(H)$$
, then  $vv_i \in E(H)$ , for all  $v \in S(v_i) - v_{i-1}$ .

Note that, if  $v_i v_{i+1}$  is a double edge, then so are all  $v_i v, v \in S(v_i) - v_{i-1}$ . Observe that there are no restrictions on  $v_0$ , other than those arising from the restrictions on  $v_1$ . Indeed, all edges  $v_1 v$  for  $v \in S(v_1) - v_0$  must follow the direction of the edge  $v_1 v_2$  (forward, backward, or double)—with the possible exception of a single vertex v, which must be the vertex  $v_0$ . Thus, such a  $v_0$  can be chosen if and only if the restrictions on  $v_1$  have at most one exception. Similarly, there are no restrictions on  $v_{k+1}$ , other than those arising from the restrictions on  $v_k$ . All edges  $v_k v$  for  $v \in S(v_k)$  must follow the direction of the edge  $v_k v_{k+1}$ . It is easy to see that such a  $v_{k+1}$  can be chosen if and only if between  $v_k$  and  $S(v_k)$  there does not exist at the same time a single forward edge and a single backward edge. Finally, we note that the exceptional case, when H is a star, also conforms to the general definition; we have chosen to state it separately only for convenience.

**Theorem 5.9.** Let *H* be a reflexive digraph that is a tree. Then the following statements are equivalent.

1. *H* is a good caterpillar.

2. *H* is an adjusted interval digraph.

3. *H* has no invertible pair.

4. H does not contain, as an induced subgraph, any of the trees  $T_1, \ldots, T_7$ , or their reverses.

**Proof.** 1 implies 2 via Theorem 2.4, as a good caterpillar can be ordered starting from  $v_0$  and proceeding to  $v_1, v_2, \ldots, v_k$ , with listing the double edges of  $S(v_i) - v_{i-1}$  first, as suggested by Corollary 2.3. The definition of a good caterpillar ensures that the listing for  $S(v_i) - v_{i-1}$  can be chosen to end with  $v_{i+1}$ .

2 and 3 are equivalent by Theorem 3.1, and 3 implies 4 by inspection. (We have already shown that a, a' is an invertible pair in  $T_2$ . We leave it to the reader to find invertible pairs in the other trees.)

Theorem 2.4 allows us to derive 4 from 2: none of the forbidden subtrees allows a min ordering. To see this, in the trees  $T_1$ ,  $T_3$ ,  $T_4$ , focus on the vertices 0, 1, 2, and in the trees  $T_2$ ,  $T_5$ ,  $T_6$ ,  $T_7$ , focus on the vertices a, a', b, b'.

It remains to show that 4 implies 1. Thus, suppose that H is a reflexive tree which does not contain any of  $T_1-T_7$  or their reverses. Since H does not contain  $T_1$ , U(H) is a caterpillar. If H is a star, the conclusion now follows. Thus assume that H is not a star: when all leaves of H are removed we obtain a path P, say  $P = p, r, s, \ldots, y, z$ . We will prove that one of p, z can be chosen as  $v_1$  and the other as  $v_k$ . Suppose first that p cannot be chosen to satisfy the condition for  $v_1$ . Then, in S(p), there must be two vertices v, v' such that the edges pv, pv' do not follow the direction of the edge pr on P. If pr is a double edge, this means that pv, pv' are single edges. Since H does not contain  $T_3$ , both are forward (or both backward) edges. This implies that all edges  $pv, v \in S(p)$  follow the direction of pr, and thus p can be chosen to satisfy the condition for  $v_k$ . Similarly, if pr is a single (forward or backward) edge, p can be chosen as  $v_k$ , since H does not contain  $T_2$ . Therefore, each of p, z satisfies the condition for  $v_1$  or for  $v_k$ . Suppose next that neither p nor z satisfies the condition for  $v_1$ . Then each contains two single edges whose direction does not follow the direction of pr; this contradicts the fact that H does not contain  $T_5$  and  $T_6$  or their reverses. Similarly, the absence of  $T_7$  implies that at least one of p, z satisfies the condition for  $v_k$ . The absence of  $T_4$  (and its reverse) implies that each intermediate vertex  $r, s, \ldots, y$  of P satisfies the condition for  $v_i$  if its left or its right neighbour, we again obtain a contradiction as above with the fact that H does not contain the trees  $T_5, T_6, T_7$ .  $\Box$ 



Fig. 1. Minimal trees with invertible pairs. Dashed edges are optional; edges without direction can be forward, backward or double; dotted lines denote paths; loops are omitted.

We now prove the following dichotomy classification.

#### **Corollary 5.10.** *Let H be a reflexive digraph that is a tree.*

If H is a good caterpillar, then H has a min ordering and LHOM(H) is polynomial-time solvable. Otherwise, H contains one of the trees  $T_1-T_7$  in Fig. 1, or their reverses, as an induced subgraph, and LHOM(H) is NP-complete.

**Proof.** If *H* is a good caterpillar, the theorem implies that it has a min ordering, and hence LHOM(*H*) is polynomial-time solvable. Otherwise, the theorem implies that *H* contains  $T_1$ – $T_7$ . We now claim that for each reflexive digraph *H* containing one of the trees  $T_1$ – $T_7$ , the problem LHOM(*H*) is NP-complete.

If H contains  $T_1$ , then S(H) is not an interval graph, and hence LHOM(H) is NP-complete.

If *H* contains  $T_2$ , then *H* has a DAT on the vertices *a*, *a'*, *b*. Indeed, the walk  $a \leftarrow a \leftarrow a$  is congruent to and avoids both walks  $a' \leftarrow a' \leftarrow a'$  and  $b \leftarrow c \leftarrow a'$ . (Here we write *c* for the central vertex of  $T_3$ .) Similarly,  $a' \leftarrow a' \leftarrow a'$  is congruent to and avoids both  $a \leftarrow a \leftarrow a$  and  $b \leftarrow c \leftarrow a$ ; and  $b \rightarrow b$  is congruent to and avoids both  $a \rightarrow c$  and  $a' \rightarrow c$ . Finally, we observe that two pairs *a*, *a'* and *b*, *c* are both invertible: for *a*, *a'* we have already noted this in the illustration of invertible pairs below Lemma 2.1. For *b*, *c*, we observe that the walk  $b \rightarrow b \leftarrow c \leftarrow a \leftarrow a \leftarrow a$  is congruent to and avoids the walk  $c \rightarrow b' \leftarrow b' \leftarrow c' \leftarrow a'$ , and so, by symmetry, the pair *b*, *c* is invertible. Therefore, *H* contains a DAT, and LHOM(*H*) is NP-complete by [20]. Observe that even though the DAT is defined on the three vertices *a*, *a'*, *b*, all the vertices of  $T_2$ , including *b'*, are involved in the walks defining the DAT.

If *H* contains *T*<sub>3</sub>, then it has a DAT on 0, 1, 2. Specifically, the walk  $2 \leftarrow 2 \rightarrow 2$  is congruent to and avoids both walks  $0 \leftarrow 0 \rightarrow a$  and  $1 \leftarrow a \rightarrow a$ ; the walk  $1 \rightarrow 1 \rightarrow 1-1$  is congruent to and avoids both walks  $0 \rightarrow a \rightarrow b-2$  and  $2 \rightarrow 2 \rightarrow 2-2$ ; and the walk  $0 \leftarrow 0 \leftarrow 0-0$  is congruent to and avoids both walks  $1 \leftarrow a \leftarrow b-2$  and  $2 \leftarrow 2 \leftarrow 2-2$ . All three pairs 2, *a* and 1, 2 and 0, 2 are invertible, as can be seen from Lemma 2.1, using the fact that the walks  $a \leftarrow 0-0 \leftarrow 0 \rightarrow 0 \rightarrow a \leftarrow b-2 \leftarrow 2$  and  $2 \leftarrow 2-b \leftarrow a \rightarrow 1 \rightarrow 1 \leftarrow 1-1 \leftarrow a$  avoid each other. Note that this proof applies regardless of the direction(s) of the arc(s) between *b* and 2 (as suggested by the notation *b*-2).

Similar proofs apply to the trees  $T_4, \ldots, T_7$ . There is always a DAT with vertices 0, 1, 2 or *a*, *a'*, *b*. The details are technical but not difficult to find.  $\Box$ 

#### **Corollary 5.11.** Let *H* be a reflexive digraph that is a tree.

If H has an invertible pair, then LHOM(H) is NP-complete.

Thus Conjecture 5.4 also holds for trees.

### Acknowledgements

We are grateful to the anonymous referees for several helpful suggestions that improved the presentation.

#### References

- [1] K.S. Booth, G.S. Lueker, Testing for the consecutive ones property, interval graphs, and graph planarity using PQ-tree algorithms, J. Comput. Syst. Sci. 13 (1976) 335-379.
- A.A. Bulatov, Tractable conservative constraint satisfaction problems, ACM Trans. Comput. Logic 12 (4) (tentative) (in press).
- [3] C. Carvalho, Personal communication.
- [4] D.G. Corneil, S. Olariu, L. Stewart, The LBFS structure and recognition of interval graphs, SIAM J. Discrete Math. 23 (2009) 1905–1953.
- [5] S. Das, M. Francis, P. Hell, J. Huang, Recognition and characterization of chronological interval digraphs, manuscript 2011.
- [6] T. Feder, P. Hell, List homomorphisms to reflexive graphs, J. Combin. Theory B 72 (1998) 236–250.
- [7] T. Feder, P. Hell, The retraction and subretraction problems for reflexive digraphs, manuscript 2008.
- [8] T. Feder, P. Hell, J. Huang, List homomorphisms and circular arc graphs, Combinatorica 19 (1999) 487–505.
- [9] T. Feder, P. Hell, J. Huang, Bi-arc graphs and the complexity of list homomorphisms, J. Graph Theory 42 (2003) 61–80.
- [10] T. Feder, P. Hell, J. Huang, A. Rafiey, Adjusted interval digraphs, Electron. Notes Discrete Math. 32 (2009) 83-91.
- [11] T. Feder, P. Hell, K. Tucker-Nally, Digraph matrix partitions and trigraph homomorphisms, Discrete Appl. Math. 154 (2006) 2458–2469.
- [12] D.R. Fulkerson, O.A. Gross, Incidence matrices and interval graphs, Pacific J. Math. 15 (1965) 835-855.
- [13] M.C. Golumbic, Algorithmic Graph Theory and Perfect Graphs, Academic Press, New York, 1980.
- [14] W. Gutjahr, E. Welzl, G. Woeginger, Polynomial graph-colorings, Discrete Appl. Math. 35 (1992) 29-45.
- [15] M. Habib, R. McConnell, C. Paul, L. Viennot, Lex-BFS and partition refinement, with applications to transitive orientation, interval graph recognition and consecutive ones testing, Theoret. Comput. Sci. 234 (2000) 59-84.
- [16] R. Halin, Some remarks on interval graphs, Combinatorica 2 (1982) 297–304. [17] P. Hell, From graph colouring to constraint satisfaction: there and back again, in: Topics in Discrete Mathematics, in: Algorithms and Combinatorics Series, vol. 26, Springer Verlag, 2006, pp. 407-432.
- [18] P. Hell, J. Nešetřil, Graphs and Homomorphisms, Oxford University Press, 2004. [19] P. Hell, A. Rafiev, List homomorphisms to irreflexive digraphs, manuscript 2007.
- [20] P. Hell, A. Raffey, The dichotomy of list homorphisms for digraphs, in: ACM-SIAM SODA 2011, pp. 1703–1713.
  [21] B. Larose, Taylor operations on finite reflexive structures, Int. J. Math. Comput. Sci. 1 (2006) 1–26.
- [22] C.G. Lekkerkerker, J.C. Boland, Representation of a finite graph by a set of intervals on the real line, Fund. Math. 51 (1962) 45-64.
- [23] H. Mueller, Recognizing interval digraphs and interval bigraphs in polynomial time, Discrete Appl. Math. 78 (1997) 189–205.
- [24] E. Prisner, A characterization of interval catch digraphs, Discrete Math. 73 (1989) 285–289.
- [25] M. Sen, S. Das, A.B. Roy, D.B. West, Interval digraphs: an analogue of interval graphs, J. Graph Theory 13 (1989) 581–592.