Hereditarily hard $H$-colouring problems

Jørgen Bang-Jensen$^a$, Pavol Hell$^b$, Gary MacGillivray$^c$.*

$^a$Department of Mathematics and Computer Science, Odense University, Odense, Denmark
$^b$School of Computing Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6
$^c$Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4

Received 7 July 1993; revised 27 April 1994

Abstract

Let $H$ be a graph (respectively digraph) whose vertices are called 'colours'. An $H$-colouring of a graph (respectively digraph) $G$ is an assignment of these colours to the vertices of $G$ so that if $u$ is adjacent to $v$ in $G$, then the colour of $u$ is adjacent to the colour of $v$ in $H$. We continue the study of the complexity of the $H$-colouring problem: 'Does a given graph (respectively digraph) admit an $H$-colouring?'. For graphs it was proved that the $H$-colouring problem is NP-complete whenever $H$ contains an odd cycle, and is polynomial for bipartite graphs. For directed graphs the situation is quite different, as the addition of an edge to $H$ can result in the complexity of the $H$-colouring problem shifting from NP-complete to polynomial. In fact, there is not even a plausible conjecture as to what makes directed $H$-colouring problems difficult in general. Some order may perhaps be found for those digraphs $H$ in which each vertex has positive in-degree and positive out-degree. In any event, there is at least, in this case, a conjecture of a classification by complexity of these directed $H$-colouring problems. Another way, which we propose here, to bring some order to the situation is to restrict our attention to those digraphs $H$ which, like odd cycles in the case of graphs, are hereditarily hard; i.e., are such that the $H'$-colouring problem is NP-hard for any digraph $H'$ containing $H$ as a subdigraph. After establishing some properties of the digraphs in this class, we make a conjecture as to precisely which digraphs are hereditarily hard. Surprisingly, this conjecture turns out to be equivalent to the one mentioned earlier. We describe several infinite families of hereditarily hard digraphs, and identify a family of digraphs which are minimal in the sense that it would be sufficient to verify the conjecture for members of that family.

1. Introduction and preliminaries

Let $H$ be a graph (respectively digraph) whose vertices are called 'colours'. An $H$-colouring of a graph (respectively digraph) $G$ (or a homomorphism $G \rightarrow H$) is an assignment of these colours to the vertices of $G$ so that if $u$ is adjacent to $v$ in $G$, then...
the colour of \( u \) is adjacent to the colour of \( v \) in \( H \). Note that a \( K_n \)-colouring of a graph \( G \) is just the usual \( n \)-colouring of \( G \). This motivated the 'colouring' terminology (cf. [28]).

Let \( H \) be a fixed (directed) graph. The \( H \)-colouring problem is stated as follows.

**H-colouring (H-COL)**

*INSTANCE*: A (directed) graph \( G \).

*QUESTION*: Does there exist an \( H \)-colouring of \( G \)?

Each \( H \)-colouring problem is clearly in NP.

The complexity of the \( H \)-colouring problem for undirected graphs was completely determined by Hell and Nešetřil [16] (also see [22, 28, 29, 32]), who proved that the \( H \)-colouring problem is NP-complete whenever \( H \) contains an odd cycle, and is polynomial otherwise. This result was extended to infinite graphs of finite maximum degree in [25].

Attention has subsequently shifted to the \( H \)-colouring problem for digraphs [1–4, 11–13, 18–21, 23–27, 33], and edge-coloured graphs [5]. Several large families of digraphs have been completely classified, and many partial results are known. For example, \( H \)-colouring is polynomial for any digraph \( H \) which has a loop — simply colour all vertices of \( G \) by a vertex of \( H \) with a loop. If \( C_n \) denotes the directed cycle of length \( n \), then there is a straightforward polynomial algorithm solving \( C_n \)-COL; the correctness proof for that algorithm implies that \( G \) admits a \( C_n \)-colouring if and only if the net length (difference between the number of forward edges and the number of backward edges) of any cycle in \( G \) is divisible by \( n \) (cf. [15, 28]). If \( T \) is a tournament, then \( T \)-COL is polynomial if \( T \) is transitive or contains a unique directed cycle, and is NP-complete if \( T \) contains at least two directed cycles [3]. There is much other evidence which suggests that the presence of two directed cycles in \( H \) (often) makes \( H \)-colouring hard, while the problem is (usually) tractable whenever \( H \) has only one directed cycle [2, 3, 4, 11–13, 23–28]. On the other hand, there are digraphs without any directed cycles, e.g., oriented cycles and oriented trees, for which the \( H \)-colouring problem is NP-complete [12, 13, 18, 19, 27]. However, for connected (every two vertices are joined by an oriented walk) smooth digraphs (directed graphs without sources — vertices of in-degree zero — and sinks — vertices of out-degree zero), the following conjecture proposes a classification of \( H \)-colouring problems.

**Conjecture 1.1** (Bang–Jensen and Hell [2]). Let \( H \) be a connected smooth digraph. If \( H \) is homomorphically equivalent to a directed cycle, then \( H \)-COL is polynomial. Otherwise, \( H \)-COL is NP-complete.

For arbitrary directed graphs the situation seems more complex, as additions to \( H \) can result in the complexity of the \( H \)-colouring problem shifting from NP-complete to polynomial. For example, if \( H \) is the digraph constructed from \( C_4 \cup C_6 \) by identifying a vertex on each directed cycle, then \( H \)-COL is NP-complete [2, 13]. On the other hand, it is easy to see that a given digraph is \((H \cup C_2)\)-colourable if and only if it is
C₂-colourable, hence, \((H \cup C_2)\)-COL is polynomial. We propose to bring some order to the situation by restricting our attention to those digraphs \(H\) which, like odd cycles in the case of graphs, are hereditarily hard, i.e., are such that the \(H'\)-colouring problem is NP-hard for any digraph \(H'\) containing \(H\) as a subdigraph.

After introducing preliminaries in the remainder of this section, we define, in Section 2, the class of \(H\)-colouring problems we call 'hereditarily hard' and study some of their properties; we make a conjecture as to precisely which digraphs belong to this class. The concept of hereditary hardness is extended in Section 3. Some families of hereditarily hard colouring problems, consistent with the conjecture, are described in Section 4, and tools which enable the construction of other such families are described. Finally we describe, in Section 5, a set \(S\) of digraphs which has the property that a directed graph \(H\) satisfies the hypotheses of (the unproved part of) Conjecture 2.5 if and only if there is a digraph in \(S\) which admits a homomorphism to \(H\). It follows from our results that it would suffice to prove the conjecture for the directed graphs in \(S\).

We use the terminology of [6,10], subject to the exceptions and additions mentioned below.

The equivalent digraph of a graph \(G\) is the directed graph obtained from \(G\) by replacing each edge by two oppositely oriented directed edges. Similarly, if \(D\) is a directed graph, and for vertices \(u\) and \(v\) both of the directed edges \(uv\) and \(vu\) exist, we sometimes say that \(u\) and \(v\) are joined by a double edge, or an undirected edge; in our diagrams these are drawn as undirected edges. If a subdigraph \(H\) of a digraph \(D\) is the equivalent digraph of an undirected graph, it is referred to as an undirected \(H\).

The directed cycle of length \(n\) (\(n \geq 1\)) is denoted by \(C_n\). It is assumed to have vertex-set \(V(C_n) = \{0, 1, \ldots, n-1\}\), and edge-set \(E(C_n) = \{i(i+1)\mod n: i = 0, 1, \ldots, n-1\}\). Similarly, \(P_n\), the directed path of length \(n\), (\(n \geq 1\)) is assumed to have vertex-set \(V(P_n) = \{0, 1, \ldots, n\}\), and edge-set \(E(P_n) = \{i(i+1): i = 0, 1, \ldots, n-1\}\).

Directed graphs \(D\) and \(H\) are homomorphically equivalent if there exist homomorphisms \(D \to H\) and \(H \to D\). This defines an equivalence relation on the set of directed graphs; we denote the equivalence class of \(G\) by \([G]\). The set of equivalence classes is partially ordered by the order \(\prec_h\), where \([G]\prec_h[G']\) if there is a homomorphism \(G \to H\); we refer to this as the homomorphism order. If there is no danger of confusion we will write \(G \prec_h H\) for \([G]\prec_h[H]\). Digraphs belonging to minimal elements with respect to this order are called h-minimal digraphs. If \(D\) and \(H\) are homomorphically equivalent, then a digraph \(G\) is \(D\)-colourable if and only if it is \(H\)-colourable, and consequently the complexity of \(D\)-COL and \(H\)-COL is the same. Since every digraph \(H'\) is homomorphically equivalent to a unique (up to isomorphism) minimal subdigraph \(H\), called the core of \(H'\) [16,32], we can restrict our attention to those directed graphs which are not homomorphically equivalent to any proper subdigraph. A digraph \(H\) is a core (or retract-free [4,27], of a minimal graph [8,32]) if it is not homomorphically equivalent to any proper subdigraph (i.e., it is its own core).
In light of the definitions and discussion so far, the first (i.e., polynomial) part of Conjecture 1.1 is clearly true. The unproved part states that the $H$-colouring problem is NP-complete for every smooth core which is not a directed cycle.

For our NP-completeness proofs we will use the following reductions. Let $I$ be a fixed digraph, and let $u$ and $v$ be distinct vertices of $I$. The indicator construction with respect to $(I, u, v)$ transforms a given digraph $H$ into the digraph $H^*$, defined to have the same vertex set as $H$, and to have as the edge set all pairs $hh'$ for which there is a homomorphism of $I$ to $H$ taking $u$ to $h$ and $v$ to $h'$. The triple $(I, u, v)$ is called an indicator, and, if some automorphism of $I$ maps $u$ to $v$ and $v$ to $u$, it is called a symmetric indicator. (The result of the indicator construction with respect to a symmetric indicator can be defined to be an undirected graph [16]). Certain types of restricted indicators play an important role in the construction of hereditarily hard $H$-colouring problems (see Section 4). It is proved in [16] that if the $H^*$-colouring problem is NP-complete, then so is the $H$-colouring problem. To use this construction to prove NP-completeness, care must be taken to assure that $H^*$ has no loops, i.e., that no homomorphism of $I$ to $H$ can map $u$ and $v$ to the same vertex, because if $H^*$ has a loop then $H^*$-COL is polynomial. An indicator $(I, u, v)$ such that no homomorphism of $I$ to $H$ maps $u$ and $v$ to the same vertex is a good indicator (for $H$).

Let $J$ be a fixed digraph, with specified vertices $x$ and $j_1, j_2, \ldots, j_t$. The sub-indicator construction with respect to $(J, x, j_1, j_2, \ldots, j_t)$, and $h_1, h_2, \ldots, h_t$ transforms a given core $H$ with specified vertices $h_1, h_2, \ldots, h_t$, to its subdigraph $H^*$ induced by the vertex set $V^*$ defined as follows: Let $W$ be the digraph obtained from the disjoint union of $J$ and $H$ by identifying $j_i$ and $h_i$, $i = 1, 2, \ldots, t$. A vertex $v$ of $H$ belongs to $V^*$ just if there is a retraction of $W$ to $H$ which maps $x$ to $v$. The structure $(J, x, j_1, j_2, \ldots, j_t)$ is called a subindicator. If all of the vertices $j_1, j_2, \ldots, j_t$ are isolated, then the outcome of the subindicator construction is independent of the specified vertices $h_1, h_2, \ldots, h_t$. In this case we call $(J, x, j_1, j_2, \ldots, j_t)$ a free subindicator, and refer to the subindicator construction with respect to $(J, x, \text{free})$. It is proved in [16] that if $H$ is a core and the $H^*$-colouring problem is NP-complete, then so is the $H$-colouring problem.

Let $G$ be a digraph, and let $W$ be a walk in $G$. The net length of $W$, $nl(W)$, is equal to the number of 'forward' edges in $W$ minus the number of 'backward' edges in $W$ (an edge $xy$ is forward if $x$ precedes $y$ in $W$, otherwise it is backward). It is proved in [15] (and follows from our earlier remark) that a digraph $G$ does not admit a homomorphism to any directed cycle $C_n$ with $n > 1$ if and only if it possesses a collection of oriented cycles (equivalently: closed walks) $C^1, C^2, \ldots, C^m$ such that $\gcd\{nl(C^i)\} = 1$.

In this paper we will sometimes use the term NP-hard instead of NP-complete. This is because sometimes we are able to give a polynomial time Turing reduction to the problem in question, but unable to give a polynomial time transformation (cf. [10, pp. 113, 118–120]). Since any NP-complete problem can be considered to be NP-hard, the use of this terminology is justified. It should be noted, however, that all $H$-colouring problems are in NP, and that all of our NP-hard $H$-colouring problems are at least as hard as any other problem in NP. (Some authors would replace the term
NP-hard by the more descriptive (and cumbersome) term ‘NP-complete with respect to Turing reduction’.

2. The definition and some properties

Let $H$ be a graph (respectively digraph). We say that $H$-COL (and $H$) is hereditarily hard if $H'$-COL is NP-hard for every loopless graph (respectively digraph) $H'$ which contains $H$ as a subgraph. It follows from the main theorem in [16] that each undirected graph that contains an odd cycle is hereditarily hard. The proof that theorem can be viewed as having two main steps. First it is proved that if the graph $H$ contains an odd cycle, then there is a graph $H'$, which contains $K_3$, such that $H'$-COL polynomially transforms to $H$-COL. It is then proved that $H'$-COL is NP-complete whenever $H'$ contains $K_3$, i.e., that $K_3$ is hereditarily hard.

A polynomial extension of a digraph $H$ is a loopless digraph $H'$ such that $H'$-COL is polynomial and $H'$ contains $H$ as a subdigraph. Unless $P=NP$, a directed graph is hereditarily hard if and only if does not admit a polynomial extension. Note that if $H'$-COL is polynomial, then $H$ is a polynomial extension of itself. In particular then, every undirected bipartite graph has a polynomial extension. As a second example, consider an acyclic digraph $A$ such that $A$-COL is NP-hard (see [12, 13, 18, 9, 27]). A transitive tournament with the same number of vertices as $A$ is a polynomial extension of $A$.

Our first proposition describes an infinite family of hereditarily hard digraphs. More such families are described in Section 4.

**Proposition 2.1.** Let $H$ be the equivalent digraph of an undirected odd cycle. Then $H$ is hereditarily hard.

**Proof.** Suppose that $H$ is a subdigraph of $G$. Let $G^*$ be the undirected graph that results from applying the indicator construction with respect to $(C_2, 0, 1)$ to $G$. Since $G^*$ contains an odd cycle, $G^*$-COL is NP-complete. (According to our definitions, $G^*$ is actually the equivalent digraph of an undirected graph. This is not a problem, since it is clear that a graph $F$ admits a homomorphism to the underlying simple graph corresponding to $G^*$ if and only if the equivalent digraph of $F$ admits a homomorphism to $G^*$.) The result now follows. □

It follows that the equivalent digraph of any nonbipartite undirected graph is hereditarily hard.

The set of directed graphs is partially ordered with respect to inclusion, that is, $G <_i H$ if $G$ is a subdigraph of $H$. Minimal elements with respect to this order are called $i$-minimal digraphs. The set of hereditarily hard graphs is, by definition, an upper order ideal with respect to this order. The next result states that the set of equivalence classes of hereditarily hard digraphs is also an ideal with respect to the homomorphism order.
Lemma 2.2. Let $G$ be a fixed digraph. These are equivalent:

1. The $H$-colouring problem is NP-hard whenever $G \leq_{h} H$.
2. The $H$-colouring problem is NP-hard whenever $G \leq_{i} H$.

Proof. (1) $\Rightarrow$ (2): Assume (1), and suppose that there is a homomorphism of $G$ to $H$. Consider $H' = G \cup H$. Then $G$ is a subdigraph of $H'$, so the $H'$-colouring problem is NP-hard. Since $H$ and $H'$ are homomorphically equivalent, the $H$-colouring problem is also NP-hard.

(2) $\Rightarrow$ (1): The inclusion map $i: G \to H$ is a homomorphism. This completes the proof. $\Box$

It follows from Lemma 2.2 that if $H$ is hereditarily hard, then so is every $G$ which is a homomorphic image of $H$. In particular, if $[G]$ contains a hereditarily hard digraph, then every element of $[G]$ is hereditarily hard, and if some digraph in $[G]$ has a polynomial extension, then so does every element of $[G]$.

According to [16], the $i$-minimal hereditarily hard graphs are odd cycles, and the only $h$-minimal element is $[K_3]$. In the next two lemmas we prove that $i$-minimal hereditarily hard digraphs are connected and smooth.

Lemma 2.3. Let $H$ be a digraph with connected components $C_1, C_2, \ldots, C_n$. Then $H$ is hereditarily hard if and only if $C_i$ is hereditarily hard, for some $i \in \{1, 2, \ldots, n\}$.

Proof. ($\Rightarrow$) If $P = NP$, every digraph is hereditarily hard, so there is nothing to prove. Otherwise, assume that $P \neq NP$, and suppose $H$ is hereditarily hard but there is no $i \in \{1, 2, \ldots, n\}$ such that $C_i$ is hereditarily hard. Thus, for $i = 1, 2, \ldots, n$, $C_i$ has a polynomial extension $X_i$. But then the disjoint union $H' = H \cup X_1 \cup X_2 \cup \ldots \cup X_n$ is a polynomial extension of $H$ (because $H' = \bigcup_{1 \leq i \leq n} (C_i \cup X_i)$ and, for $i = 1, 2, \ldots, n$, the $(C_i \cup X_i)$-colouring problem is polynomial), a contradiction.

($\Leftarrow$) Obvious. This completes the proof. $\square$

Lemma 2.4. Let $v$ be a source (sink) of $H$. Then $H$ is hereditarily hard if and only if $H - v$ is hereditarily hard.

Proof. ($\Rightarrow$) If $P = NP$, there is nothing to prove. Assume that $P \neq NP$, and suppose that $H - v$ is not hereditarily hard; thus it admits a polynomial extension $G$. Let $G^* = G$ be the superdigraph of $G$ constructed by adding a new vertex $x$ and the edges $\{xy: y \in V(G)\}$. Then $G^*-COL$ is polynomial [3, 13]. Thus $H$ also has a polynomial extension. This contradiction proves the implication.

($\Leftarrow$) Obvious. This completes the proof. $\square$

Consider $D$ such that $D-COL$ is polynomial. If there is a homomorphism $H \to D$, then $D$ and $(H \cup D)$ are homomorphically equivalent, therefore the $(H \cup D)$-colouring problem is polynomial and $(H \cup D)$ is a polynomial extension of $H$. Hence any digraph
that admits a homomorphism to a directed cycle of length greater than one has a polynomial extension. Every smooth digraph which we know to have a polynomial extension admits a homomorphism to a directed cycle of length greater than one.

Assuming that $P \neq NP$, the above two lemmas imply that a hereditarily hard digraph must have a connected component with at least two directed cycles, since otherwise the digraph obtained by repeatedly applying Lemma 2.4 to remove sources and sinks is either empty or a disjoint union of directed cycles, and by Lemma 2.3 such a digraph is not hereditarily hard. (Also note that all of the digraphs covered by Conjecture 1.1 have a connected component with at least two directed cycles.) It follows from the above discussion that the $i$-minimal hereditarily hard digraphs are those connected smooth digraphs which do not admit a homomorphism to any directed cycle of length greater than one, and are critical in the sense that every proper subdigraph lacks at least one of these properties.

Let $H$ be a graph. Let $R(H)$, the reduction of $H$, be the result of applying the subindicator construction with respect to $(P_{|V(H)|}, |V(H)|, \text{free})$ to $H$ (a similar use of the subindicator construction appears in [1]). By this definition, $R(H)$ is unique. Furthermore, $R(H)$ is smooth. It is not difficult to see that the digraph $R(H)$ may also be derived from $H$ by iteratively deleting all sources and sinks, until a smooth digraph remains. By Lemma 2.4, $H$ is hereditarily hard if and only if $R(H)$ is hereditarily hard. The following conjecture proposes a complete classification of the hereditarily hard $H$-colouring problems.

**Conjecture 2.5.** Let $H$ be a connected digraph. If $R(H)$ does not admit a homomorphism to a directed cycle of length greater than one, then $H$ is hereditarily hard. Otherwise $H$ has polynomial extension.

The second statement is true since, as we have noted, any digraph which admits a homomorphism to a directed cycle of length greater than one has a polynomial extension.

The reduction of $H$ can be used to extend the implications of Conjecture 1.1 to digraphs which are not smooth. Since $R(H)$ is obtained from $H$ via the subindicator construction, it follows that if $R(H)$-COL is NP-hard, then so is $H$-COL. Thus the truth of the unproved of Conjecture 1.1 implies that if the core of $R(H)$ is not a directed cycle, then $H$-COL is NP-hard. It is, however, possible for the core of $R(H)$ to be a directed cycle (or even empty) and $H$-COL to be NP-hard (see [27]).

We will prove that, surprisingly, Conjectures 1.1 and 2.5 are equivalent.

Let $H$ be a smooth digraph; then $H$ has a directed cycle. Let $g$ be the directed girth (i.e., the length of a shortest directed cycle) of $H$. Since no directed cycle admits a homomorphism to a larger directed cycle, $H$ is not $C_n$-colourable for any $n$ greater than $g$. This, together with the observation that any directed graph is $C_1$-colourable, allows us to talk about the largest $d$ for which there is a homomorphism of $H$ to $C_d$. 
Lemma 2.6. Let \( H \) be a connected smooth digraph, and let \( d \) be the largest positive integer such that \( H \) is \( C_d \)-colourable. Assume \( H \) and \( C_d \) are not homomorphically equivalent. Let \( H* \) be the digraph obtained by applying the indicator construction with respect to \( (P_d, 0, d) \) to \( H \). Then \( H* \) is a smooth digraph with exactly \( d \) connected components, none of which admits a homomorphism to a directed cycle of length greater than one. Furthermore, if \( H \) is strong then so is each component of \( H* \).

Proof. If \( d = 1 \), \( H = H* \). Hence assume \( d > 1 \).

We are given that there is a homomorphism of \( H \) to \( C_d \), but \( H \) and \( C_d \) are not homomorphically equivalent; thus there is no homomorphism \( C_d \rightarrow H \). In particular, \( H \) has no directed \( d \)-cycle. Therefore \( H* \) is loopless.

Fix a \( C_d \)-colouring of \( H \). Since any two adjacent vertices of \( H* \) are joined in \( H \) by a directed path of length \( d \), they are assigned the same colour. Therefore \( H* \) has at least \( d \) connected components.

We now prove that \( H* \) has precisely \( d \) connected components \( H^0, H^1, \ldots, H^{d-1} \), where \( H^j \) is the subdigraph of \( H* \) induced by the set of vertices of \( H \) of colour \( j \). Without loss of generality, consider the case \( j = 0 \). Let \( u, w \) be distinct vertices of colour \( 0 \). Since \( H \) is connected, there exists a \((u, w)\)-path \( P \). Let \( v \) be the first vertex in \( P \) which is different from \( u \), and also has colour 0. Let \( Q \) be the \((u, v)\)-section of \( P \). It suffices to show that \( H* \) contains a \((u, v)\)-walk.

Let us call an intermediate vertex of \( Q \) a source of \( Q \) (resp. sink of \( Q \)) if it is the tail (resp. head) of two consecutive edges of \( Q \). We also call \( u \) a source of \( Q \) (resp. sink of \( Q \)) if it is the tail (resp. head) of the first edge of \( Q \). Similarly, \( v \) is called a sink of \( Q \) (resp. source of \( Q \)) if it is the head (resp. tail) of the last edge of \( Q \). Let \( s_0, s_1, \ldots, s_k \) be the list of sources and sinks of \( Q \) in the order they are encountered when traversing \( Q \) (thus \( u = s_0 \) and \( v = s_k \)). For \( i = 0, 1, \ldots, k-1 \), the \((s_i, s_{i+1})\)-section of \( Q \) is a directed path. For \( i = 1, 2, \ldots, k-1 \), define the vertex \( t_i \) such that if \( s_i \) is a sink of \( Q \), then there is a directed \((s_i, t_i)\)-path of length \( d - c_i \) (mod \( d \)), where \( c_i \) is the colour of \( s_i \), and if \( s_i \) is a source of \( Q \), then there is a directed \((t_i, s_i)\)-path of length \( c_i \) (mod \( d \)). The vertex \( t_i \) always exists since \( H \) is smooth, and, further, each vertex \( t_i \) is coloured 0. It is not difficult to see from the definitions that, for \( i = 1, 2, \ldots, k-1 \), the vertices \( t_i \) and \( t_{i+1} \) are joined in \( H \) by a directed path of length \( d \). Therefore they are adjacent in \( H* \). It is also clear that the edges \( u t_1 \) and \( t_{k-1} v \) exist (since \( u = s_0 \) and \( v = s_k \)). Thus there exists a \((u, v)\)-walk in \( H* \).

It follows that the set of vertices coloured 0 induces a connected component of \( H* \). Further, if \( H \) is strong, then the path \( P \) can be chosen to be a directed path, and so each component of \( H* \) is also strong.

We now show that no component of \( H* \) admits a homomorphism to a directed cycle of length greater than one. Again without loss of generality consider \( H^0 \) as above. Assume that \( u \) is a source of \( Q \), the argument being similar if \( u \) is a sink of \( Q \). By our choice of \( v \), the path \( Q \) has net length zero or \( d \). The vertices \( s_0, s_1, \ldots, s_k \) are alternately sources and sinks of \( Q \). Let \( T = u, t_1, t_2, \ldots, t_{k-1}, v \) be the derived walk in \( H* \). By definition of \( H* \), the \((u, t_i)\)-section of \( T \) has net length zero when \( i \) is even, and net length one when \( i \) is odd, for \( i = 1, 2, \ldots, k-1 \).
Suppose first that \( n_l(Q) = 0 \). Then among \( s_0, s_1, \ldots, s_k \) there is one more source of \( Q \) than sink of \( Q \). This implies that \( k \) must be even. Hence \( k - 1 \) is odd and \( n_l(T) = 0 \).

Now suppose that \( n_l(Q) = d \). Then among \( s_0, s_1, \ldots, s_k \) there are an equal number of sources and sinks of \( Q \). This implies that \( k \) is odd. Hence \( k - 1 \) is even and \( n_l(T) = 1 \).

Therefore every walk \( W \) in \( H \) whose origin and terminus are coloured 0 gives rise to a walk in \( H^0 \) with net length \((1/d) n_l(W)\).

By the definition of \( d, H \) contains a collection of oriented cycles \( W_1, W_2, \ldots, W_n \), such that \( \gcd\{n_l(W_i) : i = 1, 2, \ldots, n\} = d \). By the above argument, each of these gives rise to a closed walk \( W_{i_0} \) in \( H^0 \) such that \( n_l(W_{i_0}) = (1/d) n_l(W_i) \). Since \( \gcd\{W_{i_0} : i = 1, 2, \ldots, n\} = 1, H^0 \) does not admit a homomorphism to any directed cycle of length greater than one.

Finally, since \( H \) is smooth, every vertex is the origin of a directed path of length \( d \) and the terminus of a directed path of length \( d \). Hence each component of \( H^* \) is also smooth. This completes the proof. \( \Box \)

**Theorem 2.7.** Conjectures 1.1 and 2.5 are equivalent.

**Proof.** (2.5) \( \Rightarrow \) (1.1): Assume Conjecture 2.5 is true, and let \( H \) satisfy the hypotheses of Conjecture 1.1. If \( H \) does not admit a homomorphism to a directed cycle of length greater than one, there is nothing to prove. Hence assume that \( H \) admits a homomorphism to such a directed cycle. Let \( d \) be the largest positive integer such that there is a homomorphism of \( H \) to \( C_d \) (see the comment preceding Lemma 2.6 regarding the existence of \( d \)). The digraph \( H \) has no directed \( d \)-cycle, otherwise \( C_d \) and \( H \) would be homomorphically equivalent. Let \( H^* \) be the result of applying the indicator construction with respect to \((P_d, 0, d)\) to \( H \). Let \( H^0 \) be a connected component of \( H^* \). By Lemma 2.6 the digraph \( H^0 \) is smooth, and does not admit a homomorphism to a directed cycle of length greater than one; hence \( H^0 = R(H^0) \). Since we are assuming Conjecture 2.5 is true, \( H^0 \)-COL is hereditarily hard, and therefore \( H \)-COL is NP-hard.

(1.1) \( \Rightarrow \) (2.5): Assume that Conjecture 1.1 is true, and let \( H' \) satisfy the hypotheses of Conjecture 2.5. Let \( G' \) be a digraph that contains \( H' \), and let \( G \) be the core of \( G' \). It is not hard to see that \( G \) contains a homomorphic image \( H \) of \( H' \). Consider \( R(G) \). Since \( H \) is smooth, \( R(G) \) contains \( H \); hence \( R(G) \) is not homomorphically equivalent to a directed cycle. Moreover, since \( R(G) \) is smooth, it satisfies the hypotheses of Conjecture 1.1. Since we assuming Conjecture 1.1 is true, \( R(G) \)-COL is NP-hard. Therefore the \( G \)-colouring problem is also NP-hard. This completes the proof. \( \Box \)

**Corollary 2.8.** It suffices to prove Conjecture 1.1 for digraphs that do not admit a homomorphism to any directed cycle of length greater than one.

**Proof.** Suppose Conjecture 1.1 is true for all connected smooth digraphs that do not admit a homomorphism to a directed cycle of length greater than one. Let \( H \) be a connected smooth digraph, and let \( d \) be the largest positive integer such that there is
a homomorphism of $H$ to $C_\lambda$. Let $H^*$ be the result of applying the indicator construction with respect to $(P_d, 0, d)$ to $H$. By Lemma 2.6 the digraph $H^*$ has exactly $d$ connected components, none of which admit a homomorphism to a directed cycle of length greater than one. Let $K$ be a connected component of the core of $H^*$. Then $K$ is smooth and does not admit a homomorphism to a directed cycle of length greater than one. By hypothesis, $K$-COL is NP-hard. In [2] it is proved that if $D = A \cup B$ is retract-free, and $A$-COL is NP-hard, then $D$-COL is also NP-hard. Therefore the $H$-colouring problem is NP-hard. This completes the proof.

3. An extension of hereditary hardness

Let $H$ be a digraph that admits a homomorphism to a directed cycle of length $n$. Then $H$ has a polynomial extension, namely $H \cup C_n$, and so $H$-COL is not hereditarily hard unless $P = NP$. In this section we introduce a generalization of hereditary hardness that enables us to establish complexity theorems, similar to hereditary hardness, for some directed graphs which have a polynomial extension. Our strategy is to impose enough restrictions on the superdigraphs of $H$ to be considered so that, in this restricted family of digraphs, the presence of $H$ as a subdigraph of a digraph $G$ is sufficient for $G$-COL to be NP-hard. For example, let $H$ be the digraph constructed from the equivalent digraph of $K_3$ by subdividing every directed edge. Suppose $G$ is a superdigraph of $H$ that contains no directed two-cycle. Then $(C_4, 0, 2)$ is a good indicator. The result $G^*$ of applying the indicator construction with respect to $(C_4, 0, 2)$ to $G$ contains an undirected three-cycle. By Proposition 2.1, the $G^*$-colouring problem is NP-complete. Thus, if we restrict our attention to digraphs with no directed two-cycle, the $G$-colouring problem is NP-complete for any superdigraph $G$ of $H$. More examples are given in Section 4.

Motivated by the above discussion, we make the following definitions. Let $S$ be a set of directed graphs. A digraph $H$ is hereditarily hard with respect to $S$ if the following two conditions are satisfied: (i) if $G$ is a loopless digraph in $S$, then the $G$-colouring problem is NP-hard whenever $H$ is a subdigraph of $G$, and (ii) at least one such $G$ exists. A digraph $G$ is called a polynomial extension of $H$ with respect to $S$ if $G$ is in $S$, the digraph $H$ is a subdigraph of $G$, and the $G$-colouring problem is polynomial. It is clear that if $P \neq NP$, a digraph is hereditarily hard with respect to $S$ if and only if it has no polynomial extension with respect to $S$.

It follows from the definition that if $H$ is hereditarily hard, $H$-COL is NP-hard. This need not be the case if $H$ is hereditarily hard with respect to a collection of digraphs. For example, it is proved in [4] that any oriented odd cycle is hereditarily hard with respect to the set of 'partitionable' digraphs, while on the other hand, it is easy to construct oriented odd cycles $C$ for which the $C$-colouring problem is polynomial (e.g. see [12, 13, 17–20]).

Most of the results in Section 2 hold in this more general setting, although not necessarily for arbitrary sets (some statements may not make sense with respect to sets
defined to forbid some of the hypotheses). More specifically, Lemmas 2.2 and 2.3 hold for any set \( S \) such that \( G \cup H \) is in \( S \) whenever \( G \) and \( H \) are both in \( S \), and Lemma 2.4 is true with respect to the set of digraphs with no closed directed walk of length \( k \) (for some fixed \( k \)), among others. In all instances, the modifications needed to the proofs are minor, and the reader should have little difficulty adding the missing details.

4. Some families of hereditarily hard digraphs

The purpose of this section is to give some examples of the digraphs discussed in Sections 2 and 3. Although the focus is on hereditarily hard digraphs, we also give some examples of digraphs which are hereditarily hard with respect to \( L_k = \{ G : G \) has no closed directed walk of length \( k \} \).

Let \( S \) be a set of directed graphs. An \emph{hh-indicator with respect to} \( S \) is an indicator \((I, u, v)\) such that for every loopless digraph \( G \) that contains a homomorphic image of \( I \) in which \( u \) and \( v \) are identified, either \( G \) is not in \( S \), or \( G\)-COL is NP-hard. It follows from this definition that if \( G \) is in \( S \) and \( G^* \) has a loop, then \( G \) is hereditarily hard with respect to \( S \). An \emph{hh-indicator} is an indicator \((I, u, v)\) such that every loopless digraph \( G \) that contains a homomorphic image of \( I \) in which \( u \) and \( v \) are identified is hereditarily hard.

The importance of hh-indicators is illustrated in the following lemma.

**Lemma 4.1.** Let \((I, u, v)\) be an hh-indicator with respect to \( S \). Let \( H^* \) be the digraph that results from applying the indicator construction with respect to \((I, u, v)\) to \( H \). If \( H^* \) is hereditarily hard, then \( H \) is hereditarily hard with respect to \( S \).

**Proof.** Let \( G \in S \), and suppose \( H \) is a subdigraph of \( G \). Let \( G^* \) be the result of applying the indicator construction with respect to \((I, u, v)\) to \( G \). There are two possibilities, depending on whether \( G^* \) contains a loop. If \( G^* \) contains a loop, then \( G \) must contain a subdigraph which is a homomorphic image of \( I \) such that \( u \) and \( v \) map to the same vertex. Thus \( G\)-COL is NP-hard. Otherwise, \( G^* \) is a loopless digraph that contains the hereditarily hard digraph \( H^* \), so the \( G^*\)-colouring problem is NP-hard. Consequently \( G\)-COL is also NP-hard. This completes the proof. \( \square \)

Lemma 4.1 can be used to construct new hereditarily hard digraphs from old. For example, let \( H \) be the undirected three-cycle with \( V(H) = \{0, 1, 2\} \) and \( E(H) = \{[0, 1], [1, 2], [2, 0]\} \), and let \( 3 \) be a new vertex. Set \( I = \langle H - 01 \rangle + 03 \) (see Fig. 1). Then any homomorphic image of \( I \) in which the vertices 1 and 3 are identified is also an image of an undirected three-cycle. Hence, by Lemma 2.2, \((I, 1, 3)\) is an \emph{hh-indicator}. Let \( G = H \), and let \( G' \) be the digraph constructed by replacing each edge \( xy \) of \( G \) by a copy of \( I \), and identifying 1 with \( x \) and 3 with \( y \). The result of applying the indicator construction with respect to \((I, 1, 3)\) to \( G' \) is \( G \) (an undirected three-cycle). Hence \( G' \) is hereditarily hard.
The general procedure is as follows. Suppose $H$ is hereditarily hard, and let $wu$ be an edge of $H$. Let $v$ be a new vertex, and set $I = (H - wu) + wv$. Any homomorphic image of $I$ in which $u$ and $v$ are identified is also an image of $H$. Thus, by Lemma 2.2, $(I, u, v)$ is an $hh$-indicator. Now, let $G$ be hereditarily hard, and let $G'$ be the digraph obtained by replacing each edge $xy$ of $G$ by a copy of $I$, and identifying $u$ with $x$, and $v$ with $y$. The result of applying the indicator construction with respect to $(I, u, v)$ to $G'$ contains $G$. Hence $G'$ is hereditarily hard.

Lemma 4.1 can also sometimes be used to construct digraphs which are hereditarily hard with respect to a given set $S$ of directed graphs. As an example, we construct a digraph $H$ which is hereditarily hard with respect to $L_2 = \{G: G$ has no closed directed walk of length two$\}$. The procedure is analogous to the above. Since any loopless homomorphic image of $C_4$ in which vertices 0 and 2 are identified contains a directed two-cycle, $(C_4, 0, 2)$ is an $hh$-indicator with respect to $L_2$. Let $H^*$ be the undirected three-cycle, and let $H$ be the digraph obtained by replacing each (undirected) edge $xy$ of $H^*$ by a copy of $C_4$, identifying 0 with $x$ and 2 with $y$ (see Fig. 2). Since $H$ is in $L_2$, there exists a superdigraph of $H$ with the appropriate property. It is easy to verify that the result of applying the indicator construction with respect to $(C_4, 0, 2)$ to $H$ contains an undirected 3-cycle, which is hereditarily hard. Thus $H$ is hereditarily hard with respect to $L_2$. This procedure works in general. Let $(I, u, v)$ be an $hh$-indicator with respect to $S$, and let $H^*$ be a hereditarily hard digraph. If $H$ is the digraph which results from replacing every (directed) edge $xy$ of $H^*$ by a copy of $(I, u, v)$, and identifying the pairs of vertices $u, x$ and $v, y$, then the result of applying the indicator construction with respect to $(I, u, v)$ to $H$ contains $H^*$. It is, however, not clear that there is a superdigraph of $H$ that belongs to $S$. Suppose such a digraph $G$ exists. Let $G^*$ be the result of applying the indicator construction with respect to $(I, u, v)$ to $G$. If since $G^*$ has a loop, then $G$-COL is NP-hard, since $(I, u, v)$ is an $hh$-indicator. If $G^*$ is loopless, then since $G^*$ contains $H^*$, $G^*$-COL is NP-hard. Therefore $H$ is hereditarily hard with respect to $S$. 

![Fig. 1. An example hh-indicator.](image-url)
We now describe several infinite families of hereditarily hard digraphs. Each of these directed graphs in turn gives rise to a collection of infinite families of hereditarily hard digraphs (constructed as above, via Lemma 4.1), and also to the infinite family of hereditarily hard digraphs that contain it.

Let $n$ be an integer greater than or equal to three. The digraph $W_n$, the wheel with $n$ spokes, is defined to be the digraph constructed from $C_n \cup \{v\}$ by adding the undirected edges $\{[v, c] : c \in V(C_n)\}$. The digraph $W_4$ is shown in Fig. 3.

Theorem 4.2. If $n$ is not divisible by four, then $W_n$ is hereditarily hard.

Proof. Let $(I, u, v)$ be the symmetric indicator shown in Fig. 4 with $i = 0$. The digraph that results from identifying $u$ and $v$ is an undirected three-cycle. Thus any loopless homomorphic image of $(I, u, v)$ in which $u$ and $v$ are identified is also an image of an undirected three-cycle, and is therefore hereditarily hard. Since there is an automorphism of $I$ that exchanges $u$ and $v$, $(I, u, v)$ is a symmetric hh-indicator. The result
Fig. 4. A useful symmetric indicator.

Theorem 4.4. For any positive integer $i$, the digraph $X_i$ is hereditarily hard.

Proof. The argument is similar to Theorem 4.2. Let $(I, u, v)$ be the $hh$-indicator shown in Fig. 4 (the digraph that results from identifying $u$ and $v$ is an undirected $(2i+3)$-cycle). Let $X_i^*$ be the digraph which results from applying the indicator construction with respect to $(I, u, v)$ to $X_i$. It is not hard to check that $X_i^*$ contains the undirected $(2i+1)$-cycle $0, 2, 4, \ldots, 4i, 0$. Therefore $X_i^*$ is hereditarily hard, and the result follows from Lemma 4.1. \[\Box\]

Similarly, let $T_j$ be a digraph constructed from $C_{4j+2}$ by adding the edges $(0, 2j), (2j, 4j), \ldots, (2j+2, 0)$ and undirected odd paths between vertices and $2j+2$, $3$ and $2j+4, \ldots, 4j+1$ and $2j$. A prototype of $T_i$ is shown in Fig. 6.

Theorem 4.5. For any positive integer $j$, the digraph $T_j$ is hereditarily hard.

Proof. The proof is similar to the proofs of the previous two theorems. Let the longest of the undirected odd paths have length $2i+1$, and let $(I, u, v)$ be the $hh$-indicator $W_{n^*}$ of applying the indicator construction with respect to $(I, u, v)$ to $W_n$ is the undirected graph with edge-set $\{[x, y] : y - x \equiv 2 (\text{mod } n)\}$. If $n$ is odd, $W_{n^*}$ is an undirected $n$-cycle, and if $n \equiv 2 (\text{mod } 4)$ it is the union of two undirected $(n/2)$-cycles. Since undirected odd cycles are hereditarily hard (Proposition 2.1), the result follows from Lemma 4.1. This completes the proof. \[\Box\]
shown in Fig. 4. The result \( T_j^* \) of applying the indicator construction with respect to \( (I, u, v) \) to \( T_j \) contains the undirected odd cycle \( 0, 2j, 4j, \ldots, 2j + 2, 0 \), which is hereditarily hard. This completes the proof.

We now describe examples of digraphs which are hereditarily hard with respect to \( L_k = \{ G: G \text{ has no closed directed walk of length } k \} \). The set of digraphs with property \( L_2 \) is precisely the set of orientations of simple graphs.

Let \( G \) be a digraph. An edge \( uv \) of \( G \) is said to be bypassed if there is a vertex \( w \) such that the edges \( uw \) and \( wv \) exist. For any positive integer \( i \), let \( B_{2i+1} \) be any digraph constructed from \( C_{2i+1} \) by adding a bypass to at least one out of every \( i \) consecutive edges (of the \( (2i + 1) \)-cycle). The bypasses may use existing or new vertices.

**Theorem 4.6.** For any positive integer \( i \), any digraph \( B_{2i+1} \) is hereditarily hard with respect to \( L_{i+1} \).
Proof. Any homomorphic image of $C_{2i+2}$ in which vertices 0 and $i+1$ are identified contains a closed directed walk of length $i+1$. Hence $(C_{2i+2}, 0, i+1)$ is a symmetric $hh$-indicator with respect to $L_{i+1}$. Let $B^*$ be the result of applying the indicator construction with respect to $(C_{2i+2}, 0, i+1)$ to $B_{2i+1}$. Since, for any $x$, the directed $(x, x+i+1)$-path along $C_{2i+2}$ contains at least one bypassed edge, the undirected edge $[x+i+1, x]$ is present in $B^*$. Thus $B^*$ contains the undirected $(2i+1)$-cycle $0, i, 2i, ..., i+1, 0$. The results follows. □

The next result provides a method to construct digraphs that are hereditarily hard with respect to different sets than the ones considered so far. In Lemma 4.1 we showed that if the result of the indicator construction (with respect to an $hh$-indicator) is hereditarily hard, then $H$ is hereditarily hard with respect to a given set. Below we show that if the result of the indicator construction is a hereditarily hard digraph, then we can find a set $S$ of directed graphs such that $H$ is hereditarily hard with respect to $S$.

Lemma 4.7. Let $(I, u, v)$ be an indicator. Let $H^*$ be the result of applying the indicator construction with respect to $(I, u, v)$ to $H$. If $H^*$ is hereditarily hard, then $H$ is hereditarily hard with respect to $S = \{ G: G$ contains no homomorphic image of $I$ in which $u$ and $v$ are identified $\}$.

Proof. Since $H^*$ has no loops, $H$ is in $S$. Furthermore, $H$-COL is NP-hard. Let $G$ be a superdigraph of $H$. Let $G^*$ be the result of applying the indicator construction to $G$. Then two possibilities arise: either $G^*$ has a loop, in which case $G$ contains a homomorphic image of $I$ in which $u$ and $v$ are identified, or $G^*$ is hereditarily hard. That is, the $G$-colouring problem is NP-hard whenever $H$ is a subdigraph of $G$ and $G$ is in $S$. Therefore $H$ is hereditarily hard with respect to $S$. This completes the proof. □

We conclude this section with an example of this construction. Let $I$ be the four-vertex oriented path with edge set $ux, xy, vy$, and let $H$ be the digraph constructed from an undirected three-cycle by replacing each edge $xy$ with a copy of $I$ and identifying $u$ with $x$, and $v$ with $y$. The result $H^*$ of applying the indicator construction with respect to $(I, u, v)$ to $H$ is an undirected three-cycle, which is hereditarily hard. Thus $H$ is hereditarily hard with respect to $\{ G: G$ contains no homomorphic image of $I$ in which $u$ and $v$ are identified $\}$ = $\{ G: G$ has no transitive triple $\}$. Thus the $G$-colouring problem is NP-hard whenever the loopless directed graph $G$ contains $H$ and has no transitive triple.

5. The $h$-minimal digraphs

In Section 2 the $i$-minimal hereditarily hard digraphs were described, under the assumption $P \neq NP$. It would suffice to prove Conjecture 2.5 for the directed graphs in this class.
In this section we describe a family $S$ of directed graphs which contains the $h$-minimal digraphs, again, assuming $P \neq NP$. The set $S$ has the property that a digraph $D$ satisfies the conditions of Conjecture 2.5 if and only if some member of $S$ admits a homomorphism to $D$. By Lemma 2.2 it would suffice to prove Conjecture 2.5 for these digraphs. A set of $h$-minimal digraphs can be constructed by choosing from $S$ a representative from each minimal equivalence class with respect to the homomorphism order.

Let $D$ be a digraph, and let $C$ be a directed cycle in $D$. A vertex of $C$ is a vertex of attachment if it is adjacent with some vertex in $D - C$. If every strong component of $D$ is a vertex or a directed cycle, and every directed cycle has exactly one vertex of attachment, then $D$ is called singly attached. The set $S$ consists of all connected singly attached smooth digraphs which do not admit a homomorphism to a directed cycle of length greater than one. By definition, each member of $S$ satisfies the hypotheses of Conjecture 2.5.

A maximal strong component (resp. minimal strong component) of a digraph $G$ is a strong component $C$ of $G$ such that there exists no edge $dc$ (resp. $cd$), where $c$ is in $C$ and $d$ is in $G - C$. Every maximal strong component of a smooth digraph $G$ contains a directed cycle, as does every minimal strong component of $G$.

**Theorem 5.1.** Suppose $D$ satisfies the conditions of Conjecture 2.5. Then there is a digraph $H$ in $S$ which admits a homomorphism $H \rightarrow D$.

**Proof.** Without loss of generality $D = R(D)$, that is, $D$ is smooth and connected. Since $D$ is not homomorphically equivalent to a directed cycle of length greater than one, it has a collection $W^1, W^2, \ldots, W^t$ of oriented cycles such that $\gcd\{nl(W^i)\} = 1$. For $i = 1, 2, \ldots, t$, let $L^i$ be an oriented cycle such that there is a homomorphism $f_i$ of $L^i$ onto $W^i$. We can choose these cycles $L^i$ so that each of them has a source and a sink. From the definitions, $\gcd\{nl(L^i)\} = 1$. Let $M^1, M^2, \ldots, M^r$ (resp. $N^1, N^2, \ldots, N^s$) be a collection of directed cycles, one from each maximal (resp. minimal) strong component of $D$. For $i = 1, 2, \ldots, r$, let $m_i$ be a vertex on $M^i$ and, for $j = 1, 2, \ldots, s$, let $n_j$ be a vertex on $N^j$. The digraph $H$ is constructed from $M^1, M^2, \ldots, M^r, N^1, N^2, \ldots, N^s, L^1, \ldots, L^t$ by adding directed paths as follows: For $i = 1, 2, \ldots, t$, let $r$ be a source of $L^i$, or a sink of $L^i$. For $k = 1, 2, \ldots, r$, if there is a directed $(m_k, f_i(r))$-path of length $l$ in $D$, then add a path of length $l$ from $m_k$ to $r$ in $H$ (all added paths are disjoint, and add $l - 2$ new vertices to $H$). Similarly, for $j = 1, 2, \ldots, s$, if there is a directed $(f_j(r), n_j)$-path of length $l$ in $D$, then add a directed path of length $l$ from $r$ to $n_j$ in $H$. Observe that new directed cycles are created by this construction and, further, $H$ is in $S$. Moreover, since in the construction every vertex of $H$ corresponds to a vertex of $D$, there is a natural homomorphism $H \rightarrow D$. This completes the proof. □

**References**