Some problems not definable using structure homomorphisms

Florent Madelaine [a,1] and Iain A. Stewart [a,1]

[a] *Department of Mathematics and Computer Science, Leicester University, Leicester LE1 7RH, U.K.*

Abstract

We exhibit some problems definable in Feder and Vardi's logic MMSNP that are not in the class CSP of constraint satisfaction problems. Whilst some of these problems have previously been shown to be in MMSNP (that is, definable in MMSNP) but not in CSP, existing proofs are probabilistic in nature. We provide explicit combinatorial constructions to prove that these problems are not in CSP and we use these constructions to exhibit yet more problems in MMSNP that are not in CSP.

1 Introduction

Many problems, i.e., classes of finite structures over some relational signature (throughout, our signatures are finite and only ever consist of relation symbols), can be formulated as constraint satisfaction problems where the instances of a *constraint satisfaction problem* are pairs (I, T) of finite structures over some fixed relationial signature, with the subset of yes-instances consisting of those pairs (I, T) for which there is a homomorphism from I to T, i.e., a mapping φ from the domain of I to the domain of T such that for every relation R^I, of arity a, of I and every a-tuple $\mathbf{u}$ of elements from the domain of I, if $\mathbf{u} \in R^I$ then $\varphi(\mathbf{u}) \in R^T$ (where R^T is the corresponding relation of T). In practice, one often encounters the situation where the finite structure T, the *template*, is fixed and only the finite structure I varies. For example, if we consider finite structures over $\sigma_2 = \langle E \rangle$, where E is a relation symbol of arity 2, and regard a finite σ_2-structure as an undirected graph via "there is an edge joining vertex u and vertex v if, and only if,

[1] Supported by EPSRC Grant GR/M12933.

either $E(u, v)$ or $E(v, u)$ holds" (and we ignore self-loops) then if T is the σ_2-structure $\langle\{0, 1, 2\}, \{(0, 1), (1, 0), (1, 2), (2, 1), (0, 2), (2, 0)\}\rangle$ then we have essentially a realisation of the well-known abstract decision problem consisting of those graphs that can be properly 3-coloured.

The class of problems CSP, as defined in [1], is the class of constraint satisfaction problems where there is a fixed template. This class has some interesting subclasses. Schaefer [3] showed that every problem in the sub-class of CSP where the template is restricted to be a boolean structure, i.e., a finite structure with just two elements, is either in **P** or is **NP**-complete; that is, there is a dichotomy. Hell and Nešetřil [2] showed that every problem in the sub-class of CSP where the template is restricted to be a σ_2-structure is either in **P**, if the template is bipartite, or **NP**-complete otherwise (where we identify a graph as a σ_2-structure as above). It is unknown whether every problem of CSP is either in **P** or **NP**-complete. However, these "dichotomy" results for CSP have prompted the following question: "*What is the most general sub-class of* **NP** *with the property that every problem in this class is either in* **P** *or* **NP**-*complete*?". Whilst Feder and Vardi [1] failed to answer this question completely, they did provide evidence that a logically-defined sub-class of **NP**, which they termed MMSNP or *monotone monadic SNP without inequality*, might exhibit the dichotomy in question.

A problem, over some signature σ, is in MMSNP if, and only if, it can be defined by a sentence of monadic existential second-order logic where the first-order part consists of a universally quantified quantifier-free formula in which the equality relation is forbidden and in which every occurrence of a relation symbol from the signature σ does not appear within the scope of an even number of negations. That is, any problem in MMSNP, over some signature σ, can be defined by a second-order sentence of the form:

$$\exists M_1 \exists M_2 \ldots \exists M_k \forall x_1 \forall x_2 \ldots \forall x_m (C_1 \wedge C_2 \wedge \ldots \wedge C_t),$$

with each M_i a new relation symbol of arity 1 and each C_j a conjunction of atoms and negated atoms involving the relation symbols of $\{M_1, M_2, \ldots, M_k\}$ and negated atoms involving the relation symbols of σ. The reader is referred to [1] for more details.

Amongst other things, Feder and Vardi proved that: every problem in CSP is in MMSNP; there are problems in MMSNP that are not in CSP; and every problem in MMSNP has a randomised polynomial-time equivalent problem in CSP. Hence, if "randomised polynomial-time" in the latter result could be replaced with "polynomial-time" (which is not an unreasonable expectation) then CSP has a dichotomy if, and only if, MMSNP has a dichotomy. However, it is with the fact that MMSNP properly contains CSP that we are concerned in this brief note. Realisations (as problems over some signature) of the abstract decision problems consisting of those triangle-free undirected graphs and of those undirected graphs for which there exists a 2-colouring so that the vertices of every triangle are monochromatically coloured (the problems TRI-FREE and NO-MONO-TRI, defined later) were the problems that Feder and Vardi showed to be in MMSNP but not in CSP. However, the arguments produced by Feder and Vardi were rather scant and also probabilistic in nature; that is, the existence of a suitable infinite class of graphs was shown to exist but not explicitly constructed. In this note, we prove that these problems are in MMSNP but not CSP, but we prove these facts by explicit combinatorial constructions of suitable infinite classes of graphs. Whilst the probabilistic method can be very powerful, our explicit constructions give rise to proofs that other problems are in MMSNP but not CSP. One cannot deduce these facts from Feder and Vardi's arguments.

2 Some separating problems

We begin by defining some problems in MMSNP and then go on to prove that these problems are not in CSP.

The problem TRI-FREE is the problem over σ_2 defined by the following first-order sentence:

$$\forall x(\neg E(x,x)) \wedge \forall x\forall y\forall z(\neg(E(x,y) \vee E(y,x)) \vee \neg(E(x,z) \vee E(z,x))$$
$$\vee\neg(E(y,z) \vee E(z,y))).$$

Note that the above sentence is also a sentence of MMSNP and that TRI-FREE can be considered to be a realisation of the abstract decision problem consisting of those undirected graphs in which there is no triangle.

The problem NO-MONO-TRI is the problem over σ_2 defined by the following sentence of MMSNP:

$$\exists C(\forall x(\neg E(x,x)) \wedge \forall x\forall y\forall z(((E(x,y) \vee E(y,x)) \wedge (E(x,z) \wedge E(z,x))$$
$$\wedge(E(y,z) \vee E(z,y))) \Rightarrow (\neg(C(x) \wedge C(y) \wedge C(z)) \wedge \neg(\neg C(x) \wedge$$
$$\neg C(y) \wedge \neg C(z))))).$$

Note that the problem NO-MONO-TRI can be considered as a realisation of the abstract decision problem consisting of those undirected graphs for which there exists a 2-colouring of the vertices so that the vertices of every triangle in the graph are not monochromatically coloured.

The problem TRI-FREE-TRI is the problem over σ_2 defined by the following sentence of MMSNP:

$$\exists R \exists W \exists B(\forall x((R(x) \wedge \neg W(x) \wedge \neg B(x)) \vee (\neg R(x) \wedge W(x) \wedge \neg B(x)) \vee(\neg R(x) \wedge \neg W(x) \wedge B(x))) \wedge \forall x \forall y((E(x,y) \vee E(y,x)) \Rightarrow (\neg(R(x) \wedge R(y)) \wedge \neg(W(x) \wedge W(y)) \wedge \neg(B(x) \wedge B(y)))) \wedge \forall x(\neg E(x,x)) \wedge\forall x \forall y \forall z(\neg(E(x,y) \vee E(y,x)) \vee \neg(E(x,z) \vee E(z,x)) \vee \neg(E(y,z) \vee E(z,y)))).$$

Note that the problem TRI-FREE-TRI can be considered as a realisation of the abstract decision problem consisting of those undirected graphs that are tripartite and in which there is no triangle; that is, as a restriction of TRI-FREE to tripartite graphs.

The problem NO-WALK-5 is the problem over σ_2 defined by the following first-order sentence:

$$\forall x(\neg E(x,x)) \wedge \forall x_1 \forall x_2 \forall x_3 \forall x_4 \forall x_5(\neg((E(x_1,x_2) \vee E(x_2,x_1)) \wedge (E(x_2,x_3) \vee E(x_3,x_2)) \wedge (E(x_3,x_4) \vee E(x_4,x_3)) \wedge (E(x_4,x_5) \vee E(x_5,x_4)) \wedge(E(x_5,x_1) \vee E(x_1,x_5)))).$$

Note that the above sentence is also a sentence of MMSNP and that NO-WALK-5 can be considered to be a realisation of the abstract decision problem consisting of those undirected graphs in which there is no closed walk of length 5. The problem NO-WALK-7 is defined similarly (and is also in MMSNP). Moreover, the problems NO-WALK-5-TRI and NO-WALK-7-TRI, being the restrictions of NO-WALK-5 and NO-WALK-7 to tripartite graphs, as above, are in MMSNP too.

In what follows, we construct infinite classes of graphs with specific properties. Henceforth, when we talk about an undirected graph G, we actually mean the unique σ_2-structure $\mathcal{A}$ where:

- $|\mathcal{A}|$ is the set of vertices of G;
- $\mathcal{A} \models \forall x(\neg E(x,x))$;
- there is an edge joining u and v in G if, and only if, $E(u,v)$ and $E(v,u)$ are both true; and
- there is not an edge joining u and v in G if, and only if, both $E(u,v)$ and $E(v,u)$ are false.

Our first observation is that if the template defining a problem in CSP over σ_2 has a self-loop then the problem must consist of the class of all σ_2-structures. Hence, we may assume that any template has no self-loops. Our second observation is that the template defining a problem in CSP over σ_2 must be in the problem (as the identity map from the template to the template is a homomorphism).

Lemma 1 *Let G be a triangle-free graph and let T be a σ_2-structure such* *that:*

$$\forall x(\neg E(x,x)) \wedge \forall x \forall y \forall z(\neg(E(x,y) \vee E(y,x)) \vee \neg(E(x,z) \vee E(z,x)) \vee \neg(E(y,z) \vee E(z,y)));$$

i.e., $T \in$ TRI-FREE. Furthermore, suppose that in G, there is a path of length 3 joining two non-adjacent vertices u and v. Then any homomorphism φ from G to T is such that $\varphi(u) \neq \varphi(v)$.

Proof. Suppose that there is a homomorphism φ from G to T such that $\varphi(u) = \varphi(v)$. By definition, there is a path u, w_1, w_2, v in G. Because T has no self-loops, we must have that $\varphi(u)$, $\varphi(w_1)$ and $\varphi(w_2)$ are distinct in T and $E(\varphi(u), \varphi(w_1))$, $E(\varphi(w_1), \varphi(w_2))$ and $E(\varphi(w_2), \varphi(u))$ hold in T. This yields a contradiction. □

Suppose that some problem Ω over σ_2 is such that:

- every σ_2-structure in Ω is in TRI-FREE; and
- for every n, Ω contains a graph H_n with n mutually non-adjacent vertices where there is a path of length 3 joining every pair of such vertices.

Then, by Lemma 1, Ω is not in CSP (any homomorphism from H_n to the template must have an image of size at least n).

Define the graph H_n as follows. The vertices of H_n consist of the union of the sets:

- $V_n = \{1, 2, \ldots, n\}$;
- $U_n^1 = \{(i,j) : 1 \leq i,j \leq n, i < j\}$; and
- $U_n^2 = \{(i,j) : 1 \leq i,j \leq n, i > j\}$.

The edges of H_n consist of the union of the sets:

- $\{(i,(i,j)) : 1 \leq i,j \leq n, i < j\}$;
- $\{(i,(i,j)) : 1 \leq i,j \leq n, i > j\}$; and
- $\{((i,j),(j,i)) : 1 \leq i,j \leq n, i \neq j\}$.

The graph H_n can be depicted as in Fig. 1. Note that: H_n is triangle-free; there is a path of length 3 joining any two distinct vertices of V_n; V_n forms an independent set in H_n; and H_n is tripartite.

Lemma 2 *There does not exist a closed walk of length 5 or 7 in H_n.*

Proof. Suppose that there exists a closed walk W of length 5 or 7 in H_n. As H_n is tripartite, W must have at least one vertex, w_1 say, in V_n. Hence, there is $w_2 \in V_n \setminus \{w_1\}$ such that either:

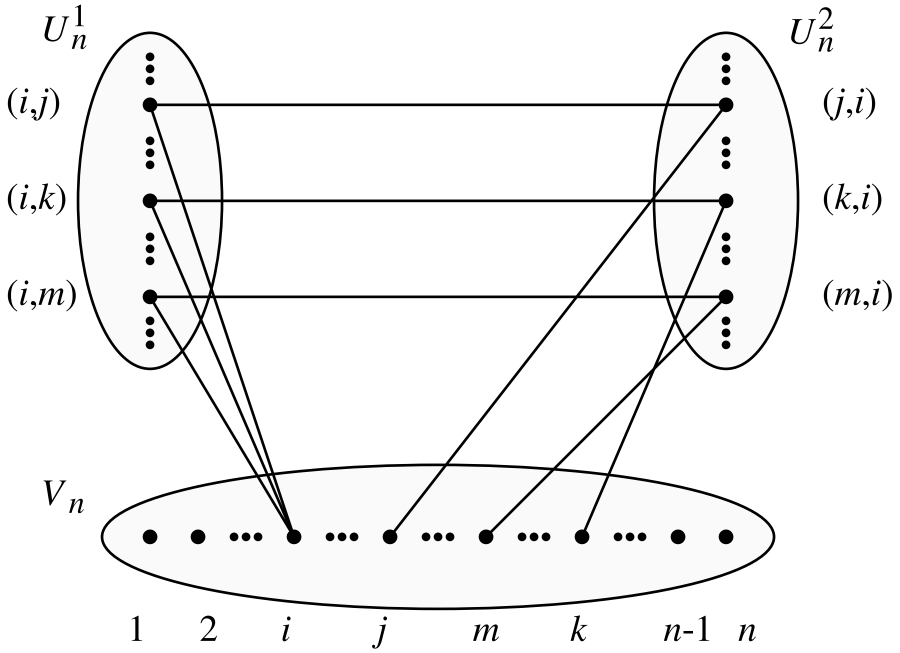

Figure 1. The graph H_n.

(1) w_1, (w_1, w_2), (w_2, w_1) is a sub-walk of W; or
(2) w_1, (w_1, w_2), w_1 is a sub-walk of W.

Suppose that the length of W is 5. In case (1), we obtain a contradiction as every vertex of U_n^1 and U_n^2 is joined to exactly one vertex of V_n. In case (2), we also obtain a contradiction as this would imply that H_n has a triangle. Hence, H_n has no closed walk of length 5.

Suppose that the length of W is 7. In case (1), we must have a closed walk of length 4 between w_1 and w_2. As every vertex of U_n^1 and U_n^2 has exactly one neighbour in V_n, this yields a contradiction. Case (2) yields a contradiction as it implies that there must be a closed walk of length 5 in H_n. □

Our observation immediately after the proof of Lemma 1 yields the following corollary.

Corollary 3 *The problems TRI-FREE, TRI-FREE-TRI, NO-WALK-5, NO-WALK-7, NO-WALK-5-TRI and NO-WALK-7-TRI are in MMSNP but not in CSP.* □

This only leaves the problem NO-MONO-TRI. Let G_n be obtained from H_n

by adding in two extra vertices, a and b, and joining both a and b to every other vertex (this means that we have an edge (a,b) too).

Lemma 4 *Suppose that u and v are vertices of V_n in the graph G_n and let T be a σ_2-structure in NO-MONO-TRI such that there is a homomorphism φ from G_n to T. Then $\varphi(u) \neq \varphi(v)$.*

Proof. Suppose that $\varphi(u) = \varphi(v)$. By arguing as in Lemma 1, there are vertices w_1 and w_2 of $G_n \setminus \{a,b\}$ such that $\varphi(w_1) \neq \varphi(u) \neq \varphi(w_2)$. Also, both $\varphi(a)$ and $\varphi(b)$ must different from the image of any other vertex of G_n. Hence, $E(x,y)$ holds in T for every distinct pair of elements x and y from the set $\{\varphi(u), \varphi(w_1), \varphi(w_2), \varphi(a), \varphi(b)\}$ of 5 elements. We obtain a contradiction as this implies that $T \notin$ NO-MONO-TRI. □

Lemma 5 *For every $n \geq 2$, $G_n \in$ NO-MONO-TRI.*

Proof. Colour the elements a and b 'black' and the other elements 'white'. □

By arguing as above, we immediately obtain the following.

Corollary 6 *NO-MONO-TRI is in MMSNP but not in CSP.* □

References

[1] T. Feder and M.Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: a study through Datalog and group theory. *SIAM J. Computing*, 28:57–104, 1998.

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

[3] T.J. Schaeffer. The complexity of satisfiability problems. In *Proc. 10th ACM Symp. on Theory of Computing*, pages 216–226. ACM, 1978.