Further Effects of Two Directed Cycles on the Complexity of H-Colouring

Jørgen Bang-Jensen
Department of Computer Science
University of Copenhagen
DK-2100, Denmark

Gary MacGillivray
Department of Mathematics and Statistics
University of Regina
Saskatchewan, Canada, S4S 0A2

Abstract. Let H be a digraph whose vertices are called colours. Informally, an H-colouring of a digraph G is an assignment of these colours to the vertices of G so that adjacent vertices receive adjacent colours. In this paper we continue the study of the H-colouring problem, that is, the decision problem "Does there exist an H-colouring of a given digraph G?". In particular, we prove that the H-colouring problem is NP-complete if the digraph H consists of a directed cycle with two chords, or two directed cycles joined by an oriented path, or is obtained from a directed cycle by replacing some arcs by directed two-cycles, so long as H does not retract to a directed cycle. We also describe a new reduction which yields infinitely many new infinite families of NP-complete H-colouring problems

1. Introduction

Let G and H be (directed) graphs. A *homomorphism* of G to H is a function $f : V(G) \to V(H)$ such that $f(x)f(y) \in E(H)$ whenever $xy \in E(G)$. Since a k-colouring of a graph G is a homomorphism of G to K_k, the term *H-colouring of G* has been employed to describe a homomorphism $G \to H$.

We study the H-colouring problem, which is described as follows.

$H - COL$ (H – colouring)

INSTANCE: A directed graph G.

QUESTION: Does ther exist an H-colouring of G?

Each H-colouring problem clearly belongs to NP.

The complexity of the H-colouring problem for undirected graphs was completely determined by Hell & Nešetřil [Hell & Nešetřil, 1990], who proved that the problem is NP-complete whenever H contains an odd cycle, and is polynomial otherwise.

Attention has subsequently shifted to attempting to classify directed graphs according to the complexity of the H-colouring problem (for a survey see [Bang-Jensen, 1989; MacGillivray, 1989]). In view of the fact that there are trees T for which $T - COL$ is NP-complete [Gutjahr et al., 1988], it seems that a complete classification may be difficult to accomplish. An important development is the following conjecture, due to Bang-Jensen and Hell, which proposes a classification for a large family of digraphs (the terminology is reviewed in the next section).

1.1. Conjecture [Bang-Jensen & Hell, 1988]. *Let H be a connected digraph in which each vertex has in-degree at least one and out-degree at least one. If H does not admit a retraction to a directed cycle, then the H-colouring problem is NP-complete. Otherwise the H-colouring problem is polynomial.*

(The last statement is easy to see, cf. the comment following the definition of retraction.) The conjecture is known to be true for several large classes of digraphs (cf. above).

In this paper we generalise a result of Maurer, Sudborough and Welzl [Maurer et al, 1981] (cf. Theorem 3.1), and also a result of Bang-Jensen and Hell [Bang-Jensen & Hell, 1988; Gutjahr et al., 1989], (cf. Theorems 3.3 and 3.8). Our results add to the list of sparse digraphs H with two directed cycles for which the H-colouring problem is NP-complete. In addition, we describe a new reduction (cf. Theorem 4.5) which provides infinitely many new infinite families of NP-complete H-colouring problems.

2. Preliminaries

The purpose of this section is to state our definitions and describe the reductions we use. The terminology is fairly standard. For terms not defined here the reader should consult [Bondy & Murty, 1976] or [Garey & Johnson, 1979].

Let D be a digraph and let W be a closed walk in D. The *net-length* of W, $nl(W)$, is the number of "forwards" arcs of W minus the number of "backwards" arcs of W. (An arc uv of W is *forwards* if u preceeds v on W, otherwise it is *backwards*).

A digraph D is *connected* if any two vertices are joined by an oriented path. If any two vertices of D are joined by a directed path, we say that D is *strongly connected* or *strong*.

Let u and v be vertices of a digraph D. If the arcs uv and vu are both present, we say that u and v are joined by a *double arc* or an *undirected edge*. We denote this situation by $[u, v]$.

A *source* of a digraph D is a vertex of in-degree zero. A *sink* of D is a vertex of out-degree zero. We call a digraph D *smooth* if it has no sources and no sinks.

We use $N^{+d}(x)$ (resp. $N^{-d}(x)$) to denote the set of all vertices u for which there is a directed (x, u)-path (resp. (u, x)-path) of length d. Thus $N^{+1}(x)$ is the out-neighbourhood of x, etc.

We use C_n to denote the directed cycle of length n. We assume $V(C_n) = \{0, 1, \ldots, n-1\}$, and $E(C_n) = \{i(i+1) : i = 0, 1, \ldots, n$, where addition is modulo $n\}$. Similarly, we use P_n to denote the directed path of length n. We assume $V(P_n) = \{0, 1, \ldots, n\}$, and $E(P_n) = \{i(i+1) : i = 0, 1, \ldots, n-1\}$.

The *directed girth* (resp. *directed odd girth*) of a digraph D is the length of its shortest directed cycle (resp. directed odd cycle). If D has no directed cycle (resp. directed odd cycle), we define the directed girth (resp. directed odd girth) to be infinite.

A *semi-complete digraph* is a directed graph such that, for all pairs of vertices u, v, at least one of the arcs uv or vu exists. That is, a semicomplete digraph is a digraph with a spanning tournament. The following theorem regarding the complexity of H-colouring by semi-complete digraphs was proved in [Bang-Jensen et al., 1988].

Theorem 2.1. *Let H be a semi-complete digraph. If H has two or more directed cycles, then the H-colouring problem is NP-complete. Otherwise (H is acyclic or unicyclic), the H-colouring problem is polynomial.*

Let H' be a directed graph and let H be a subdigraph of H'. A *retraction of H' to H* is a homomorphism r of H' to H such that $r(h) = h$ for all vertices h of H. If H' admits a retraction to H, we say that H is a *retract* of H'. A directed graph is *retract-free* (or a *core* [Hell & Nešetřil, 1990], or a *minimal graph* [Welzl, 1982]) if it does not admit a retraction to a proper subdigraph. Every directed graph H contains a unique (up to isomorphism) subdigraph C which is retract-free, and for which there is a retraction of H to C [Welzl, 1982]. Following [Hell & Nešetřil, 1990] we call C the *core* of H. If H is a retract of H', there are homomorphisms i: $H \to H'$ (the *inclusion*) and r: $H' \to H$ (a retraction); thus a given digraph is H'-colourable if and only if it is H-colourable. This allows us, when we choose, to restrict our attention to retract-free digraphs. In particular, as the C_n-colouring problem is polynomial for any positive integer n [Maurer et al. 1981], this proves the last statement in Conjecture 1.1 .

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 arc 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 the digraph H^* is loopless (*i.e.*, if no homomorphism of I to H can map u and v to the same vertex), it is called a *good indicator*. If some automorphism of I maps u to v and v to u, we say that the indicator (I, u, v) is *symmetric*. (The result of the indicator construction with respect to a symmetric indicator is the equivalent digraph of an undirected graph, and can be defined to be an undirected graph [Hell & Nešetřil, 1990].)

2.2. Lemma [Hell & Nešetřil, 1990]. *$H^* - COL$ polynomially transforms to $H - COL$.*

In applying Lemma 2.2 care must be taken to assure that H^* has no loops, i.e., that (I, u, v) is a good indicator. If H^* has a loop, then there is a polynomial time algorithm for H^*-colouring; map all vertices of G to a vertex with a loop.

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 *retract-free* digraph 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 *sub-indicator*. The digraph J is not required to be connected. If the vertices $j_1, j_2, \ldots, j_t$. are all isolated, the outcome of the sub-indicator construction is independent of the choice of $h_1, h_2, \ldots, h_t$. In this case we call $(J, x, j_1, j_2, \ldots, j_t)$ a *free sub-indicator* and, in order to reflect the independence of the specified vertices, refer to it as the sub-indicator construction with respect to $(J, x, free)$.

2.3. Lemma [Hell & Nešetřil, 1990]. *$H^- - COL$ polynomially transforms to $H - COL$.*

Similarly, let J be a fixed digraph with a specified arc xy and specified vertices $j_1, j_2, \ldots, j_t$. The *edge sub-indicator construction with respect to $(J, xy, j_1, j_2, \ldots, j_t)$, and* $h_1, h_2, \ldots, h_t$ transforms a given retract-free digraph H with specified vertices $h_1, h_2, \ldots, h_t$ into its subdigraph $H^\wedge$ induced by the arcs of H which are images of the arc xy under retractions of W (as defined above) to H. The structure $(J, xy, j_1, j_2, \ldots, j_t)$ is called an *edge sub-indicator*. A *free edge sub-indicator* is defined and denoted similarly to the above.

2.4. Lemma [Hell & Nešetřil, 1990]. *$H^\wedge - COL$ polynomially transforms to $H - COL$.*

We conclude this section by mentioning an NP-complete problem that will be used in our transformations.

NOT-ALL-EQUAL k-SAT ($k \geq 3$ fixed) [Schaefer, 1978]
INSTANCE: A set U of variables, and a collection C of clauses over U such that each clause $c \in C$ involves k variables.
QUESTION: Is there a satisfying truth assignment for C in which each clause contains at least one true literal and at least one false literal?
Comment: The problem remains NP-complete even if no clause contains a negated literal [Lovász, 1973]. In this case it is the problem of two-colouring a k-regular hypergraph.

3. Results

We begin by extending some work of Maurer, Sudborough and Welzl. Let $C_{n,k}$ be a digraph obtained from a directed n-cycle by replacing k arcs with double arcs. It has been proved [Maurer et al, 1981] that if n is odd, then $C_{n,1} - COL$ is NP-complete. When n is even and $k \geq 1$, the core of $C_{n,k}$ is a directed two-cycle (a double arc). Thus $C_{n,k} - COL$ is polynomial. A complete classification of the complexity of $C_{n,k} - COL$ is given below.

3.1. Theorem. *If n is even or $k = 0$, then $C_{n,k} - COL$ is polynomial. Otherwise (n is odd and $k > 0$) $C_{n,k} - COL$ is NP-complete.*

Proof: It remains to prove that if n is odd and $2 \leq k \leq n$, then $C_{n,k} - COL$ is NP-complete. Let C^* be the digraph that results from applying the indicator construction with respect to $(P_{n-2}, 0, n-2)$ to $C_{n,k}$. Since the directed odd girth of $C_{n,k}$ is n, the digraph C^* is loopless. Furthermore, each double arc of $C_{n,k}$ is also a double arc of C^*.

We claim that the vertices incident with double arcs induce a semi-complete digraph. Suppose $[u, v]$ and $[x, y]$ are distinct double arcs. The arcs of $C_{n,k}$ belonging to the directed n-cycle give rise to a directed (u, x)-path and a directed (x, u)-path. Moreover, exactly one of these paths has odd length. Since both u and x are incident with double arcs, this implies that there is either a directed (u, x)-walk of length $n-2$ or a directed (x, u)-walk of length $n-2$. Hence one of ux and xu is an arc of C^*. This proves the claim.

Let C^{*-} be the digraph which results from applying the sub-indicator construction with respect to $(C_2, 0, free)$ to C^*. Then C^{*-} is a semi-complete digraph with at least two directed cycles, and therefore, by Theorem 2.1, $C^{*-} - COL$ is NP-complete. Thus $C_{n,k} - COL$ is also NP-complete. This completes the proof. ∎

We now generalise the following result in two ways (cf. Theorems 3.3 and 3.8)

3.2. Theorem [Bang-Jensen & Hell, 1988; Gutjahr et al., 1989]. *Let H be a digraph of the form D_1 or D_2 (see figure 3.1). If H does not admit a retraction to a directed cycle, then $H - COL$ is NP-complete. Otherwise, $H - COL$ is polynomial.* ∎

Theorem 3.2 states, as a special case, that if D is a digraph constructed from a directed cycle by adding a chord, then $D - COL$ is NP-complete unless D admits a retraction to a directed cycle. That is, Conjecture 1.1 is true for directed cycles with one chord.

Let H be a directed graph constructed from a directed n-cycle by adding two chords. Then, depending on the relative orientation of the chords, H is of one of four types; an example of each type is shown in figure 3.2. We now prove that the H-colouring problem is NP-complete unless H retracts to a directed cycle. That is, Conjecture 1.1 is also true for directed cycles with two chords.

3.3. Theorem. *Let H be a directed graph that is constructed from a directed cycle by adding two chords. If H does not admit a retraction to a directed cycle, then $H - COL$ is NP-complete. Otherwise, $H - COL$ is polynomial.*

We have previously noted the second statement (cf. the paragraph following Theorem 2.1). The proof of the first statement is divided into four lemmas, depending on the type of H.

Note that, if H does not admit a retraction to a directed cycle then any retract of H is of the form D_1 or D_2. In view of Theorem 3.2, it may therefore be assumed

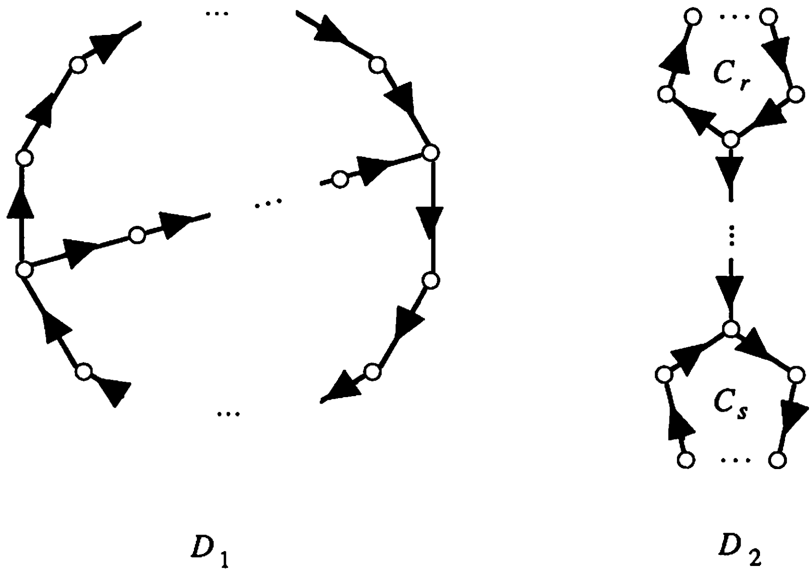

Figure 3.1. Digraphs with two directed cycles.

without loss that if H admits no retraction to a directed cycle, then H is retract-free.

3.4. Lemma. *If H is of type I and does not admit a retraction to a directed cycle, then $H - COL$ is NP-complete.*

Proof: Let H be of type I. Then H has exactly three directed cycles, say of lengths n, a, and b, respectively. Without loss of generality assume $n > a > b$. Suppose that H does not admit a retraction to a directed cycle. Then b does not divide both a and n. There are two cases to consider.

Case 1: b does not divide a.

Let H^- be the result of applying the sub-indicator construction with respect to $(C_a, 0, free)$ to H. Then H^- is the subdigraph of H induced by the vertex set of the directed a-cycle. Since b does not divide a, the digraph H^- does not admit a retraction to a directed cycle. Hence $H^- - COL$ is NP-complete by Theorem 3.2, and therefore $H - COL$ is also NP-complete.

Case 2: b divides a.

Since the directed b-cycle is not a retract of H, b does not divide n. Let $H^\wedge$ be the result of applying the edge sub-indicator construction with respect to $(C_{n+b}, 01, free)$ to H. It is clear that every arc, except the chord e that forms the directed a-cycle,

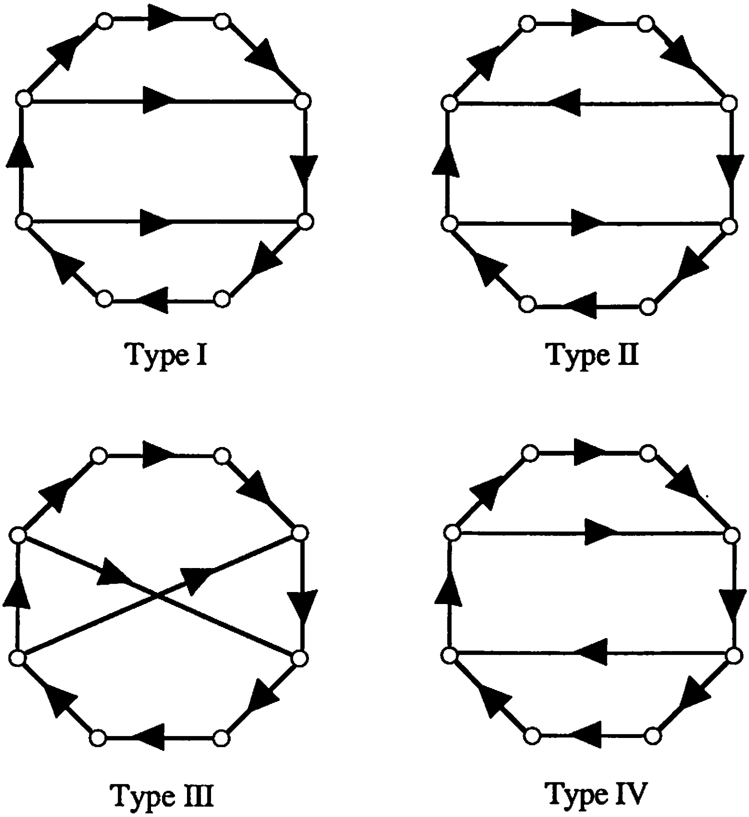

Figure 3.2. The four possible orientations of the chords.

belongs to a closed directed walk of length $n+b$. If e also belongs to such a closed walk, then there are integers a and b such that $n+b = \alpha a + \beta b = \gamma b$ (since b divides a). Therefore b divides n, which is a contradiction. Hence $H^{\wedge} = H - e$, and so $H^{\wedge} - COL$ is NP-complete by Theorem 3.2. Therefore $H - COL$ is also NP-complete.

All cases have been considered. ∎

3.5. Lemma. *If H is of type II and does not admit a retraction to a directed cycle, then $H - COL$ is NP-complete.*

Proof: Let H be of type II. The digraph H can have three or four directed cycles. When H has four directed cycles, the chords have both endpoints in common. In this case, H is also of type IV. We defer consideration of this case to Lemma 3.7. Hence assume H has exactly three directed cycles, say of length n, a, and b.

Without loss of generality $n > a \geq b$. Suppose that H does not admit a retraction to a directed cycle. Then b does not divide both a and n. Let the directed n-cycle be $0, 1, \ldots, n-1, 0$. There are four cases to consider.

Case 1: b does not divide a.

Let uv (resp. xy) be the chord that belongs to the directed a-cycle (resp. directed b-cycle).Without loss of generality assume $u \neq y$; otherwise consider the converse of H. Suppose first that the path from u to y on C_n has at least two edges. Let J be the directed graph constructed by identifying the terminal vertex of a directed path of length $n-a-1$ with a vertex on a directed a-cycle. Let 0 be the label of the initial vertex of the directed path. Let H^{-} be the result of applying the sub-indicator construction with respect to $(J, 0, \textit{free})$ to H. It is not hard to see that the core of H^{-} is of the form D_2. Since b does not divide a, $H^{-} - COL$ is NP-complete by Theorem 3.2. Hence $H - COL$ is also NP-complete. Suppose now that H contains the edge uy. If $x \neq v$, then by an argument similar to the one above we may assume that H also contains the edge xv. Now $n = a + b$, so a does not divide $n + b$, since $a > b$. Let J be the directed graph consisting of a directed cycle C_{n+b} and let e be a fixed edge of J. Let $H^{\wedge}$ be the result of applying the edge-sub-indicator construction with respect $(J, e, \textit{free})$ to H. Then $H^{\wedge} = H - uv$, since a does not divide $n + b$. Also, b does not divide n, since $n = a + b$ and b does not divide a. Now it follows from Theorem 3.2 that $H^{\wedge}$-colouring, and hence also H-colouring, is NP-complete. If $x = v$ then $n = a + b - 1$. Let H' be the result of applying the edge-sub-indicator construction with respect to $(C_{a+b}, e, \textit{free})$ to H, where e is a fixed edge of the cycle. Then $H' = H - uy$, and H'-colouring is NP-complete by Theorem 3.2. Hence H-colouring is NP-complete.

Case 2: b divides a, and $b < a$.

Since the directed b-cycle is not a retract, b does not divide n. Let e be the chord that belongs to the directed a-cycle. Every arc except e belongs to a closed directed walk of length $n + b$. If the arc e also belongs to such a closed directed walk, then either $n + a \leq n + b$, or a divides $n + b$. The former case is impossible since $a > b$, and the latter case is also impossible since, if b divides a, then b divides n. Let $H^{\wedge}$ be the result of applying the edge sub-indicator construction with respect to $(C_{n+b}, 0, \textit{free})$ to H. Then $H^{\wedge} = H - e$, where e is the chord of the n-cycle that forms the a-cycle. Since $H^{\wedge}$ is of the form D_1, and does not admit a retraction to a directed cycle, $H^{\wedge} - COL$ is NP-complete by Theorem 3.2. Therefore $H - COL$ is also NP-complete.

Case 3: $a = b$, and there exists a vertex x on an a-cycle such that the vertex $x + a$ is also on an a-cycle.

Since C_a is not a retract, the digraph H is retract-free. Relabel the vertices so that x is labelled 0, $x + 1$ is labelled 1, and so on. (That is, subtract x from the label of each vertex, where computations are modulo n). Let $k \geq 0$. If u is a vertex on a directed a-cycle then the set of vertices reachable from u by a directed walk of length ka is $\{u+0, u+a, u+2a, \ldots, u+ka\}$, where computations are modulo n. Let m be the order of the element a in Z_n. Note that $m > 2$ (if $m = 2$ then $2a = n$, whence C_a is a retract). We show that NOT-ALL-EQUAL m-SAT without negated variables polynomially transforms to $H-COL$. Suppose an instance of NOT-ALL-EQUAL m-SAT without negated variables is given, with variables $x_1, x_2, \ldots, x_p$, and clauses $K^1, K^2, \ldots, K^q$. Construct a digraph G from H, $\{x_1, x_2, \ldots, x_p\}$, and q copies of C_n, say $C^1, C^2, \ldots, C^q$, by adding directed paths as follows. Vertex 0 in H is joined to each vertex x_j $(j = 1, 2, \ldots, p)$ by a directed path of length a, starting at vertex 0. Vertex 0 in H is also joined to vertex 0 on each C^l $(l = 1, 2, \ldots, q)$ by a directed path of length ma, starting at vertex 0 of H. If the r^{th} variable in clause K^s is x_t, then join x_t to vertex ra of C^s by a directed path of length $(m-2)a$, starting at x_t. Clearly the digraph G is constructible in polynomial time.

Claim. The digraph G is H-colourable if and only if there is a satisfying truth assignment in which each clause contains at least one true variable and at least one false variable.

Proof: ($\Rightarrow$) Consider a homomorphism of G to H. Since H is retract-free, the copy of H in G must map onto H. We may therefore assume that every vertex of H maps to itself. Thus each vertex x_i $(i = 1, 2, \ldots, n)$ maps to 0 or a. Moreover, each C^j maps onto the directed n-cycle in H (because a does not divide n), and vertices $0, a, 2a, \ldots, (m-1)a$ of C^l map, in cyclic order, to the corresponding set of vertices of H. Define a truth assignment be setting $x_i = T$ just if x_i maps to 0. Consider an arbitrary clause C^s. Recall that there is no directed $(0, -a)$-walk of length $(m-2)a$ in H. Let v be the vertex of C^s that maps to $-a$, and let x_r be the vertex joined to v by a copy of $P(m-2)a$. Then x_r must map to a. Hence K^s contains a false variable. Similarly K_s contains a true variable (there is no directed walk of length $(m-2)a$ from vertex a to vertex 0 in H).

($\Leftarrow$) Suppose such a truth assignment exists. Define an H-colouring of G as follows. Every vertex of the copy of H is coloured by itself. For $i = 1, 2, \ldots, t$, if $x_i = T$, then colour x_i by 0, otherwise colour x_i by a. This partial colouring extends to all of the directed paths joining the x_j's to the copy of H. Consider K^s. There exists t such that the t^{th} variable l_t in K^s is true, and the $(t+1)^{\text{st}}$ variable l_{t+1} is false. Colour vertex la of C^s by $(-2a)$ and vertex $(l+1)a$ of C^s by $(-a)$. This completely determines the colouring of C^s (each n-cycle of G must map onto the n-cycle of H, since the a-cycle is not a retract). Furthermore, this partial colouring can be extended to all of the directed $(m-2)a$-paths joining C^s to x_i $(x_i \in K^s)$, and to the directed m-paths joining H to C^s. Therefore G is H-colourable.

This completes the proof of case 3.

Case 4: $a = b$ and for every vertex x on a directed a-cycle the vertex $x + a$ is not on a directed a-cycle.

Without loss of generality, the vertex 0 is on a directed a-cycle. Since C_a is not a retract, the digraph H is retract-free. Let m be the order of the element a in Z_n. Note that $m > 2$ (if $m = 2$ then $2a = n$, whence C_a is a retract).
Claim. Each directed a-cycle contains at least two elements of $\langle a \rangle$.

Proof: Each directed a-cycle contains the same number of elements of $\langle a \rangle$. If this number is one, then a divides n and C_a is a retract, which is a contradiction.

Let I be the directed graph constructed from a directed path of length $(m-1)a$ as follows. Identify vertex 0 (on the path) with a vertex on a directed a-cycle, and identify $(m-1)a$ with a vertex on a second directed a-cycle. For $i = 1, 2, \ldots, a$, add a directed path of length $n-1$ from i to $i-1$. Let H^* be the result of applying the indicator construction with respect to $(I, 0, (m-1)a)$ to H. There is no H-colouring of I such that the vertex a is coloured by a vertex on a directed a-cycle (otherwise *colour*(0) and *colour*(0)+a = *colour*(a) are vertices of H that are both on a directed a-cycle, which is a contradiction). Let A be the set of vertices of H which are on directed a-cycles. Let $x \in A$ and consider an H-colouring of I such that colour(0)=x. Since vertex a of I does not map to a vertex on a directed a-cycle, the possible images of vertex $(m-1)a$ of I are those vertices which also lie on a directed a-cycle, and are reachable from vertex $x+a$ of H by a directed walk of length $(m-2)a$. (Note that the first a vertices of I must map to $x+1, \ldots, x+a$ in H, because a does not divide n). Thus

$$colour((m-1)a) \in \{x + 2a, x + 3a, \ldots, x + (m-1)a\} \cap A = Y.$$

so $colour((m-1)a) \neq colour(0)$. Hence H^* is loopless. Moreover the vertex $(m-1)a$ can be coloured by any vertex in the set Y. The claim now implies that H^* contains the equivalent digraph of K_4. The result H^{**} of applying the sub-indicator construction with respect to $(C_2, 0, free)$ to H^* is a loopless undirected graph that has an odd cycle. Thus $H^{**}-COL$ is NP-complete, and so H^*-COL and $H - COL$ are also NP-complete.

All cases have been considered. ∎

3.6. Lemma. *If H is of type III and does not admit a retraction to a directed cycle, then $H - COL$ is NP-complete.*

Proof: Let H be of type III. Then H has three directed cycles, say of lengths n, a, and b. Without loss of generality assume $n > a \geq b$. Suppose that the core of H is not a directed cycle. We may further assume that the chords have no commom vertex, since this occurrance is covered under Lemmas 3.4 and 3.7. There are three cases to consider.

Case 1. b does not divide a.

The argument is similar to case 1 of Lemma 3.5, and uses similar sub-indicators.

Case 2. b divides a, and $b < a$.

Since the directed b-cycle is not a retract, b does not divide n. The remaining details are identical to those of case 2 of Lemma 3.5.

Case 3. $a = b$.

Then there is a vertex x on a directed a-cycle such that $x + a$ is also on a directed a-cycle. The remaining details are identical to those of case 3 of Lemma 3.5.

This completes the proof. ∎

3.7. Lemma. *If H is of type IV and does not admit a retraction to a directed cycle, then $H - COL$ is NP-complete.*

Proof: The digraph H has four directed cycles, say of lengths n, a, b, and c. Without loss of generality assume $n > a \geq b > c$. Note that $n = a + b - c$. Suppose that the core of H is not a directed cycle. There are four cases to consider.

Case 1. c divides b.

Then the subdigraph of H induced by the vertex set of the directed a-cycle is a retract. Since C_c is not a retract of H, c does not divide a. Consequently the core of H is of the form D_1, and $H - COL$ is NP-complete by Theorem 3.2.

Case 2. c does not divide b, and $b < a$.

Let H^- be the result of applying the sub-indicator construction with respect to $(C_b, 0, free)$ to H. Then H^- consists of a directed b-cycle plus a chord that belongs to the directed c-cycle. That is, H^- is of the form D_1. Since c does not divide b, $H^- - COL$ is NP-complete by Theorem 3.2, and therefore $H - COL$ is also NP-complete.

Case 3. $a = b$, c does not divide b, and c does not divide n.

Note c does not divide a. Let m be the order of the element a in Z_n. If $m = 2$, then $2a = n$, and hence $c = 0$, which is a contradiction. Therefore $m > 2$. Let Q_r $(r \geq a - 1)$ denote the $(r + 1)$-vertex digraph constructed from P_r by adding the arcs $\{i(i - a + 1) : i = a - 1, a, \ldots, r\}$. Since any $a - 1$ consecutive arcs along the directed r-path must belong to an image of a directed a-cycle and c does not divide a, no image of Q_r in H contains a directed c-cycle. This effectively eliminates the use of the directed c-cycle. The transformation is from NOT-ALL-EQUAL m-SAT without negated variables, and is identical to case 3 of Lemma 3.5, except that wherever P_r appears in the construction, Q_r should be used.

Case 4. $a = b$, and c divides n.

Since C_c is not a retract, c does not divide a. Let J be the directed graph constructed by identifying the initial vertex of a directed path of length $a - 2$ with a vertex on a directed a-cycle. Let x be the terminal vertex of the directed path. Let the vertices of H be numbered cyclically such that vertex 0 is the terminal vertex of one of the chords. Let H^- be the result of applying the sub-indicator construction with respect to $(J, x, free)$ to H. It may be directly verified that $H = H - \{2a - 1\}$. Consequently the core of H is of the form D_1 and, since c does not divide a, $H^- - COL$ is NP-complete. Therefore $H - COL$ is also NP-complete.

All cases have been considered. ∎

Theorem 3.3 generalises Theorem 3.2 for digraphs of the form D_1 (cf. figure 3.1). Our next result generalises the same theorem for digraphs of the form D_2. Let $Q = q_0, q_1, \ldots, q_q$ be an oriented path, and let r and s be integers. Let H be the digraph constructed from $C_r \cup C_s \cup Q$ as follows. Let $V(C_r) = \{r_0, r_1, \ldots, r_{r-1}\}$, and $E(C_r) = \{r_i r_{i+1} : i = 0, 1, \ldots, r-1\}$. Similarly, let $V(C_s) = \{s_0, s_1, \ldots, s_{s-1}\}$ and $E(C_s) = \{s_i s_{i+1} : i = 0, 1, \ldots, s-1\}$. Identify the vertices q_0 and q_q with r_0 and s_0, respectively. That is, H is of the form D_2 except that the directed path has been replaced by an oriented path.

3.8. Theorem. *If r divides s or s divides r, then $H - COL$ is polynomial. Otherwise (r does not divide s and s does not divide r) $H - COL$ is NP-complete.*

Proof: If r divides s, then the directed r-cycle is a retract, and if s divides r, then the directed s-cycle is a retract. In either case, the H-colouring problem is polynomial.

Suppose r does not divide s and s does not divide r. (Note that this implies that H is retract-free.) By Theorem 3.2 we may assume that Q is not a directed path. Let $k = nl(Q)$, and let a and b be any integers such that $a, b > |V(Q)|$, $a \equiv 1 - k \pmod{rs}$, and $b \equiv 0 \pmod{rs}$. Let I be the digraph constructed from $P_a \cup P_b \cup Q$ and two rs-cycles with special vertices x and y respectively, by identifying the terminal vertex of P_a with q_0 and x and the initial vertex of P_b with q_q and y. Let u be the initial vertex of P_a and let v be the terminal vertex of P_b. Then the net-length of the (u, v)-path P in I is congruent to 1 modulo rs. Let H^* be the result of applying the indicator construction with respect to (I, u, v) to H. We make the following assertions about the digraph H^*.

"(1)" **Every internal vertex of Q is either isolated, a source of H^*, or a sink of H^*.**

Let q_i be an internal vertex of Q. If there is no directed path between q_i and either q_0 or q_t, then the vertex q_i is isolated in H^*, since there is no homomorphism of P_a to Q that maps u to q_i. Suppose there is a directed path from q_i to q_q. Since Q is not a directed path, there is no directed walk of length b that ends at q_i. Similarly, if there is a directed path from q_i to q_0, there is no directed walk of length b that ends at q_i. Hence q_i is a source of H^*. The existence of a directed path from q_0 or q_q to q_i similarly impies that q_i is a sink of H^*.

(2) **H^* is loopless.**

By (1) no internal vertex of Q is incident with a loop. Suppose r_i is incident with a loop. Then there is a homomorphism of I to H which takes both u and v to r_i. Let W be the walk in H determined by the image of I. Since I and Q have the same number of sources (and sinks), no vertex of W is on the directed s-cycle. That is, W is contained in the subdigraph

$(C_r \cup Q) - q_q$. Since the net length of each (q_0, q_0)-section of W is zero, it follows that $nl(W) \equiv 0 \pmod r$, which is a contradiction (recall that $nl(P) \equiv 1 \pmod{rs}$). Similarly, no vertex of C_s is incident with a loop.

(3) **H^* contains both C_r and C_s.**

This is clear since $nl(P) \equiv 1 \pmod{rs}$.

(4) **Neither C_r nor C_s has a chord.**

Consider a homomorphism of I into H that takes u to r_i and v to r_j. By (2), $i \neq j$. Arguing as in (2), the image of I is contained in the subdigraph $(C_r \cup Q) - q_q$. Since the net length of each (q_0, q_0)-section of the image of I is zero, the net-length of the walk defined by the image of I is congruent to $(j - i)$ modulo r. Therefore $j = i + 1 \pmod r$, and C_r has no chord. Similarly C_s has no chord.

(5) **The arc $r_{r+k-1} s_0$ exists.**

We describe the necessary homomorphism of I to H. Map u to r_{r+k-1}. Since $a \equiv 1 - k \pmod{rs}$ the first vertex in the copy of Q in I maps to $r_0 = q_0$. Now map each vertex of the copy of Q in I to the corresponding vertex of Q, and map the copy of P_b in I to C_s. Since $b \equiv 0 \pmod{rs}$, the vertex v maps to s_0. Clearly the rs-cycles at x and y can be mapped to C_r (respectively C_s).

(6) **The arc $s_{s+k-1} r_0$ exists if and only if Q is self-converse.**

($\Rightarrow$) If the arc exists, then the copy of Q in I must map onto the copy of Q^{-1} in H.

($\Leftarrow$) The argument is similar to (5).

If Q is self-converse, then $nl(Q) = 0$. Hence the arcs from (5) and (6) are $s_{s-1} r_0$ and $r_{r-1} s_0$, respectively.

(7) **There are no other arcs between C_r and C_s.**

Consider a homomorphism of I to H in which u maps to a vertex on one of the directed cycles and v maps to a vertex on the other directed cycle. It is not hard to see that the copy of Q in I must map onto Q. Since homomorphisms to directed cycles are completely determined by the image of a single vertex, this forces u to map to s_{s+k-1} and v to map to r_0, or u to map to r_{r+k-1} and v to map to s_0, (depending on the orientation of the supposed arc).

Thus the structure of H^* is completely determined. Let G be the core of H^*, and let G^- be the result of applying the sub-indicator construction with respect to $(P_2, 1, \mathit{free})$ to G. Then G^- is the subdigraph of H^* induced by $V(H^*) - \{q_1, q_2, \ldots, q_{q-1}\}$. If Q is not self-converse, G^- consists of a directed r-cycle and a directed s-cycle joined by an arc. Since r does not divide s and s does not divide r, $G^- - COL$ is NP-complete by Theorem 3.2. On the other hand, if Q is self converse, G^- consists of a directed cycle with two chords and is of type II. The lengths of the cycles are $r + s$, r, and s. Since r does not divide s and s does not divide r, $G^{r} - COL$ is NP-complete by Theorem 3.3. Therefore $H - COL$ is also NP-complete. This completes the proof. ∎

4. A new reduction

In this section we describe a new tool for proving NP-completeness (resp. NP-hardness) results in directed H-colouring (cf. Theorem 4.5). In particular, we show that it is often sufficient for NP-hardness of $H - COL$ that H have a strong component C for which $C - COL$ is NP-hard. In a sense, this result is related to the concept of "hereditarily hard" H-colouring problems introduced in [Bang-Jensen et al. 1989] (we say $H - COL$ is hereditarily hard if $G - COL$ is NP-hard whenever H is a subdigraph of G). On combining this theorem with other results which have appeared in the literature, we obtain infinitely many new infinite families of NP-complete (resp. NP-hard) H-colouring problems.

We begin by specializing the following lemma to strong digraphs (cf. Corollary 4.3).

4.1. Lemma [Häggkvist et al, 1987]. *There is a homomorphism of a directed graph H to C_d if and only if the net-length of every (oriented) cycle is divisible by d.* ∎

Therefore a given directed graph does not admit a homomorphism to C_n just if it has a cycle of net-length not divisible by n, and does not admit a homomorphism to any directed cycle of length greater than one if and only if it has a collection $C^1, C^2, \ldots, C^k$ of cycles such that $gcd\{nl(C^i) : i = 1, 2, \ldots, k\} = 1$.

4.2. Lemma. *Let H be strong. There is no homomorphism of H to C_d if and only if there exists an integer k such that d does not divide k, and there is a homomorphism of C_k to H.*

Proof:

($\Rightarrow$) Suppose H does not admit a homomorphism to C_d. Let W be a closed walk in H with net length not divisible by d (the walk W exists by Lemma 4.1), and with the minimum number of backwards arcs among all such closed walks. If W has no backwards arcs there is nothing to prove, so we may assume that W has at least one backwards arc, xy say. Since H is strong, there is a directed (y, x)-path P. By our assumption on W, the length of the directed closed walk $P + xy$ is a multiple of d, say qd. Let $W' = W - xy$ (*i.e.*, the (x, y)-section of W). Then $W'P$ is a closed walk with one fewer backwards edge than W, and $nl(W'P) = nl(W) + 1 + (qd - 1)$ which is not divisible by d. This contradicts the choice of W, and completes the proof of the implication.

($\Leftarrow$) The image of C_k in H is a union of directed cycles. Since d does not divide k, the digraph H has a cycle of length not divisible by d. Consequently there is no homomorphism of H into C_d. ∎

4.3. Corollary. *Let H be strong. There is a homomorphism of H into C_d if and only if the length of every directed cycle is divisible by d.* ∎

Therefore a given strong digraph does not admit a homomorphism to C_n just if it has a directed cycle of net length not divisible by n, and does not admit a homomorphism to any directed cycle of length greater than one if and only if it has a collection $C^1, C^2, \ldots, C^k$ of directed cycles such that $gcd\{nl(C^i) : i = 1,2,\ldots,k\} = 1$.

Let H be a strong digraph. Then H has a directed cycle. Let g be the directed girth 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.

4.4. Lemma. *Let H be strong, and let d be the largest integer such that there is a homomorphism f of H to C_d. For any vertex v of H there is an integer l_v (resp. b_v) such that, for every vertex x in $f^{-1}(f(v))$, there is a directed (v,x)-walk of length l_v (resp. directed (x,v)-walk of length b_v).*

Proof: We prove only the existence of l_v; the existence of b_v may be established similarly. First we find an integer l such that there is a directed (v,y)-walk of length l for every vertex y in $\{v\} \cup N^{+d}(v)$. We then use l to define l_v.

By Corollary 4.3 the digraph H has a collection $C^1, C^2, \ldots, C^n$ of directed cycles such that $gcd\{|V(C^i)| : i = 1,2,\ldots,n\} = d$. Since H is strong, the vertex v lies on a directed cycle K of length kd, for some k. Let $\langle d \rangle$ denote the subgroup of Z_{kd} generated by d. Then $\langle d \rangle = \langle \{|V(C^i)| (\mathrm{mod}\ kd) : i = 1,2,\ldots,n\} \rangle$, so there exist directed cycles $D^1, D^2, \ldots, D^t \in \{C^1, C^2, \ldots, C^n\}$ (not necessarily all distinct) such that

$$|V(D^1)| + |V(D^2)| + \cdots + |V(D^t)| \equiv d\ (\mathrm{mod}\ kd).$$

For $j = 1,2,\ldots,t$, let v_j be a vertex on D^j, let W^j be a (v,v_j)-path, and let X^j be a (v_j,v)-path. Define

$$\begin{aligned} l = &|V(D^1)| + |V(D^2)| + \cdots + |V(D^t)| \\ &+ |V(W^1)| + |V(W^2)| + \cdots + |V(W^t)| \\ &+ |V(X^1)| + |V(X^2)| + \cdots + |V(X^t)|. \end{aligned}$$

It is not hard to see that $S = W^1 D^1 X^1 W^2 D^2 X^2 \ldots W^t D^t X^t$ is a directed (v,v)-walk of length l. Let $u \in N^{+d}(v)$. There is a directed (v,u)-walk of length l, namely $T = W^1 X^1 W^2 X^2 \ldots W^t X^t P$, where P is the (v,u)-walk of length $|V(D^1)| + |V(D^2)| + \cdots + |V(D^t)|$ formed by traversing K repeatedly, and then using the last d arcs of P to traverse a (v,u)-path of length d.

Next we define

$$l_v = l \cdot k_v, \quad k_v = \max\{k : \exists u \in f^{-1}(f(v)) \text{ such that } d(v,u) = k\}.$$

Let x be in $f^{-1}(f(v))$. Then $d(v, x)$ is divisible by d. There is clearly a directed (v, x)-walk of length l_v formed by traversing S $k_v - d(v, x)/d$ times, traversing $T - P$ $d(v, x)/d$ times, traversing K repeatedly, and then using the last $d(v, x)$ arcs to traverse a directed (v, x)-path of length $d(v, x)$. The result follows. ∎

4.5. Theorem. *Let H be a strong component of a retract-free digraph D. Then $H - COL$ polynomially Turing reduces to $D - COL$. Furthermore, if H does not admit a homomorphism to a directed cycle of length greater than one, $H - COL$ polynomially transforms $D - COL$.*

Proof: Let d be the largest integer such that there is a homomorphism of H to C_d, and fix a homomorphism $f\colon H \to C_d$. Let G be a given digraph. We define a collection of digraphs, $G_1, G_2, \ldots, G_d$, such that there is a homomorphism of G to H if and only if there is a homomorphism of some G_i to D.

There exists in H a directed path $v_1, v_2, \ldots, v_d$ and f assigns a different colour to each of these vertices. For $i = 1, 2, \ldots, d$ let l_i, and b_i be the lengths from Lemma 4.3 corresponding to v_i. If G admits a homomorphism to H, then there is a homomorphism of G to C_d. Since the C_d-colouring problem is polynomial, it may be assumed that a C_d-colouring m of G is known. The digraph G_i is constructed from the disjoint union of G and D by adding directed paths as follows: Let g be a vertex of G, and suppose that $m(g) = x$. Let $k = i + x \pmod d$. Add a directed (g, v_k)-path of length b_k, and a directed (v_k, g)-path of length l_k. The digraph G_i results from applying this construction to every vertex of G.

Claim: There is a homomorphism of G to H if and only if there is a homomorphism of some G_i to D.

Proof.

($\Rightarrow$) Let h be an H-colouring of G. Then $c = f \circ h$ is a C_d-colouring of G. Let g be a vertex of G and, without loss of generality, suppose $m(g) = 0$. Let $j = c(g)$. We claim that there is a homomorphism of G_j to D. Namely, map the copy of D in G_j identically onto itself. Map each vertex of G to its image in the H-colouring of G. By Lemma 4.4, this partial colouring can be extended to all of the paths. Hence $G_j \to D$.

($\Leftarrow$) Without loss of generality assume that G_1 admits a homomorphism to D. Since D is retract-free, we know that G_1 maps onto D. As $H \cup G$ is contained in a strong component of G_1, it is mapped to a strong component of D. Since D is retract-free, this component is isomorphic to H. Hence there is a homomorphism of G to H.

Since the collection $G_1, G_2, \ldots, G_d$ can be constructed in polynomial time, the result follows. Furthermore, if H does not admit a homomorphism to a directed cycle of length greater than one, then $d = 1$ and the construction described above is a polynomial transformation. ∎

Hence whenever we prove that $H - COL$ is NP-hard for some strong digraph H, we obtain, via Theorem 4.5, an infinite family of NP-hard H-colouring problems. In particular, this implies that $H - COL$ is NP-hard whenever the retract-free digraph H has a strong component which belongs to any one of the following classes of digraphs.

- strong semi-complete digraphs with two or more directed cycles.
- strong bipartite tournaments that do not retract to a a directed cycle.
- vertex-transitive digraphs that do not retract to a directed cycle.
- partitionable digraphs (see [Bang-Jensen et al. 1989]) that contain an oriented odd cycle.
- digraphs of the form D_1 or D_2 (cf. figure 3.1) that do not retract to a directed cycle (*i.e.*, the length of the shorter directed cycle does not divide the length of the longer directed cycle).
- directed cycles with two chords that do not retract to a directed cycle (*i.e.*, the length of the shortest directed cycle does not divide the length of all other directed cycles).
- digraphs of the form $C_{n,k}$, where n is odd and $k > 0$.

It should be noted, however, that there are NP-complete H-colouring problems such that $C - COL$ is polynomial for every strong component C of H (e.g. see Theorem 3.8).

5. Acknowledgement

Both authors thank Pavol Hell for many helpful conversations. The second author thanks the Natural Sciences faculty of Capilano College in North Vancouver for their encouragement and assistance during the 1989–90 academic year.

References

1. Bang-Jensen, J., *On the complexity of generalized colourings of directed graphs*, Proc. Twente Workshop on Graphs and Combinatorial Optimization, Memorandum no. 787, Universiteit Twente, Faculteit der Togepaste Wiskunde (1989), 13–20.
2. Bang-Jensen, J. and P. Hell, *The effect of two cycles on the complexity of colourings by directed graphs*, Discrete Applied Math. **26** (1990), 1–23.
3. Bang-Jensen, J., P. Hell and G. MacGillivray, *The complexity of colourings by semicomplete digraphs*, SIAM J. Discrete Math. **1** (1988), 281–298.
4. Bang-Jensen, J., P. Hell and G. MacGillivray, *On the complexity of colouring by superdigraphs of bipartite graphs*, Discrete Math. (to appear).
5. Bang-Jensen, J., P. Hell and G. MacGillivray, *Hereditarily hard colouring problems (submitted).*

6. Bondy, J. A. and U. S. R. Murty, "Graph Theory with Applications", Elsevier Press, New York, 1976.
7. Brouwer, A. E. and H. J. Veldman, *Contractability and NP-completeness*, J. Graph Theory **11** (1987), 71–79.
8. Fellner, W.D., *On minimal graphs*, Theoretical Comp. Sci. **17** (1982), 103–110.
9. Fortune, S., J. Hopcroft and J. Wyllie, *The directed subgraph homeomorphism problem*, Theoretical Comp. Sci. **10** (1980), 111–121.
10. Garey, M. R. and D. S. Johnson, "Computers and Intractability", W. H. Freedman, San Francisco, 1979.
11. Gutjahr, W., *Farbung durch gerichtete Graphen*, Diplomarbeit, Institute for Information Processing, IIG, Technical University Graz (1988).
12. Gutjahr, W. Ph.D. Thesis. Technical University Graz.
13. Gutjahr, W., E. Welzl and G. Woeginger, *Polynomial graph colourings*, Technical Report B-88-06, Freie Universitat Berlin.
14. Gutjahr, W., E. Welzl and G. Woeginger, Personal communication, 1989.
15. Häggkvist, R., P. Hell, D.J. Miller and V. Neuman Lara, *On multiplicative graphs and the product conjecture*, Combinatorica **8** (1988), 63–74.
16. Hell, P. and J. Nešetřil, *On the complexity of H-colouring*, J.C.T. Ser B **48** (1990), 92–110.
17. Johnson, D.S., *The NP-completeness column - An ongoing guide*, J. Algorithms **3** (1982), 89–99.
18. Lovász, *L. Coverings and colorings of hypergraphs*, Proc. 4th Southeast Conference on Combinatorics, Graph Theory and Computing, Utilitas Mathematica Publishing, Winnipeg (1973).
19. MacGillivray, G., *The complexity of generalised colourings*, Ph.D. Thesis, Simon Fraser University, Burnaby, Canada (1989).
20. MacGillivray, G., *The complexity of colouring by transitive digraphs*, SIAM J. on Disc. Math. **4** (1991), 397–408.
21. Maurer, H. A., J. H. Sudborough and E. Welzl, *On the complexity of the general colouring problem*, Information and Control **51** (1981), 123–145.
22. Maurer, H.A., A. Salomaa and D. Wood, *Colourings and interpretations: a connection between graphs and grammar forms*, Discrete Appl. Math. **3** (1981), 119–135.
23. Nešetřil, J., *Representations of graphs by means of products and their complexity*, MFCS (1982), 94–102.
24. Nešetřil, J. and A. Pultr, *On classes of relations determined by subobjects and factorobjects*, Discrete Math. **22** (1978), 287–300.
25. Schaefer, T.J., *The complexity of satisfiability problems*, Proc. 10th ACM Symp. on Theory of Computing (1978), 216–226, A.C.M., New York.
26. Welzl, E., *Colour families are dense*, Theoretical Comp. Sci. **17** (1982), 29–41.
27. Welzl, E. Personal communication, 1987.