A Dichotomy for Minimum Cost Graph Homomorphisms

Gregory Gutin* Pavol Hell† Arash Rafiey‡ Anders Yeo§

Abstract

For graphs G and H, a mapping $f: V(G) \to V(H)$ is a homomorphism of G to H if $uv \in E(G)$ implies $f(u)f(v) \in E(H)$. If, moreover, each vertex $u \in V(G)$ is associated with costs $c_i(u), i \in V(H)$, then the cost of the homomorphism f is $\sum_{u \in V(G)} c_{f(u)}(u)$. For each fixed graph H, we have the *minimum cost homomorphism problem*, written as MinHOM(H). The problem is to decide, for an input graph G with costs $c_i(u)$, $u \in V(G), i \in V(H)$, whether there exists a homomorphism of G to H and, if one exists, to find one of minimum cost. Minimum cost homomorphism problems encompass (or are related to) many well studied optimization problems. We prove a dichotomy of the minimum cost homomorphism problems for graphs H, with loops allowed. When each connected component of H is either a reflexive proper interval graph or an irreflexive proper interval bigraph, the problem MinHOM(H) is polynomial time solvable. In all other cases the problem MinHOM(H) is NP-hard. This solves an open problem from an earlier paper.

1 Motivation and Terminology

We consider finite graphs (and digraphs) without multiple edges, but with loops allowed. For a graph (or digraph) H, we use $V(H)$ and $E(H)$) to

*Department of Computer Science, Royal Holloway University of London, Egham, Surrey TW20 OEX, UK, gutin@cs.rhul.ac.uk and Department of Computer Science, University of Haifa, Israel

†School of Computing Science, Simon Fraser University, Burnaby, B.C., Canada, V5A 1S6, pavol@cs.sfu.ca

‡Department of Computer Science, Royal Holloway University of London, Egham, Surrey TW20 OEX, UK, arash@cs.rhul.ac.uk

§Department of Computer Science, Royal Holloway University of London, Egham, Surrey TW20 OEX, UK, anders@cs.rhul.ac.uk

denote the set of vertices and edges of G. A graph (or digraph) without loops will be called *irreflexive*; a graph (or digraph) in which every vertex has a loop will be called *reflexive*. In this paper our focus will be on graphs, but we shall make some remarks about digraphs as well.

The *intersection graph* of a family $\mathcal{F} = \{S_1, S_2, \ldots, S_n\}$ of sets is the graph G with $V(G) = \mathcal{F}$ in which S_i and S_j are adjacent just if $S_i \cap S_j \neq \emptyset$. Note that by this definition, each intersection graph is reflexive. (This is not the usual interpretation [10, 29].) A graph isomorphic to the intersection graph of a family of intervals on the real line is called an *interval graph*. If the intervals can be chosen to be inclusion-free, the graph is called a *proper interval graph*. Thus both interval graphs and proper interval graphs are reflexive. The *intersection bigraph* of two families $\mathcal{F}_1 = \{S_1, S_2, \ldots, S_n\}$ and $\mathcal{F}_2 = \{T_1, T_2, \ldots, T_m\}$ of sets is the bipartite graph with $V(G) = \mathcal{F}_1 \cup \mathcal{F}_2$ in which S_i and T_j are adjacent just if $S_i \cap T_j \neq \emptyset$. Note that by this definition an intersection bigraph is irreflexive (as are all bipartite graphs). A bipartite graph isomorphic to the intersection bigraph of two families of intervals on the real line is called an *interval bigraph*. If the intervals in each family $\mathcal{F}_i$ can be chosen to be inclusion-free, the graph is called a *proper interval bigraph*. Thus both interval bigraphs and proper interval bigraphs are irreflexive.

For graphs (or digraphs) G and H, a mapping $f : V(G) \to V(H)$ is a *homomorphism of G to H* if $uv \in E(G)$ implies $f(u)f(v) \in E(H)$. Recent treatment of graph (and digraph) homomorphisms can be found in [19, 21]. Let H be a fixed graph (or digraph). The *homomorphism problem* for H asks whether an input graph (or digraph) G admits a homomorphism to H. The *list homomorphism problem* for H asks whether an input graph (or digraph) G with lists (sets) $L_u \subseteq V(H), u \in V(G)$ admits a homomorphism f to H in which all $f(u) \in L_u, u \in V(G)$.

There have been several studies of homomorphism (and more generally constraint satisfaction) problems with costs. Most frequently, each edge ij of the graph H has a cost $c(i, j)$ [1, 2]. (It is then natural to take H to be a complete (reflexive) graph.) In this context, given an input graph G, one seeks a homomorphism f of G to H with minimum cost, i.e., a homomorphism for which the sum, over all $uv \in E(G)$, of $c(f(u)f(v))$ is minimized. These are typified by problems such as finding a maximum bipartite subgraph, or, in the context of more general constraints, finding an assignment satisfying a maximum number of clauses [2]. More generally, [5] considers instead of costs of edges ij of H, the costs of mapping an edge uv of G to an edge ij of H. Of course, we typically assume again that H is a complete graph. In this way, the constraint on the edge uv is 'soft' - it

may map to any pair ij of H, but with cost that depends both on uv and on ij. Nonbinary constraints are treated the same way in [5]. This general 'soft' constraint satisfaction context of [5] allows for vertex weights as well, since they can be viewed as unary constraints. We describe this model in greater detail below. Nevertheless, in combinatorial optimization it makes sense to investigate vertex weights alone, insisting that binary (and higher order) constraints are hard, or 'crisp'. This is the path we take, focusing on problems in which each possible assignment of a value to a variable has an associated cost.

We now formulate our problem, in the context of graph homomorphims. (Of course, there is a natural counterpart for constraint satisfaction problems in general.) Suppose G and H are graphs (or digraphs). As in the above discussion, we shall reserve the letters u, v, etc., for the vertices of G, and the letters i, j, etc., for the vertices of H. Let $c_i(u)$, $u \in V(G)$, $i \in V(H)$, be a nonnegative real number, which we shall think of as the *cost* of mapping u to i. The *cost of a homomorphism* f of G to H is $\sum_{u \in V(G)} c_{f(u)}(u)$. If H is fixed, the *minimum cost homomorphism problem*, MinHOM(H), for H is the following decision problem. Given an input graph G, together with costs $c_i(u)$, $u \in V(G)$, $i \in V(H)$, and an integer k, decide if G admits a homomorphism to H of cost not exceeding k.

We shall also use MinHOM(H) to denote the corresponding optimization problem, in which we want to minimize the cost of a homomorphism of G to H, or state that none exists. The minimum cost of a homomorphism of G to H (if one exists) will be denoted by $\mathrm{mch}(G, H)$. For simplicity, we shall always assume the graph G to be irreflexive. (Note that we can always solve a problem in which some vertices u of G have loops, by changing the weights $c_i(u)$ to be infinite on all vertices i of H which do not have a loop.)

Returning briefly to the problem of minimum cost *soft* homomorphism problem of [5], in the context of *graphs*, we may reformulate it as follows. Suppose H is a fixed complete graph. Given an input graph G, together with nonnegative costs $c_i(u), u \in V(G)$, $i \in V(H)$, and nonnegative costs $c_{ij}(uv), uv \in E(G)$, $ij \in E(H)$, find a mapping f of $V(G)$ to H which minimizes the sum $\sum_{u \in V(G)} c_{f(u)}(u) + \sum_{uv \in E(G)} c_{f(u)f(v)}(uv)$. This generalizes our problem MinHOM(H) for a *graph* H, as we can set the (hard) values $c_{f(u)f(v)}(uv) = 0$ if $f(u)f(v) \in E(H)$ and $c_{f(u)f(v)}(uv) = \infty$ if $f(u)f(v) \notin E(H)$. As mentioned earlier, we do not study this problem, focusing instead on the simpler problem MinHOM(H); the interested reader should consult [5].

The problem MinHOM(H) was introduced in [13], where it was motivated by a real-world problem in defence logistics. We believe it offers a

practical and natural model for optimization of weighted homomorphisms. It is easy to see that the homomorphism problem (for H) is a special case of MinHOM(H), obtained by setting all weights to 0 (and taking $k = 0$). Similarly, the list homomorphism problem (for H) is obtained by setting $c_i(u) = 0$ if $i \in L_u$ and $c_i(u) = 1$ otherwise (and taking $k = 0$). When H is an irreflexive complete graph, the problem MinHOM(H) becomes the so-called *general optimum cost chromatic partition* problem, which has been intensively studied [16, 22, 23], and has a number of applications, [25, 30]. Two special cases of that problem that have been singled out are the *optimum cost chromatic partition problem*, obtained when all $c_i(u), u \in V(G)$, are the same (the cost only depends on the colour i) [25], and the *chromatic sum problem*, obtained when each $c_i(u) = i$ (the cost of the colour i is the value i, i.e., we are trying to minimize the sum of the assigned colours) [23].

Recall that we do admit loops in a graph (or digraph). For the homomorphism problem for graphs H, the following dichotomy classification is known: if H is bipartite or has a loop, the problem is polynomial time solvable; otherwise it is NP-complete [20]. For the list homomorphism problem for graphs H, a similar dichotomy classification is also known [8]. None of the weighted versions of homomorphism problems cited above has a known dichotomy classification. This includes the soft constraint satisfaction problem of [5], although the authors do identify a class of polynomially solvable constraints that is in a certain sense maximal. We shall provide a dichotomy classification of the complexity of MinHOM(H) for graphs.

Preliminary results on MinHOM(H) for irreflexive graphs were obtained by Gutin, Rafiey, Yeo and Tso in [13]: it was shown there that MinHOM(H) is polynomial time solvable if H is an irreflexive bipartite graph whose complement is an interval graph, and NP-complete when H is either a nonbipartite graph or a bipartite graph whose complement is not a circular arc graph. This left as unclassified a large class of irreflexive graphs, settled in this paper. In fact, we shall provide a general classification which applies to graphs with loops allowed.

Theorem 1.1 *Let H be a graph (with loops allowed). If each component of H is a proper interval graph or a proper interval bigraph, then the problem MinHOM(H) is polynomial time solvable. In all other cases, the problem MinHOM(H) is NP-complete.*

The theorem will be proved in the following two sections. The next section provides the polynomial time algorithms, and the following section proves the NP-completeness.

We note that in the two polynomial cases, each component of the graph H is either irreflexive or reflexive. Indeed, it is easy to see that if H contains an edge rs where r has a loop and s doesn't, then the problem MinHOM(H) is NP-complete. It suffices to notice that if G has all vertex costs $c_s(u) = 0, u \in V(G)$, and all other vertex costs $c_i(u) = 1, u \in V(G), i \neq s$, then there exists a homomorphism of cost not exceeding k if and only if G has an independent set of size $|V(G)| - k$. Thus it suffices to consider the reflexive and irreflexive graphs separately, and we shall do so in the remainder of the paper.

In Section 4, we discuss the situation for digraphs. At this point the classification is open, although we do mention some partial results.

2 Polynomial Algorithms

We say that a digraph H has the *Min-Max property* if its vertices can be ordered $w_1, w_2, \ldots, w_p$ so that if $i < j$, $s < r$ and $w_iw_r, w_jw_s \in E(H)$, then $w_iw_s \in E(H)$ and $w_jw_r \in E(H)$.

This property was first defined in [11], where it was identified as an important property of digraphs, as far as the problem MinHOM(H) is concerned. (We should point out that the original definition, which is easily seen equivalent to the one given above, required that if $w_iw_r, w_jw_s \in E(H)$, then also $w_xw_y \in E(H)$ for $x = \min(i, j), y = \min(r, s)$ and for $x = \max(i, j)$, $y = \max(r, s)$.)

Using an algorithm of [5], the authors of [11] proved the following result. (The proof in [11] is only stated for irreflexive digraphs, but it is literally the same for digraphs in general.)

Theorem 2.1 *[11] Let H be a digraph. If H satisfies the Min-Max property, then MinHOM(H) is polynomial time solvable.*

The Min-Max property is very closely related to a property of digraphs that has long been of interest [15]. We say that a digraph G has the *X-underbar property* if its vertices can be ordered $w_1, w_2, \ldots, w_p$ so that if $i < j$, $s < r$ and $w_iw_r, w_jw_s \in E(H)$, then $w_iw_s \in E(H)$. (In other words, $w_iw_r, w_jw_s \in E(H)$ implies that $w_xw_y \in E(H)$ for $x = \min(i, j), y = \min(r, s)$). It is interesting to note that the X-underbar property is sufficient to ensure that the list homomorphism problem for H has a polynomial solution [21].

We first apply Theorem 2.1 to reflexive graphs. It is important to keep in mind that we may view graphs as special digraphs, by replacing each edge

uv of the graph by the two opposite edges uv, vu of the digraph; this does not affect which mappings are homomorphisms [21]. Under this interpretation, we observe the following fact.

Proposition 2.2 *A reflexive graph H has the Min-Max property if and only if its vertices can be ordered $w_1, w_2, \ldots, w_p$ so that $i < j < k$ and $w_i w_k \in E(H)$ imply that $w_i w_j \in E(H)$ and $w_j w_k \in E(H)$.*

Proof: To see that the condition is necessary, consider the directed edge $w_i w_k$ and the loop $w_j w_j$ and apply the definition in digraphs. To see that it is sufficient, suppose $i < j$, $s < r$ and $w_i w_r, w_j w_s \in E(H)$. Observe that, up to symmetry, there are only two nontrivial cases possible - typified by $s < i < r < j$ and $s < i < j < r$. In both cases, the condition in the theorem and the loops $w_i w_i$ and $w_r w_r$ (respectively $w_j w_j$) ensure that $w_i w_s \in E(H)$ and $w_j w_r \in E(H)$. ◇

The condition in Proposition 2.2 is known to characterize proper interval graphs [6, 17].

Corollary 2.3 *A reflexive graph H has the Min-Max property if and only if it is a proper interval graph.* ◇

For irreflexive graphs H, we observe that the standard view of H as a digraph will not work. Indeed, if both uv and vu are directed edges of the digraph H, then the Min-Max property requires that both uu and vv be loops of H. Therefore, we shall view a bipartite graph H, with a fixed bipartition into (say) *white* and *black* vertices, as a digraph in which all edges are directed from white to black vertices. Under this interpretation, we observe the following fact. (We have simply replaced one ordering of all vertices with the induced orderings on white and black vertices; note that given orderings of white and black vertices, any total ordering preserving the relative orders of white and of black vertices satisfies the condition.)

Proposition 2.4 *A bipartite digraph H, with a fixed bipartition into white and black vertices, and with all edges oriented from white to black vertices, has the Min-Max property if and only if the white vertices can be ordered as $u_1, u_2, \ldots, u_p$ and the black vertices can be ordered as $v_1, v_2, \ldots, v_q$, so that if $i < j$, $s < r$ and $u_i v_r, u_j v_s \in E(H)$, then $u_i v_s \in E(H)$ and $u_j v_r \in E(H)$.* ◇

The condition in Proposition 2.4 is known to characterize proper interval bigraphs (also known as proper interval bipartite graphs) [28, 29].

Corollary 2.5 *An irreflexive graph H has the Min-Max property if and only if it is a proper interval bigraph.* $\diamond$

It now follows that we can apply Theorem 2.1 to reflexive proper interval graphs and irreflexive bipartite proper interval bigraphs, to deduce the polynomial algorithms in Theorem 1.1. To begin with, let us formulate the result for a connected graph H.

Corollary 2.6 *If H is a connected graph which is a reflexive proper interval graph or an irreflexive proper interval bigraph, then the problem MinHOM(H) is polynomial time solvable.*

Proof: For proper interval graphs H this directly follows from Theorem 2.1, and Corollary 2.3. For proper interval bigraphs, we note that we may assume that the graph G is also bipartite, else no homomorphism to H exists. We may also assume that G is connected, as otherwise we can solve the problem for each component separately. Thus we may take G to be given with white and black vertices (only two such partitions are possible for a connected graph), and orient all edges from white to black vertices. Now we can use Theorem 2.1, and Corollary 2.5, to derive a polynomial solution. $\diamond$

Corollary 2.7 *Let H be any graph (with loops allowed). If each component of H is a reflexive proper interval graph or an irreflexive proper interval bigraph, then the problem MinHOM(H) is polynomial time solvable.*

Proof: Let $H_i, i = 1, 2, \ldots, k$, be the components of H. As above, it suffices to solve the problem for each component of G separately; thus we assume that G is connected. Now the minimum cost homomorphism of G to H is the smallest minimum cost homomorphism to any H_i. Thus a polynomial time algorithm follows from the previous corollary. $\diamond$

The polynomial algorithms for MinHOM(H) follow from [5], via the translation in [11], which depends on submodularity of the cost functions. It is often the case that a problem solved using submodularity can be solved more efficiently by another more direct method [9]. This is indeed the case here, and we give such a direct algorithm. We show how, in our case, one can solve the problem directly as a single minimum weighted cut problem. For simplicity, we shall focus on the reflexive case, although the technique applies for irreflexive graphs as well. A similar construction is given in [5], cf. also [24].

Thus suppose that H is a reflexive proper interval graph, with vertices ordered $w_1, w_2, \ldots, w_p$, so that $i < j < k$ and $w_i w_k \in E(H)$ imply $w_i w_j \in E(H)$ and $w_j w_k \in E(H)$. For simplicity we shall write i instead of w_i. We denote, for each i, by $\ell(i)$ the smallest subscript j such that j is adjacent to i; note that $j \leq i$ since H is reflexive. Also note for future reference that if $i' \leq i$, then i' is adjacent to i if and only if $\ell(i) \leq i'$.

Given a graph G with costs $c_i(u), u \in V(G), i \in V(H)$, we construct an auxiliary digraph $G \times H$ as follows. The vertex set of $G \times H$ is $V(G) \times V(H)$ together with two other vertices, denoted by s and t. The directed weighted edges of $G \times H$ are

- an edge from s to $(u, 1)$, of weight ∞, for each $u \in V(G)$,
- an edge from (u, i) to $(u, i+1)$, of weight $c_i(u)$, for each $u \in V(G)$ and $i \in V(H)$,
- an edge from (u, p) to t, of weight $c_p(u)$, for each $u \in V(G)$, and
- an edge from (u, i) to $(v, \ell(i))$, of weight ∞, for every edge $uv \in E(G)$ and each $i \in V(H)$.

(Note that each undirected edge uv of G gives rise to two directed edges $(u, i)(v, \ell(i))$ and $(v, i)(u, \ell(i))$, both of infinite weight, in the last statement.)

A *cut* in $G \times H$ is a partition of the vertices into two sets S and T such that $s \in S$ and $t \in T$; the weight of a cut is the sum of weights of all edges going from a vertex of S to a vertex of T. Let S be a cut of minimum (finite) weight, and define j_u to be the maximum value such that $(u, j_u) \in S$. Let S' be the cut containing s and all $(u, 1), (u, 2), \ldots, (u, j_u)$, for all $u \in V(G)$. If $S' \neq S$, then either the weight of S' is infinite, or at most that of S, as the only edges we might add to the cut are of the form $(u, i)(v, l(i))$. If the weight of S' is infinite, then there must be an edge of the form $(u, i)(v, \ell(i))$ in the cut S', where neither (u, i) nor $(v, \ell(i))$ belong to S. Note that $\ell(i) > j_v$ as $(v, \ell(i)) \notin S'$. Furthermore $\ell(j_u) \geq \ell(i)$, as $j_u > i$, which implies that $\ell(j_u) > j_v$. Therefore the edge $(u, j_u)(v, \ell(j_u))$ belonged to the cut S, which thus had infinite weight, a contradiction. Therefore $S' = S$. Now define a mapping f from $V(G)$ to $V(H)$ by setting $f(u) = j_u$. This must be a homomorphism of G to H; indeed, suppose that $uv \in E(G)$, but $j_u j_v \notin E(H)$. Without loss of generality assume that $j_v \leq j_u$, which implies that $j_v < \ell(j_u)$. This implies that the edge $(u, j_u)(v, \ell(j_u))$ belongs to the cut S, a contradiction. Conversely, any minimum cost homomorphism f of G to H corresponds, in this way, to a minimum weight cut of $G \times H$.

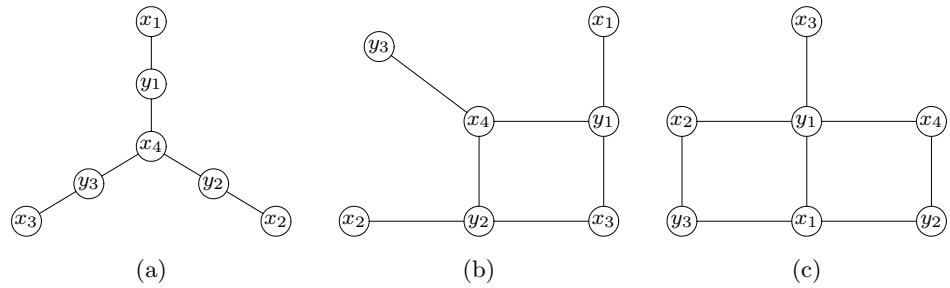

Figure 1: A bipartite claw (a), a bipartite net (b) and a bipartite tent (c).

We conclude that the minimum weight of cut in $G \times H$ is exactly equal to the minimum cost of a homomorphism of G to H. Since minimum weighted cuts can be found by standard flow techniques, we obtain a polynomial time algorithm. Specifically, we note that the graph $G \times H$ has $O(|V(G)||V(H)|)$ vertices. Using the best minimum cut (maximum flow) algorithms, we obtain minimum cost homomorphisms in time $O(|V(G)|^3|V(H)|^3)$ [27]; if H is fixed, and G has n vertices, this is $O(n^3)$.

We observe that this sort of product construction is also similar to the algorithm in [11], which transforms the minimum cost homomorphism problem into a maximum independent set problem in another kind of product $G \otimes H$. (See also Exercise 7 in Chapter 2 of [21].) Note that these kinds of algorithms, which solve the problem via a product construction involving G and H, are polynomial even if H is part of the input.

3 NP-completeness

In this section it will be more convenient to begin with the irreflexive case. Hence all graphs are irreflexive unless stated otherwise (at the end of the section).

A bipartite graph H with vertices $x_1, x_2, x_3, x_4, y_1, y_2, y_3$ is called

a bipartite claw if $E(H) = \{x_4y_1, y_1x_1, x_4y_2, y_2x_2, x_4y_3, y_3x_3\}$;

a bipartite net if $E(H) = \{x_1y_1, y_1x_3, y_1x_4, x_3y_2, x_4y_2, y_2x_2, y_3x_4\}$;

a bipartite tent if $E(H) = \{x_1y_1, y_1x_3, y_1x_4, x_3y_2, x_4y_2, y_2x_2, y_3x_4\}$.

See Figure 3.

These graphs play an important role for proper interval bigraphs. One of the equivalent characterizations is the following result [18].

Theorem 3.1 *[18] A bipartite graph H is a proper interval bigraph if and only if it does not contain an induced cycle of length at least six, or a bipartite claw, or a bipartite net, or a bipartite tent.*

It follows that to show that MinHOM(H) is NP-complete when H is not a proper interval bigraph, it suffices to prove that MinHOM(H) is NP-complete when H is either a cycle of length at least six, or a bipartite claw, or a bipartite net, or a bipartite tent. Indeed, if MinHOM(H) is NP-complete and H is an induced subgraph of H', then MinHOM(H') is also NP-complete, as we may set the costs $c_i(u) = \infty$ for all vertices u of G and all i which are vertices of H' but not of H. The NP-completeness of MinHOM(H) for bipartite cycles of length at least six follows from [7]. In the remainder of this section, we prove that MinHOM(H) is NP-complete for the bipartite claw, net, and tent.

We shall use the following tool.

Theorem 3.2 *The problem of finding a maximum independent set in a 3-partite graph G (even given the three partite sets) is NP-complete.*

Proof: Let $\mathcal{G}_3$ be the set of all graphs of degree at most 3 with at least three vertices excluding K_4. By the well-known theorem of Brooks (see, e.g., [32]), every graph in $\mathcal{G}_3$ is 3-partite. Using Lovasz' constructive proof of Brooks' theorem in [26], one can find three partite sets of a graph $G \in \mathcal{G}_3$ in polynomial time.

Nevertheless, Alekseev and Lozin showed recently in [3] that the problem of finding a maximum independent set in a graph G of $\mathcal{G}_3$ is NP-complete, which completes the proof. ◇

In the rest of this section we will use the notation of Figure 3 for the target graph H. We denote by $\alpha(G)$ the maximal number of vertices in an independent vertex set of a graph G. We will prove the following lemma using a reduction from the problem of finding a maximum independent set in a 3-partite graph.

Lemma 3.3 *If H is a bipartite claw, then MinHOM(H) is NP-complete.*

Proof: Let H be a bipartite claw, with $V(H) = \{x_1, x_2, x_3, x_4, y_1, y_2, y_3\}$ and $E(H) = \{x_4y_1, y_1x_1, x_4y_2, y_2x_2, x_4y_3, y_3x_3\}$ (see Figure 3 (a)). Let G be a 3-partite graph, with partite sets V_1, V_2, V_3. We will now build a graph G^* for which $\text{mch}(G^*, H) = |V(G)| - \alpha(G)$. This will prove the lemma, by Theorem 3.2.

Let G^* be obtained from G by inserting a new vertex m_e into every edge $e \in E(G)$. Note that $V(G^*) = V(G) \cup \{m_e \mid e \in E(G)\}$ and $E(G^*) = \{um_{uv}, m_{uv}v \mid uv \in E(G)\}$. Define costs as follows, where $i \in \{1,2,3\}$ and $j \in \{1,2,3,4\}$.

$c_{x_i}(u) = 0$	if $u \in V_i$	$c_{x_4}(u) = 1$	if $u \in V(G)$
$c_{x_i}(u) = \lvert V(G)\rvert$	if $u \notin V_i$	$c_{y_i}(u) = \lvert V(G)\rvert$	if $u \in V(G)$
$c_{y_i}(m_e) = 0$	if $e \in E(G)$	$c_{x_j}(m_e) = \lvert V(G)\rvert$	if $e \in E(G)$

Let I be an independent set in G, and define a mapping f from $V(G^*)$ to $V(H)$ as follows. For all $u \in V_i$ let $f(u) = x_i$ if $u \in I$ and $f(u) = x_4$ if $u \notin I$. Let $uv \in E(G)$ be arbitrary, and let $f(m_{uv}) = y_i$ if $\{u,v\} \cap (I \cap V_i) \neq \emptyset$, and let $f(m_{uv}) = y_1$ if $x, y \notin I$. Note that f is a homomorphism of G^* to H with cost $|V(G)| - |I|$.

Let f be a homomorphism of G^* to H of cost $|V(G)| - k$. We will now show that there exists an independent set, I in G of order at least k. If $k \leq 0$ then we are trivially done so assume that $k > 0$, which implies that all individual costs in $c(f)$ are either zero or one. Let $I = \{u \in V(G) \mid c_{f(u)}(u) = 0\}$ and note that $|I| \geq k$. Note that I is an independent set in G, as if $uv \in E(G)$, where $u \in I \cap V_i$ and $v \in I \cap V_j$ $(i \neq j)$, then $f(u) = x_i$ and $f(v) = x_j$ which implies that f is not a homomorphism, a contradiction. Therefore I is independent in G.

Observe that we have proved that $\mathrm{mch}(G^*, H) = |V(G)| - \alpha(G)$. Thus, we have now reduced the problem in Theorem 3.2 to MinHOM(H), which completes the proof. $\diamond$

In the proofs of the next two lemmas, we will again use reductions from the problem of finding a maximum independent set in a 3-partite graph.

Lemma 3.4 *If H is a bipartite net, then MinHOM(H) is NP-complete.*

Proof: Let H be a bipartite net, with $V(H) = \{x_1, x_2, x_3, x_4, y_1, y_2, y_3\}$ and $E(H) = \{x_1y_1, y_1x_3, y_1x_4, x_3y_2, x_4y_2, y_2x_2, y_3x_4\}$ (see Figure 3 (b)). Let G be a 3-partite graph, with partite sets V_1, V_2, V_3. We will now build a graph G^* such that $\mathrm{mch}(G^*, H) = 2|V_3| + |V(G)| - \alpha(G)$. This will prove the lemma, by Theorem 3.2.

Let G^* be obtained from G in the following way. For every vertex $v \in V_3$ let $P_v = s_1^v t_1^v s_2^v t_2^v s_3^v$ be a path of length 4. For every $u \in V_1$ and $v \in V_2$ with $uv \in E(G)$ we introduce a new vertex m_{uv}. We set

$$V(G^*) = V_1 \cup V_2 \cup \{m_e \mid e \in E(G)\} \cup \{V(P_v) \mid v \in V_3\}.$$

The edge set of G^* consists of the following edges. For every edge uv between V_1 and V_2 in G both um_{uv} and vm_{uv} belong to G^*. All edges in $V(P_v)$, where $v \in V_3$, belong to G^*. For all $u \in V_1$ and $v \in V_3$, where $uv \in E(G)$, the edge us_1^v belongs to G^*. For all $u \in V_2$ and $v \in V_3$, where $uv \in E(G)$, the edge us_3^v belongs to G^*.

We now define the costs of mapping vertices from $V_1 \cup V_2$ as follows, where all costs not shown are given the value $2|V_3| + |V(G)|$. For each $u \in V_i$, $i = 1, 2$, we set $c_{x_i}(u) = 0$ and $c_{x_4}(u) = 1$. We define the costs of mapping vertices from $V(G^*) - V_1 - V_2$ as follows, where $i \in \{1, 2, 3\}$ and $j \in \{1, 2\}$. For each $e \in E(G)$ and $z \in V(H)$, we set $c_z(m_e) = 0$. Finally, for each $v \in V_3$, we set

$$\begin{array}{llll} c_{y_3}(s_i^v) = 0 & \text{and} & c_q(s_i^v) = 1 & \text{for all} \quad q \in V(H) - y_3; \\ c_{x_4}(t_j^v) = 1 & \text{and} & c_q(t_j^v) = 0 & \text{for all} \quad q \in V(H) - x_4. \end{array}$$

Let I be an independent set in G, and define a mapping f from $V(G^*)$ to $V(H)$ as follows. For each $i = 1, 2$ and $u \in V_i$, let $f(u) = x_i$ if $u \in I$ and $f(u) = x_4$ if $u \notin I$. For every edge uv of G with $u \in V_1$ and $v \in V_2$, let $f(m_{uv}) = y_2$ if $v \in I$ and $f(m_{uv}) = y_1$, otherwise. For all $v \in V_3 \cap I$ let $f(s_1^v) = f(s_2^v) = f(s_3^v) = y_3$ and $f(t_1^v) = f(t_2^v) = x_4$. For all $v \in V_3 - I$ let $f(s_1^v) = f(s_2^v) = y_1$, $f(s_3^v) = y_2$ and $f(t_1^v) = f(t_2^v) = x_3$. Note that f is a homomorphism of G^* to H with cost $2|V_3| + |V(G)| - |I|$.

Let f be a homomorphism from G^* to H of cost $2|V_3| + |V(G)| - k$. We will now show that there exists an independent set I in G of order at least k. If $k \leq 0$ then we are trivially done so assume that $k > 0$, which implies that all individual costs in $c(f)$ are either zero or one. Define I as follows.

$$I = \{u \in V_1 \cup V_2 \mid c_{f(u)}(u) = 0\} \cup \{v \in V_3 \mid f(s_1^v) = f(s_3^v) = y_3\}$$

We will now show that I is independent in G and that $|I| \geq k$. First suppose that $uv \in E(G)$, where $u \in I \cap V_i$ and $v \in I \cap V_j$ $(i \neq j)$. Observe that this is not possible if $\{i, j\} = \{1, 2\}$, so without loss of generality assume that $i < j = 3$. However if $i = 1$ then we cannot have both $f(u) = x_1$ and $f(s_1^y) = y_3$ and if $i = 2$ then we cannot have both $f(u) = x_2$ and $f(s_3^y) = y_3$. Therefore I is independent.

If we could show that the cost of mapping P_v to H (denoted by $c(P_v)$) fulfills (a) and (b) below, then we would be done, as this would imply that $|I| \geq k$.

$$\begin{array}{lll} (a) & c(P_v) \geq 2 & \text{if } v \in I \cap V_3 \\ (b) & c(P_v) \geq 3 & \text{if } v \in V_3 - I \end{array}$$

Indeed,

$$\begin{aligned} c(f) &= \sum_{u \in V_1 \cup V_2} c_{f(u)}(u) + \sum_{v \in V_3} c(P_v) \\ &\geq (|V_1 \cup V_2| - |(V_1 \cup V_2) \cap I|) + 2|V_3 \cap I| + 3(|V_3| - |V_3 \cap I|) \\ &= 2|V_3| + |V(G)| - |I| \end{aligned}$$

and, thus, $|I| \geq k$.

To prove (a) and (b) assume that $v \in V_3$ is arbitrary. Note that $c_{f(s_1^v)}(s_1^v) > 0$ or $c_{f(t_1^v)}(t_1^v) > 0$ (or both), as if $f(s_1^v) = y_3$ then we must have $f(t_1^v) = x_4$. Analogously $c_{f(s_3^v)}(s_3^v) > 0$ or $c_{f(t_2^v)}(t_2^v) > 0$ (or both). This proves (a). If $c_{f(s_2^v)}(s_2^v) > 0$, then $c(P_v) \geq 3$, so assume that $c_{f(s_2^v)}(s_2^v) = 0$, which implies that $f(s_2^v) = y_3$. Thus, $f(t_1^v) = f(t_2^v) = x_4$. If $v \notin I$ then we have $c_{f(s_1^v)}(s_1^v) > 0$ or $c_{f(s_3^v)}(s_3^v) > 0$, which together with $c_{f(t_1^v)}(t_1^v) = c_{f(t_2^v)}(t_2^v) = 1$, implies (b). $\diamond$

Lemma 3.5 *If H is a bipartite tent, then MinHOM(H) is NP-complete.*

Proof: Let H be a bipartite tent with $V(H) = \{x_1, x_2, x_3, x_4, y_1, y_2, y_3\}$ and $E(H) = \{x_4y_1, y_1x_1, x_1y_2, y_2x_4, x_1y_3, y_3x_2, x_2y_1, y_1x_3\}$ (see Figure 3 (c)). Let G be a 3-partite graph, with partite sets V_1, V_2, V_3. We will now build a graph G^* such that $\text{mch}(G^*, H) = |V(G)| - \alpha(G)$. This will prove the lemma, by Theorem 3.2.

Let $E_{1,2}$ denote all edges between V_1 and V_2 in G. A graph G^* is obtained from G, by inserting a new vertex m_e into every edge $e \in E_{1,2}$. Note that $V(G^*) = V(G) \cup \{m_e \mid e \in E_{1,2}\}$. The edge set of G^* consists of all edges in G incident with a vertex in V_3 as well as of the edges $\{u_1v_{u_1u_2}, v_{u_1u_2}u_2 \mid u_1u_2 \in E_{1,2}\}$. We now define the costs of $u_i \in V_i$ as follows, where all costs not shown are given the value $|V(G)|$.

For $i = 1$:	$c_{y_2}(u_1) = 0$	$c_{y_1}(u_1) = 1$
For $i = 2$:	$c_{y_3}(u_2) = 0$	$c_{y_1}(u_2) = 1$
For $i = 3$:	$c_{x_3}(u_3) = 0$	$c_{x_1}(u_3) = 1$

For all edges $e \in E_{1,2}$ let $c_{x_1}(m_e) = |V(G)|$ and let $c_q(m_e) = 0$ for all $q \in V(H) - \{x_1\}$.

Let I be an independent set in G, and define a mapping f from $V(G^*)$ to $V(H)$ as follows.

For $u \in V_1 \cap I$:	$f(u) = y_2$	For $u \in V_1 - I$:	$f(u) = y_1$
For $u \in V_2 \cap I$:	$f(u) = y_3$	For $u \in V_2 - I$:	$f(u) = y_1$
For $u \in V_3 \cap I$:	$f(u) = x_3$	For $u \in V_3 - I$:	$f(u) = x_1$

If $u_1u_2 \in E_{1,2}$ and $u_1 \in V_1 \cap I$, then let $f(m_{u_1u_2}) = x_4$. If $u_2 \in V_2 \cap I$, then let $f(m_{u_1u_2}) = x_2$. If $u_1, u_2 \notin I$ then let $f(m_{u_1u_2}) = x_4$. Note that f is a homomorphism from G^* to H with cost $|V(G)| - |I|$.

Let f be a homomorphism from G^* to H of cost $|V(G)| - k$. We will now show that there exists an independent set, I in G of order at least k. If $k \leq 0$ then we are trivially done so assume that $k > 0$, which implies that all individual costs in f are either zero or one. Let $I = \{u \in V(G) \mid c_{f(u)}(u) = 0\}$ and note that $|I| \geq k$. Furthermore, observe that I is an independent set in G (as $f(v_e) \neq x_1$ for every $e \in E_{1,2}$). We have reduced the problem in Theorem 3.2 to MinHOM(H), which completes the proof. $\diamond$

Corollary 3.6 *If H is an irreflexive graph which is not a proper interval bigraph, then MinHOM(H) is NP-complete.*

Proof: If H is not bipartite, this follows from the fact that the homomorphism problem for H is NP-complete [20]. Otherwise, the conclusion follows from Theorem 3.1, the remarks following it, and the above three lemmas. $\diamond$

We now return to considering graphs with loops allowed. Since we have observed that a graph H with an adjacent loop and nonloop gives rise to an NP-complete problem MinHOM(H), it only remains to prove the NP-completeness of MinHOM(H) when H is a reflexive graph which is not a proper interval graph. We could proceed as before, as there is an analogous result characterizing proper interval graphs by the absence of induced cycles of length at least four, or a claw, net, or tent [31, 10, 29]. However, we instead reduce the problem to the irreflexive case, as follows.

Given a reflexive graph H, we define the bipartite graph H^* with the vertex set $\{v', v'' : v \in V(H)\}$ and edge set $\{u'v'' : uv \in E(H)\}$. (Note that each $v'v''$ is an edge of H since the graph H is reflexive.) It is proved in [18] that H is a proper interval graph if and only if H^* is a proper interval bigraph. Thus suppose a reflexive graph H is not a proper interval graph, and consider the bipartite (irreflexive) graph H^* which is then not a proper interval bigraph. We will now reduce the NP-complete problem MinHOM(H^*) to the problem MinHOM(H) as follows. Each instance of MinHOM(H^*) can also be viewed as an instance of MinHOM(H). Indeed, such an instance consists of a bipartite graph G with costs $c_{i'}(u)$ for each white vertex u of G and white vertex i' of H^*, and costs $c_{i''}(v)$ for each black vertex v of G and black vertex i'' of H^*; to see this as an instance of MinHOM(H), we only need to set $c_i(u)$ equal to $c_{i'}(u)$ if u is white and

$c_{i''}(u)$ if u is black. Now colour-preserving homomorphisms of G to H^* and to H are in a one-to-one correspondence, with the same costs, i.e., there is a homomorphism of G to H^* of cost not exceeding k if and only if there is a homomorphism of G to H of cost not exceeding k. We have proved the following fact.

Corollary 3.7 *If H is a reflexive graph which is not a proper interval graph, then MinHOM(H) is NP-complete.*

Thus, for a connected graph H (with loops allowed), we have the following situation. If H has both loops and nonloops, the problem MinHOM(H) is NP-complete by the remarks after Theorem 1.1; if H is reflexive and not a proper interval graph, then MinHOM(H) is NP-complete by Corollary 3.6; and if H is irreflexive and not a proper interval bigraph, then MinHOM(H) is NP-complete by Corollary 3.7. Of course, as observed earlier, it is enough if this happens for one component of H. Thus we conclude as follows.

Corollary 3.8 *If a graph H (with loops allowed) has a component which is neither a reflexive proper interval graph nor an irreflexive proper interval bigraph, then the problem MinHOM(H) is NP-complete.* ◇

This completes the proof of Theorem 1.1.

4 Digraphs

A digraph H (with loops allowed) satisfying the Min-Max property yields a polynomial time solvable problem MinHOM(H) (Theorem 2.1). However, there are other digraphs H for which the problem MinHOM(H) admits a polynomial solution. For instance, it is easy to see that when H is a directed cycle, we can solve MinHOM(H) in polynomial time, cf. [11]. On the other hand, a directed cycle clearly does not have the Min-Max property, as can be seen by considering the last vertex (in the Min-Max ordering) and its two incident edges. (A similar property more appropriate for cycle-like digraphs is introduced in [14].)

The classification problem for the complexity of minimum cost digraph homomorphism problems remains open. However, in [12], a partial classification has been obtained for the class of *semicomplete k-partite digraphs*. These are digraphs that can be obtained from undirected complete k-partite graphs by orienting each undirected edge in one direction or in both directions. When $k \geq 3$, the classification is given in [12]. When $k = 2$, the

situation is more complex, and the classification has only recently been completed [14]. The full classification of all minimum cost digraph homomorphism problems remains open. On the other hand, the dichotomy of *list* homomorphism problems for digraphs follows from a result of Bulatov [4].

Acknowledgements The research of the first author was supported in part by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778. The research of the second and third author was supported in part by an NSERC Discovery Grant.

References

[1] G. Aggarwal, T. Feder, R. Motwani, and A. Zhu, Channel assignment in wireless networks, and classification of minimum graph homomorphisms, Manuscript 2005.

[2] N. Alon, W. Fernandez de la Vega, R. Kannan, and M. Karpinski, Random sampling and approximation of MAX-CSP problems. *J. Comput. Syst. Sci.* 67 (2003), 212–243.

[3] V.E. Alekseev and V.V. Lozin, Independent sets of maximum weight in (p, q)-colorable graphs, *Discrete Mathematics* 265 (2003), 351–356.

[4] A.A. Bulatov, Tractable conservative constraint satisfaction problems. To appear in *ACM Trans. Comput. Logic*.

[5] D. Cohen, M. Cooper, P. Jeavons, and A. Krokhin, A maximal tractable class of soft constraints. *J. Artif. Intell. Res.* 22 (2004), 1–22.

[6] X. Deng, P. Hell and J. Huang, Linear-time representation algorithms for proper circular arc graphs and proper interval graphs, SIAM J. Comput. 25 (1996) 390–403.

[7] T. Feder, P. Hell and J. Huang, List homomorphisms and circular arc graphs. *Combinatorica* 19 (1999), 487–505.

[8] T. Feder, P. Hell, and J. Huang, Bi-arc graphs and the complexity of list homomorphisms, *J. Graph Theory* 42 (2003), 61–80.

[9] S. Fujishige, Submodular functions and optimization, North-Holland, Amsterdam, 1991.

[10] M. Golumbic, Algorithmic graph theory and perfect graphs, Elsevier, Amsterdam, 2004.

[11] G. Gutin, A. Rafiey and A. Yeo, Minimum cost and list homomorphisms to semicomplete digraphs, *Discrete Appl. Math.* 154 (2006) 881–889.

[12] G. Gutin, A. Rafiey and A. Yeo, Minimum cost homomorphisms to semicomplete multipartite digraphs, submitted.

[13] G. Gutin, A. Rafiey, A. Yeo and M. Tso, Level of repair analysis and minimum cost homomorphisms of graphs, *Discrete Appl. Math.* 154 (2006) 890–897.

[14] G. Gutin, A. Rafiey and A. Yeo, Minimum cost homomorphisms to semicomplete bipartite digraphs, submitted.

[15] W. Gutjahr, E. Welzl and G. Woeginger, Polynomial graph-colorings, *Discrete Applied Math.* 35 (1992) 29–45.

[16] M. M. Halldorsson, G. Kortsarz, and H. Shachnai, Minimizing average completion of dedicated tasks and interval graphs. Approximation, Randomization, and Combinatorial Optimization (Berkeley, Calif, 2001), Lecture Notes in Computer Science, vol. 2129, Springer, Berlin, 2001, pp. 114–126.

[17] P. Hell and J. Huang, Certifying LexBFS recognition algorithms for proper inteval graphs and proper interval bigraphs. *SIAM J. Discrete Math.* 18 (2005), 554–570.

[18] P. Hell and J.Huang, Interval bigraphs and circular arc graphs. *J. Graph Theory* 46 (2004), 313–327.

[19] P. Hell, Algorithmic aspects of graph homomorphisms, in 'Survey in Combinatorics 2003', London Math. Soc. Lecture Note Series 307, Cambridge University Press, 2003, 239–276.

[20] P. Hell and J. Nešetřil, On the complexity of H-colouring. *J. Combin. Theory B* 48 (1990), 92–110.

[21] P. Hell and J. Nešetřil, *Graphs and Homomorphisms.* Oxford University Press, Oxford, 2004.

[22] K. Jansen, Approximation results for the optimum cost chromatic partition problem. *J. Algorithms* 34 (2000), 54–89.

[23] T. Jiang and D. B. West, Coloring of trees with minimum sum of colors. *J. Graph Theory 32* (1999), 354–358.

[24] S. Khanna, M. Sudan, L. Trevisan and D. Williamson, The approximability of constraint satisfaction problems. *SIAM J. Comput.* 30 (2000), 1863–1920.

[25] L. G. Kroon, A. Sen, H. Deng, and A. Roy, The optimal cost chromatic partition problem for trees and interval graphs, Graph-Theoretic Concepts in Computer Science (Cadenabbia, 1996), Lecture Notes in Computer Science, vol. 1197, Springer, Berlin, 1997, pp. 279–292.

[26] L. Lovasz, Three short proofs in graph theory, *J. Combin. Theory, Ser. B* 19 (1975), 269–271.

[27] A. Schrijver, *Combinatorial Optimization*, Springer 2003.

[28] J.P. Spinrad, A. Brandstaedt, L. Stewart, Bipartite permutation graphs, Discrete Applied Math. 18 (1987) 279–292.

[29] J. Spinrad, *Efficient Graph Representations* AMS 2003.

[30] K. Supowit, Finding a maximum planar subset of a set of nets in a channel. *IEEE Trans. Computer-Aided Design* 6 (1987), 93–94.

[31] G. Wegner, Eigenschaften der nerven homologische-einfactor familien in R^n, Ph.D. thesis, Universität Gottigen, Gottingen, Germany (1967).

[32] D. West, *Introduction to Graph Theory*, Prentice Hall, Upper Saddle River, 1996.