The Minimal Logically-Defined NP-Complete Problem

Régis Barbanchon and Etienne Grandjean

GREYC, Université de Caen, 14032 Caen Cedex, France
{regis.barbanchon,etienne.grandjean}@info.unicaen.fr

Abstract. We exhibit an NP-complete problem defined by an existential monadic second-order (EMSO) formula over *functional structures* that is:

1. *minimal* under several syntactic criteria (i.e., any EMSO formula that further strengthens any criterion defines a PTIME problem even if all other criteria are weakened);
2. *unique* for such restrictions, up to renamings and symmetries.

Our reductions and proofs are surprisingly very elementary and simple in comparison with some recent similar results classifying existential second-order formulas over *relational structures* according to their ability either to express NP-complete problems or to express only PTIME ones.

Key words: Computational complexity, descriptive complexity, finite model theory, second-order logic, NP-completeness, parsimony.

1 Introduction and main results

1.1 Which formulas express NP-complete problems?

In the line of Fagin's Theorem [5] which states that *existential second order logic* (ESO) captures the class NP, this paper studies the following natural question: what is (are) the most simple ESO sentence(s) that define(s) some NP-complete problem(s)? This question is somewhat related to two recent papers [7, 4] that completely classified prefix classes of ESO over strings and graphs (and more generally over *relational structures*) with respect to their ability to express either some NP-complete problems or only tractable (i.e., PTIME) ones. For example, it is easy to express an NP-complete problem over graphs, such as 3-colourability, in *existential monadic second-order logic* (EMSO) with only *two* first-order variables. In contrast, one notices that ESO formulas that use only *relation* ESO variables and only *one* first-order variable can only define easy (degenerate) properties on *relational structures*. The situation completely changes if *function* symbols are allowed either in the input signature or among the ESO symbols. For example, ESO formulas with *only one first-order variable x* of one of the forms (1-2)

$$(1)\quad \langle D, E\rangle \models \quad \exists \overline{f}\ \forall x \quad \psi(x, \overline{f}, E)$$

$$(2)\quad \langle D, \overline{f}\rangle \models \quad \exists \overline{U}\ \forall x \quad \psi(x, \overline{f}, \overline{U})$$

(where x is quantified over the finite domain D, E is a binary relation symbol,

$\overline{f}$ and $\overline{U}$ are lists of unary function symbols and of monadic relation symbols respectively, and ψ is quantifier-free) can express some NP-complete problems. More precisely, [9] has recently proved that formulas of form (1) exactly define graph problems (such as the Hamiltonian cycle problem) that are recognizable in nondeterministic linear time $O(n)$ where n is the number of vertices in the graph, and [1] states that any problem is linearly reducible to SAT iff it is linearly reducible to some problem expressible by some formula of the form (2) (see also [12]). Moreover, as proved in [1], it can be assumed that the unary functions $\overline{f}$ of the input structures are permutations: the class of such problems are called LIN-LOCAL since they are linearly reducible to *local* problems.

In this paper, we exhibits a formula of the form (2) over *functional structures* that defines some NP-complete problem and is *minimal* for several criteria over the signature of input structures, the prefix and the matrix of the formula; more precisely, the further strengthening of any criterion makes the problem fall in PTIME even if all others criteria are weakened. Moreover, this problem is essentially *unique*, up to renamings and symmetries.

Finally, in contrast with our results about *functional structures*, notice that the similar question of determining the minimal ESO formulas (with two first-order variables) that define NP-complete problems over *relational structures* is, to our knowledge, widely open and seems rather difficult to us: e.g., the unicity of such a formula is very dubious.

1.2 Minimal formulas for NP-complete problems

We study the problem $\textsc{Min}_0$ defined by the *very simple* EMSO formula φ_0 of the particular form (2) that follows.

Notation 1. *Let φ_0 denote the $\{f,g\}$-formula in* conjunctive normal form *(CNF)*

$$\varphi_0 : \quad \exists U\ \forall x \quad \psi_0(x) \qquad \textit{where } \psi_0 \textit{ is the conjunction}$$

$$\psi_0 : \quad (Ux \vee Ufx) \wedge (\neg Ux \vee \neg Ufx \vee \neg Ugx),$$

and f, g are unary function symbols. Let δ_0 denote the following formula in disjunctive normal form *(DNF) which is logically equivalent to φ_0*

$$\delta_0 : \quad \exists U\ \forall x \quad (Ux \wedge \neg Ufx) \vee (Ux \wedge \neg Ugx) \vee (\neg Ux \wedge Ufx).$$

The problem $\textsc{Min}_0$ *is defined as the set of finite models $\langle D, f, g\rangle$ of φ_0 (or of δ_0).*

We shall also study the following subproblems of $\textsc{Min}_0$:

Notation 2. *Define* $\textsc{Min}_1$ *as the set of finite models $\langle D, f, g\rangle$ of φ_0, where f and g are permutations of D. For some functional structure $\langle D, f, g\rangle$, let $G(D, f, g)$ denote the graph (V, E) defined by $V = D$ and $E = \{(x, fx) : x \in D\} \cup \{(x, gx) : x \in D\} \cup \{(fx, gx) : x \in D\}$. Define* $\textsc{Min}_2$ *as the set of finite models $\langle D, f, g\rangle$ of φ_0, where f and g are permutations of D and $G(D, f, g)$ is planar.*

Our main results use the following notations:

Notation 3. *The* atoms *of a formula are its atomic subformulas. In particular, the distinct atoms of φ_0 (or δ_0) are Ux, Ufx and Ugx. The* length *of a formula is the total number of occurrences of atoms in it. The disjuncts of a DNF formula are called its* anticlauses*.*

Theorem 1 (NP-completeness). *Min$_0$, Min$_1$ and Min$_2$ are NP-complete.*

Theorem 2 (Minimality). *If $P \neq NP$, φ_0 (resp. δ_0) is, for the syntactic criteria enumerated in the table below, a* minimal *EMSO formula in CNF (resp. in DNF) of the form $\exists \overline{U}\ \forall \overline{x}\ \psi$ (where ψ is quantifier-free and $\overline{x}$ is a list of first-order variables) that defines an NP-complete problem over functional structures.*

input signature	2 unary functions	distinct atoms	3
EMSO symbols	1	clauses in CNF φ_0	2
FO variables	1	length of CNF φ_0	5
compositions of functions	0	anticlauses in DNF δ_0	3
equalities	0	length of DNF δ_0	6

That means that if any of these criteria is strenghtened and the other criterias are weakened then the problem so defined is PTIME; e.g, any formula of the form $\exists \overline{U}\ \forall \overline{x}\ \psi$ with length of $\psi < 5$ in CNF defines a PTIME problem.

Theorem 3 (Unicity). *If P$\neq$NP, φ_0 (resp. δ_0) is – up to symmetries – the* unique *minimal EMSO formula in CNF (resp. in DNF) of the form $\exists \overline{U}\ \forall \overline{x}\ \psi$ (where ψ is quantifier-free) that defines an NP-complete problem over functional structures. The symmetrical formulas are obtained by any permutation of the terms x, fx and gx and by swap of U and $\neg U$ in φ_0 (resp. δ_0).*

More precisely, all the symmetrical formulas of φ_0 essentially define the *same* minimal NP-complete problem over *permutations* (resp. planar permutations) structures. In case of (general) *functional* structures, one obtains essentially *two* minimal NP-complete problems: the one defined by φ_0 itself, and the one defined by the following formula φ_0', that is φ_0 with terms x and gx permuted:

$$\varphi_0' : \quad \exists U \quad \forall x \quad (Ugx \vee Ufx) \wedge (\neg Ugx \vee \neg Ufx \vee \neg Ux)$$

1.3 Minimal formulas for #P-complete problems

Besides NP-completeness, another important concept of the theory of complexity is #P-completeness [14]. It is also natural to look for a minimal logical formula that defines some #P-complete problem. In this regard, it is well known that the generic reduction from any NP problem to Sat can (easily) be made parsimonious with a bijective and PTIME-computable correspondence between solutions. That means that the problem Sat not only "simulates" the decision process of any problem in NP but also "reproduces" the number of its solutions and the "structure" of this set of solutions.

Notation 4. *For any problem A in NP, let us denote by $\#A$ the "natural" counting problem associated to A, i.e., the problem of counting the "natural" solutions of the instances of A. #P is the class of such counting problems; e.g.,* #Sat *is the function which maps each propositional formula F to the number of assignments I over the variables of F such that $I \models F$; similarly,* #Min$_1$ *is the function which maps each permutation structure $\mathcal{S} = \langle D, f, g \rangle$ to the number of predicates U such that $(\mathcal{S}, U) \models \forall x\ \psi_0(x)$.*

We say that an ordered pair (ρ,μ) is a *weakly parsimonious* reduction from #A to #B if ρ is a PTIME reduction from A to B, μ is a PTIME-computable function valued in positive integers such that for each instance w of A we have $\#\{\text{solutions of A for } w\} = \mu(w) \times \#\{\text{solutions of B for } \rho(w)\}$. If furthermore $\mu = 1$, then ρ is called a *parsimonious reduction.* We conjecture that:

Conjecture 1. *There exists no* parsimonious *reduction from problem* #SAT *to problems* $\#\text{MIN}_1$ *or* $\#\text{MIN}_2$.

Nevertheless, we prove in this paper that:

Theorem 4. *There exists* weakly parsimonious *reductions from problem* #SAT *to problems* $\#\text{MIN}_1$ *and* $\#\text{MIN}_2$.

In regard to Conjecture 1 concerning Formula φ_0, it is natural to look for another simple EMSO formula defining a problem to which SAT (and hence any NP problem) *parsimoniously* reduces. Let φ_{nand} denote the $\{f,g\}$-formula

$$\begin{array}{lll} \varphi_{\text{nand}}: & \exists U\ \forall x \quad \psi_{\text{nand}}(x) & \text{where } \psi_{\text{nand}} \text{ is} \\ \psi_{\text{nand}}: & Ux \iff \neg(Ufx \wedge Ugx) & \text{or equivalently in CNF} \\ \psi_{\text{nand}}: & (Ux \vee Ufx) \wedge (Ux \vee Ugx) \wedge (\neg Ux \vee \neg Ufx \vee \neg Ugx). & \end{array}$$

Clearly, ψ_{nand} (resp. φ_{nand}) implies ψ_0 (resp. φ_0). The formula φ_{nand} defines the following problems:

Notation 5. *Define* NAND_1 *as the set of finite models* $\langle D,f,g\rangle$ *of* φ_{nand} *where* f *and* g *are permutations of* D. *Define* NAND_2 *as the set of finite models* $\langle D,f,g\rangle$ *of* φ_{nand} *where* f *and* g *are permutations of* D *and* $G(D,f,g)$ *is planar.*

In contrast to Conjecture 1, we can prove that:

Theorem 5. (**i**) #SAT parsimoniously reduces *to* $\#\text{NAND}_1$ *(resp.* $\#\text{NAND}_2$*).* (**ii**) *If Conjecture 1 holds and P ≠ NP, then* φ_{nand} *is (up to symmetries) the* unique minimal *EMSO formula for which (i) holds, i.e., that defines a problem over permutation structures* $\langle D,f,g\rangle$ *to which* #SAT *parsimoniously reduces.*

Surprisingly, our proofs of completeness are rather simple and the reductions involved in Theorems 1 and 5 are essentially the same one reduction $\rho : F \mapsto \mathcal{S}(F)$ described in the next section.

2 Proofs of our results

2.1 The structures involved

Let us recall the three kinds of instances of our problems.

Definition 1. *A* function structure *is a finite structure* $\langle D,f,g\rangle$ *where* $f,g : D \longrightarrow D$ *are unary functions. A function structure* $\langle D,f,g\rangle$ *is a* permutation structure *(resp. is a planar permutation structure) if* f,g *are permutations of* D *(resp. are permutations of* D *such that the graph* $G(D,f,g)$ *is planar).*

Remark 1. *A permutation structure $\langle D, f, g\rangle$ is naturally given by its f- and g-*circuits*, where an f-circuit of length k is an orbit $a, fa, f^2a, \cdots, f^ka = a$.*

Definition 2 (Planar formula and Plan-Sat). *Let F be a propositional formula in CNF. Let $G(F)$ denote the following bipartite graph (V, E) where V is the disjoint union of the set of variables and the set of clauses of F, and E is the set of pairs (v, C) such that v is a variable that occurs in clause C. F is a planar formula if $G(F)$ is a planar graph, and* Plan-Sat *is defined as the satisfiability problem of planar formulas.*

Our proofs of completeness use the NP-complete problem Plan-Sat [13].

2.2 A gadget structure

We are going to describe a reduction $\rho : F \mapsto \mathcal{S}(F)$ that associates to each Sat (resp. Plan-Sat) instance F a permutation structure $\mathcal{S}(F)$ that contains as substructures many occurrences of the following gadget denoted True whose role is essential in our reduction.

Definition 3. True *or* True(α, β, γ) *is the gadget depicted on the left of Fig. 1.*

The symbolization means that the gadget True plays the role of the Boolean constant "true" (or "1"). More formally, the following lemma expresses that in any case, $U(\gamma)$ *can* and *should* be true whereas the value of $U(g\gamma)$ (reached via the "pending" outgoing g-edge of γ) is *free*.

Lemma 1. *Let* True(α, β, γ) *be a gadget included in a permutation structure $\mathcal{S} = \langle D, f, g\rangle$ and $U : D \longrightarrow \{0, 1\}$ be a monadic predicate*[1]*.*

1. *If $(\mathcal{S}, U) \models \varphi_0$ then we have $U(\alpha) = 1$, $U(\beta) = 0$ and $U(\gamma) = 1$;*
2. *Conversely: if $U(\alpha) = 1$, $U(\beta) = 0$ and $U(\gamma) = 1$, then the structure $(True, U)$ satisfies $\forall x\ \psi_{\text{nand}}$ (and hence $\forall x\ \psi_0$); in other words, $\psi_{\text{nand}}(x)$ is satisfied by each element $x = \alpha, \beta, \gamma$ independently of the value of $U(g\gamma)$.*

Proof. Easy and left to the reader. □

2.3 Our reduction

Let us now construct our reduction $\rho : F \mapsto \mathcal{S}(F)$ where F is a Sat (resp. Plan-Sat) instance, i.e., a conjunction of clauses $F = C_1 \wedge C_2 \wedge \cdots \wedge C_q$. In the description of the permutation structure $\mathcal{S}(F)$, we freely make use of the following notation:

Notation 6. *Whenever there exists some gadget $True(\alpha, \beta, \gamma)$ such that $g(x) = \gamma$ and $g(\gamma) = y$, we will often write $g(x) = True$ and $g(True) = y$ by commodity.*

[1] For convenience, we confuse truth values "true" and "false" with 0 and 1 and assimilate a monadic predicate $U \subseteq D$ to its characteristic function $U : D \longrightarrow \{0, 1\}$.

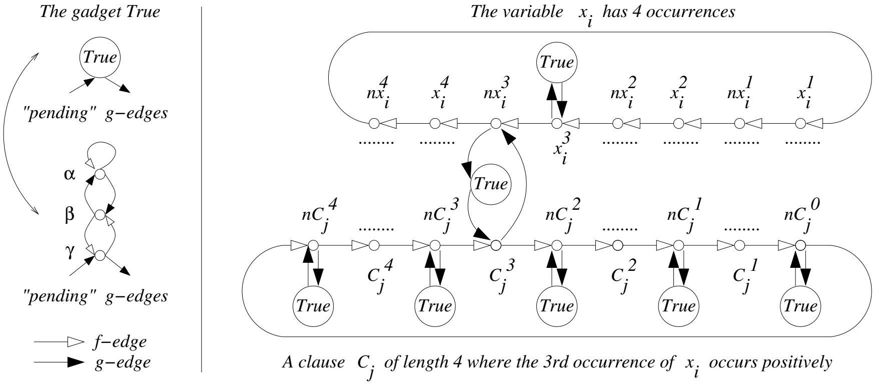

Fig. 1. The gadget True and the reduction around variable x_i and clause C_j

Let us now describe the f- and g-circuits of our permutation structure $\mathcal{S}(F)$:

• Construct a f-circuit $(x_i^1, nx_i^1, x_i^2, nx_i^2, \cdots, x_i^{r-1}, nx_i^{r-1}, x_i^r, nx_i^r)$ for each variable x_i with r occurrences in F. Vertices x_i^k, nx_i^k correspond to the k^{th} occurrence of x_i in F.

• Construct a f-circuit $(nC_j^\ell, C_j^\ell, nC_j^{\ell-1}, C_j^{\ell-1}, \cdots, nC_j^1, C_j^1, nC_j^0)$ of odd length for each clause $C_j = \lambda_1 \vee \cdots \vee \lambda_\ell$ in F, where the C_j^k and nC_j^k are new elements corresponding to the "prefix" of length k of the clause C_j defined as $\mathit{prefix}_k(C_j) = \lambda_1 \vee \cdots \vee \lambda_k$; also construct the $\ell + 1$ g-circuits (nC_j^k, True) for $0 \leq k \leq \ell$ using $\ell + 1$ new gadgets True.

• If the k^{th} literal of C_j is the h^{th} occurrence – resp. negation of the h^{th} occurrence – of x_i, construct the g-circuits $(C_j^k, nx_i^h, \text{True})$ and (x_i^h, True) – resp. $(C_j^k, x_i^h, \text{True})$ and (nx_i^h, True) – using two new gadgets True.

This completes the description of $\mathcal{S}(F)$ which is represented on the right of Fig. 1. The following lemma, that is obvious by the construction of $\mathcal{S}(F)$, means that our reduction preserves planarity.

Lemma 2. *F is a planar formula iff $\mathcal{S}(F)$ is a planar permutation structure.*

2.4 Properties of the reduction

Lemmas 3 and 4 that follow mean together that $\rho : F \mapsto \mathcal{S}(F)$ is a reduction (resp. parsimonious reduction) from SAT to the problem defined by φ_0 (resp. φ_{nand}). First, the following fact whose proof is straightforward will be useful in our study of the f-circuits of $\mathcal{S}(F)$.

Fact 1. *Let $\mathcal{S} = \langle D, f, g \rangle$ be a permutation structure and $U : D \longrightarrow \{0,1\}$ be a monadic predicate such that $(\mathcal{S}, U) \models \forall x\ \psi_0(x)$. Then, for every $a \in D$ such that $(\mathcal{S}, U) \models U(ga)$ (i.e., $U(ga) = 1$), it holds $U(a) = 1 - U(fa)$.*

Here is the first implication involved in the equivalence to be proved, i.e., $\mathcal{S}(F) \models \varphi_0$ (resp. φ_{nand}) iff F is satisfiable.

Lemma 3. *If $\mathcal{S}(F)$ satisfies φ_0 then F is satisfiable.*

In order to prove Lemma 3, we need the following two claims:

Claim 1 (Existence of a witness literal for each clause). *Let U be a predicate such that $(\mathcal{S}(F), U) \models \forall x\ \psi_0(x)$. For each clause C_j, there exists at least one literal λ in C_j for which it holds: $U(nx_i^h) = 0$ if $\lambda = x_i$, and $U(x_i^h) = 0$ if $\lambda = \neg x_i$, where λ is the h^{th} occurrence of x_i.*

Claim 2 (Coherence of the occurrences of the same variable). *Let U be a predicate such that $(\mathcal{S}(F), U) \models \forall x\ \psi_0(x)$. For each variable x_i occurring r times, it holds: $U(x_i^1) = 1 - U(nx_i^1) = U(x_i^2) = 1 - U(nx_i^2) = \cdots = U(x_i^r) = 1 - U(nx_i^r)$.*

We first prove Claims 1 and 2, and then deduce Lemma 3.

Proof (of Claim 1). Assume that the claim is false. Then there is a clause C_j such that for each literal λ, it holds $U(nx_i^h) = 1$ if $\lambda = x_i$ and $U(x_i^h) = 1$ if $\lambda = \neg x_i$. This implies $U(ga) = 1$ for each element a of the f-circuit of C_j, and hence $U(a) = 1 - U(fa)$ by Fact 1, which is impossible since the length of this f-circuit is odd. □

Proof (of Claim 2). Immediate consequence of Fact 1 applied to each element a of the f-circuit of x_i since we always have $g(a) =$ True and thus $U(ga) = 1$. □

Proof (of Lemma 3). Define the assignment I of the variables F as $I(x_i) = U(x_i^h) = 1 - U(nx_i^h)$, for each variable x_i and any $1 \leq h \leq r$, which is coherent by Claim 2. Claim 1 ensures that in each clause C_j of F, there is some literal λ such that $I(\lambda) = 1$. Hence, $I \models C_j$ and $I \models F$. □

Lemma 4 states the most precise property of our reduction $\rho : F \mapsto \mathcal{S}(F)$.

Lemma 4. *There is a bijective correspondence $I \mapsto U_I$ of the set of satisfying assignments $\{I : I \models F\}$ onto the set of monadic predicates $\{U : (\mathcal{S}(F), U) \models \forall x\ \psi_{\text{nand}}(x)\}$. That means that $\rho : F \mapsto \mathcal{S}(F)$ is a parsimonious reduction from* SAT *to the problem defined by φ_{nand}.*

Proof (of Lemma 4). For each I such that $I \models F$, let us construct its associated monadic predicate U_I, on the domain D of $\mathcal{S}(F)$. The correction will be ensured by Claim 3 and its converse Claim 4:

• Set $U_I(\alpha) = 1$, $U_I(\beta) = 0$ and $U_I(\gamma) = 1$ for each gadget True(α, β, γ) in $\mathcal{S}(F)$: this is justified by Lemma 1;

• Set $U_I(x_i^h) = I(x_i)$ and $U_I(nx_i^h) = 1 - I(x_i)$ for each variable x_i occurring r times in F and each $1 \leq h \leq r$;

• For each clause $C_j = \lambda_j^1 \vee \cdots \vee \lambda_j^\ell$, set $U_I(nC_j^0) = 1$, and for $k = 1, \cdots, \ell$, set $U_I(C_j^k) = \mathit{value}(\mathit{prefix}_k(C_j), I)$, and $U_I(nC_j^k) = 1 - \mathit{value}(\mathit{prefix}_k(C_j), I)$, where $\mathit{prefix}_k(C_j) = \lambda_j^1 \vee \cdots \vee \lambda_j^k$ and in particular $C_j = \mathit{prefix}_\ell(C_j)$.

In the following, we essentially use the well-known fact that all the Boolean connectives can be expressed by means of the NAND one only. More precisely, $1 - v = \text{NAND}(v, 1)$ and $\text{OR}(v, v') = \text{NAND}(1 - v, 1 - v')$.

Claim 3. $(\mathcal{S}(F), U_I) \models \forall x\ \psi_{\text{nand}}(x)$.

Proof (of Claim 3). For each element a of the f-circuit of any variable x_i, we have $U_I(ga) = 1$ and $U_I(a) = 1 - U_I(fa)$, and hence $(\mathcal{S}(F), U_I) \models U(a) \iff \text{NAND}(U(fa), U(ga))$. For every clause C_j of length ℓ, one easily obtains the following equalities for $1 \leq k \leq \ell$ if $C_j^k = C_j^{k-1} \vee x_i^h$:

- $U_I(nC_j^k) = 1 - U_I(C_j^k) = \text{NAND}(U_I(C_j^k), 1)$, and
- $U_I(C_j^k) = \text{NAND}(U_I(nC_j^{k-1}), U_I(nx_i^h))$;

and similarly in the case $C_j^k = C_j^{k-1} \vee \neg x_i^h$. This proves $(\mathcal{S}(F), U_I) \models \psi_{\text{nand}}(a)$ for every element $a \neq nC_j^0$ in the f-circuit of C_j. Finally, this also holds for $a = nC_j^0$ since $U_I(nC_j^\ell) = \mathit{value}(\neg C_j, I) = 0$ and, as a consequence, $U_I(nC_j^0) = 1 = \text{NAND}(U_I(nC_j^\ell), 1)$ as required. This completes the proof of Claim 3. □

It remains to prove the converse of Claim 3.

Claim 4. *Let U be a monadic predicate such that $(\mathcal{S}(F), U) \models \forall x\ \psi_{\text{nand}}(x)$. Then there is an assignment I, of course unique, such that $U = U_I$ and $I \models F$.*

Proof (of Claim 4). It is a variant of the proof of Lemma 3 and is left to the reader. This completes the proof of Lemma 4. □

Lemmas 2, 3 and 4 together imply the following:

Corollary 1. **(i)** SAT *(resp.* PLAN-SAT*) reduces to problem* MIN$_1$ *(resp.* MIN$_2$*) by the reduction $\rho : F \mapsto \mathcal{S}(F)$.* **(ii)** #SAT *(resp.* #PLAN-SAT*) parsimoniously reduces to problem* #NAND$_1$ *(resp.* #NAND$_2$*) by the same reduction.*

So, we have proved Theorems 1 and 5(i), by making use of the known result that #SAT parsimoniously reduces to #PLAN-SAT [13]. A careful analysis of our reduction $\rho : F \mapsto \mathcal{S}(F)$ from SAT (PLAN-SAT) to MIN$_1$ (MIN$_2$) shows that the only part of $\mathcal{S}(F)$ where this reduction is not parsimonious are the f-circuits of the clauses of F when at least two literals of some clause of F are true together. On the other hand, it is known that the problem $\frac{1}{3}$-SAT (also denoted one-in-three-SAT, see [6]) and its planar restriction PLAN-$\frac{1}{3}$-SAT defined below are equivalent to SAT and PLAN-SAT under parsimonious reductions (see [10]).

Definition 4. *Let $\frac{1}{3}$-*SAT *(resp.* PLAN-$\frac{1}{3}$-SAT*) denote the satisfiability problem of a conjunction of $\frac{1}{3}$-clauses (resp. planar $\frac{1}{3}$-clauses) of the form $\frac{1}{3}(a, b, c)$ whose meaning is "exactly one of the three variables a, b, c is true".*

Theorem 4 is a straightforward consequence of the following lemma:

Lemma 5. #$\frac{1}{3}$-SAT *(resp.* #PLAN-$\frac{1}{3}$-SAT*) reduces to* #MIN$_1$ *(resp.* #MIN$_2$*) under a* weakly parsimonious *reduction.*

Proof. Let $F \mapsto F'$ be the trivial parsimonious and planarity-preserving reduction from $\frac{1}{3}$-SAT (resp. PLAN-$\frac{1}{3}$-SAT) to SAT (resp. PLAN-SAT) that replaces every $\frac{1}{3}$-clause $\frac{1}{3}(a, b, c)$ by the logically equivalent conjunction $(a \vee b \vee c) \wedge$

$(\neg a \vee \neg b) \wedge (\neg b \vee \neg c) \wedge (\neg c \vee \neg a)$. One notices that in each clause of this conjunction, except one 2-clause, e.g., $C = \neg a \vee \neg b$, exactly one literal is true and both literals of C are true. Let us now consider the composed reduction $\rho' : F \mapsto \mathcal{S}(F')$ from $\frac{1}{3}$-SAT (PLAN-$\frac{1}{3}$-SAT) to MIN$_1$ (MIN$_2$). If F contains q $\frac{1}{3}$-clauses then it holds $\#\{U : (\mathcal{S}(F'), U) \models \forall x\ \psi_0(x)\} = 2^q \times \#\{I : I \models F\}$.

This is easily justified by a careful analysis of the f-circuits of clauses (of F') in $\mathcal{S}(F')$: one sees that each $\frac{1}{3}$-clause of F gives exactly 2 "local configurations" of the (union of four) f-circuits of the four corresponding clauses of F'. □

2.5 Minimality of φ_0 and δ_0 in Theorem 2

We consider EMSO formulas of the form: $\varphi : \exists \overline{U}\ \forall \overline{x}\ \psi$, where $\overline{U}$ (resp. $\overline{x}$) is a list of monadic relation symbols (resp. first-order variables) and ψ is quantifier-free.

Proof. There is nothing to prove about the absence of composition of functions and the absence of equality. We prove the minimality of:

- *the input signature* (= 2 unary function symbols)*:* a famous theorem of Courcelle [2], asserts that any MSO property of bounded tree-width structures can be checked in deterministic linear time. In particular, any EMSO property of σ-structures with $\sigma = \{f, U_1, \cdots, U_k\}$ where f is a unary function symbol and $U_1, \cdots, U_k$ are monadic relation symbols is checkable in linear time.
- *the number of EMSO symbols* (= 1)*:* immediate since any first-order (FO) property is AC$_0$ and thus is PTIME.
- *the number of FO symbols* (= 1)*:* trivial.
- *the number of clauses in* φ_0 (= 2)*:* assume φ is an ESO formula in CNF with only one clause. If some ESO symbol occurs in φ then φ defines a trivial "yes"-problem. Otherwise, φ defines a first-order property.
- *the length of* φ_0 (= 5)*:* if the length of φ in CNF is ≤ 4 then φ either: (i) contains only clauses of length ≤ 2, or (ii) contains only one clause (of length 3 or 4), or (iii) contains exactly one 3-clause and one unit clause. In case (i), φ is ESO-Krom and, as a consequence, defines a PTIME problem [8]. In case (ii), φ defines a PTIME problem as it was noticed above. Finally, in case (iii), one observes that the 3-clause either contains ≤ 1 positive literal or contains ≤ 1 negative literal. Hence, φ is either ESO-Horn or ESO-Anti-Horn, and thus defines a PTIME problem [8].
- *the number of distinct atoms* (= 3)*:* if φ in CNF contains ≤ 2 distinct atoms, then its clauses are trivially of length ≤ 2, and φ is ESO-Krom.
- *the number of anticlauses in* δ_0 (= 3)*:* notice that any formula φ in DNF that contains ≤ 2 disjuncts is equivalent to a CNF formula that consists of clauses of length ≤ 2.
- *the length of* δ_0 *in DNF* (= 6)*:* w.l.g., assume that φ in DNF is of the form $\varphi : \exists \overline{U}\ \forall \overline{x}(\psi_0 \vee \psi_1)$, where ψ_0 (resp. ψ_1) is a disjunction of anticlauses in each of which no (resp. at least one) EMSO symbol occurs. If ψ_1 contains a unit anticlause, then φ defines a trivial "yes"-problem. Moreover, if the number of anticlauses in ψ_1 is ≤ 2, then φ defines a PTIME problem. Thus, if φ defines an NP-complete problem then ψ_1 consists of at least 3 anticlauses of length ≥ 2. □

2.6 Unicity up to symmetries of φ_0 and δ_0 in Theorem 3

Let us prove the unicity of φ_0 (the proof of δ_0 is similar). Let φ be an EMSO formula in CNF that satisfies the conditions of the table of Thereom 2 and defines an NP-complete problem over functional structures $\langle D, f, g\rangle$. The list of atoms that occur in φ is Ux, Ufx, Ugx, and φ is of the form $\exists U\ \forall x\ \psi(f,g,U,x)$, where ψ is a conjunction of two clauses C_1 and C_2 with $|C_1| + |C_2| = 5$ and $|C_1| < |C_2| \leq 3$. That implies $|C_1| = 2$ and $|C_2| = 3$.

Proof. One notices that one clause consists of positive literals and the other one consists of negative literals: otherwise, φ would define a trivial "yes"-problem. That implies that φ has one of the following two forms φ_0 or φ'_0 as defined in Subsection 1.2, up to permutations of f and g and swap of U and $\neg U$:

Formulas φ_0 and φ'_0 essentially define the same problem over (planar) permutation structures $\langle D, f, g\rangle$: By replacing x by $g^{-1}x$ in the matrix of the formula φ_0, we immediately get $\langle D, f, g\rangle \models \varphi_0(f, g)$ iff $\langle D, f', g'\rangle \models \varphi'_0(f', g')$, where $f' = g^{-1}$ and $g' = fg^{-1}$. This also makes sense for planar permutation structures since $G(D, f, g)$ is planar iff $G(D, f', g')$ is planar. □

It remains to prove Theorem 5(*ii*), more precisely reformulated as follows: assume Conjecture 1 and P $\neq$ NP. Then φ_{nand} is (up to permutations of x, fx, gx and swap of U and $\neg U$) the *unique minimal* EMSO $\{f, g\}$-formula in CNF of the form $\exists U\ \forall x\ \psi(x)$ with the only atoms Ux, Ufx and Ugx that defines a problem over permutation structures to which #Sat *parsimoniously* reduces. More precisely, φ_{nand} has a minimal number of clauses (= 3), and a minimal length (= 7).

2.7 Minimality of φ_{nand} in Theorem 5(ii)

Proof. We prove the minimality of:

- *the number of clauses* (= 3)*:* clearly, any EMSO formula φ of the required form that defines an NP-complete problem (over permutation structures) with exactly two clauses has exactly one purely negative clause and one purely positive clause, and has at least one 3-clause and no unit clause[2]; so, the other one has length 2 or 3. This gives only two possible forms: our minimal formula φ_0 (and its symmetrical variants), and φ_{nae} defined as:

$$\begin{array}{lll} \varphi_{\text{nae}}: & \exists U\ \forall x\quad \psi_{\text{nae}}(x) & \text{where } \psi_{\text{nae}} \text{ is the ``not-all-equal'' formula} \\ \psi_{\text{nae}}: & (Ux \vee Ufx \vee Ugx) \wedge (\neg Ux \vee \neg Ufx \vee \neg Ugx). & \end{array}$$

One easily sees that for any function structure $\mathcal{S}$, the number $\#\{U : (\mathcal{S}, U) \models \forall x\ \psi_{\text{nae}}(x)\}$ is *even* because ψ_{nae} is invariant by inversion of U and $\neg U$. So, no reduction from Sat to the problem defined by φ_{nae} (if such a polynomial reduction exists) can be parsimonious with the standard way of counting solutions.

- *the length* (= 7)*:* it is a consequence of the fact that there should be at least three clauses of length ≥ 2 with at least one of length 3. □

[2] If φ contained a unit clause, then it would define either a trivial "yes"-problem or a trivial "no"-problem.

2.8 Unicity of $\varphi_{\mathbf{nand}}$ in Theorem 5(ii)

Proof. Clearly, any formula that meets our minimality conditions, i.e., that has three clauses and length 7, has exactly one 3-clause and two 2-clauses. Moreover:

(i) At least one clause is purely positive and at least one is purely negative;

(ii) No 2-clause subsumes the 3-clause;

(iii) Each 2-clause must disagree with the 3-clause on the sign of every literal: otherwise, if we write the 3-clause as $(\ell_1 \vee \ell_2 \vee \ell_3)$, either the 2-clause is of the form $(\ell_1 \vee \ell_2)$ and then its subsumes the 3-clause, or the 2-clause is of the form $(\overline{\ell_1} \vee \ell_2)$ and then a resolution step over ℓ_1 induces the 2-clause $(\ell_2 \vee \ell_3)$ that in turn subsumes the 3-clause. This contradicts *(ii)*;

(iv) The 2-clauses have exactly one atom in common: they clearly have at least one since there are only three atoms available. Now, if they have two, they disagree on the sign of either one literal or two literals. If we have $(\ell_1 \vee \ell_2) \wedge (\ell_1 \vee \overline{\ell_2})$, then a resolution step over ℓ_2 induces the unit clause (ℓ_1). If we have $(\ell_1 \vee \ell_2) \wedge (\overline{\ell_1} \vee \overline{\ell_2})$, then $\ell_1 \iff \overline{\ell_2}$ and the 3-clause reduces either to a 2-clause or to "true" by replacing ℓ_1 by $\overline{\ell_2}$;

(v) The 3-clause must be monotone. Otherwise, by *(i)*, the two 2-clauses must be monotone of opposite sign: Let then ε be the majoritary sign of the 3-clause. The 2-clause of sign ε cannot disagree on the sign of every literal with the 3-clause, since this latter has only one literal of sign $\overline{\varepsilon}$. This contradicts *(iii)*;

(vi) Both 2-clauses are monotone, of the same sign, opposite to the sign of the 3-clause: This is a direct consequence of *(iii)* and *(v)*.

Clearly, Remarks *(iv)*, *(v)* and *(vi)* together leave exactly ψ_{nand} and its symmetrical variants as the only candidates. □

3 Conclusion and open problems

Exhibiting "the" minimal EMSO formula that defines an NP-complete problem over functional structures is the main contribution of this paper. The "minimality" is also strengthened by the fact that this main result also holds when restricted to permutation structures or even to planar permutation structures which seem to be the simplest functional structures. A striking point is the unicity (up to symmetries) of our formula. More precisely, we have seen that all the symmetrical forms of our minimal formula essentially define only *two* distinct NP-complete problems over functional structures (see formulas φ_0 and φ_0' in Section 2.6) and only *one* such problem over permutation (resp. planar permutation) structures. This delineates a very neat frontier in logic between NP-complete problems and tractable ones. Several open problems remain:

The first one is the analogous minimality question over relational structures. The second one is Conjecture 1 and its analogue for function structures: is there a parsimonious reduction from #Sat to #Min$_0$? A difficulty in counting complexity is to define a relevant notion of reduction. Recently, Durand et al [3] have defined an interesting reduction, callled subtractive reduction, under which #P and other counting complexity classes are closed and have significant complete problems. If positively answered, the following question may be easier and more relevant than Conjecture 1: is there a subtractive reduction from #SAT to #MIN$_1$ and #MIN$_2$ (i.e., are the latter #P-complete under such reductions)?

Another interesting objective consists in looking for a necessary and sufficient decidable condition for which any EMSO formula of the form $\exists\overline{U}\ \forall\overline{x}\ \psi(\overline{U},\overline{f},\overline{x})$ and of unary signature $\overline{f}$ expresses an NP-complete problem over $\overline{f}$-structures (resp. over permutation $\overline{f}$-structures, or over planar permutation $\overline{f}$-structures.)

Finally, does the EMSO formula φ_{nae} of subsection 2.7 define a PTIME or NP-complete problem over permutation structures? Notice that φ_{nae} defines a PTIME problem over planar permutation structures since the problem NAE-SAT is PTIME for planar instances [11].

Aknowledgments: the authors thank the referees for their helpful comments that contributed to improve the presentation.

References

[1] R. Barbanchon and E. Grandjean. Local problems, planar local problems and linear time. In *Computer Science Logic*, volume 2471 of *LNCS*, pages 397–411, 2002.

[2] B. Courcelle. On the expression of graph properties in some fragments of monadic second-order logic. In N. Immermann and P. Kolaitis, editors, *Descriptive Complexity and Finite Models*, pages 33–62. American Mathematical Society, 1997.

[3] A. Durand, M. Hermann, and P. Kolaitis. Subtractive reductions and complete problems for counting complexity classes. *Theoretical Computer Science (to appear)*, 2003.

[4] T. Eiter, G. Gottlob, and Y. Gurevitch. Existential second order logic over strings. *Journal of the ACM*, 41(1):77–131, 2000.

[5] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. *Complexity and Computation*, 7:43–73, 1974.

[6] M. R. Garey and D. S. Johnson. *Computers and Intractability.* W.H. Freeman and Co., 1979.

[7] G. Gottlob, P. G. Kolaitis, and T. Schwentick. Existential second-order logic over graphs: Charting the tractability frontier. *Foundations Of Computer Science*, pages 664–674, 2000.

[8] E. Grädel. Capturing complexity classes by fragments of second order logic. *Theoretical Computer Science*, 101(1):35–57, 1991.

[9] E. Grandjean and F. Olive. Graph properties checkable in linear time in the number of vertices. *Journal of Compter System Sciences (to appear)*, 2004.

[10] H. B. Hunt III, M. V. Marathe, V. Radhakrishnan, and R. E. Stearns. The complexity of planar counting problems. *SIAM Journal on Computing*, 27(4):1142–1167, 1998.

[11] J. Kratochvíl and Z. Tuza. On the complexity of bicoloring clique hypergraphs of graphs. *Symposium On Discrete Algorithms*, pages 40–41, 2000.

[12] C. L. Lautemann and B. Weinzinger. Monadic-NLIN and quantifier-free reductions. In *Computer Science Logic*, volume 1683 of *LNCS*, pages 322–337, 1999.

[13] D. Lichtenstein. Planar formulae and their uses. *SIAM Journal on Computing*, 11(2):329–343, 1982.

[14] L. G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8(2):189–201, 1979.