

Towards a more efficient approach for the satisfiability of two-variable logic

Ting-Wei Lin*

Chia-Hsuan Lu†

Tony Tan‡

Abstract

We revisit the satisfiability problem for two-variable logic, denoted by $\text{SAT}(\text{FO}^2)$, which is known to be NEXP-complete. The upper bound is usually derived from its well known *exponential size model* property. Whether it can be determinized/randomized efficiently is still an open question.

In this paper we present a different approach by reducing it to a novel graph-theoretic problem that we call *Conditional Independent Set* (CIS). We show that CIS is NP-complete and present three simple algorithms for it: Deterministic, randomized with zero error and randomized with small one-sided error, with run time $O(1.4423^n)$, $O(1.6181^n)$ and $O(1.3661^n)$, respectively.

We then show that without the equality predicate $\text{SAT}(\text{FO}^2)$ is in fact equivalent to CIS in succinct representation. This yields the same three simple algorithms as above for $\text{SAT}(\text{FO}^2)$ without the equality predicate with run time $O(1.4423^{(2^n)})$, $O(1.6181^{(2^n)})$ and $O(1.3661^{(2^n)})$, respectively, where n is the number of predicates in the input formula. To the best of our knowledge, these are the first deterministic/randomized algorithms for an NEXP-complete decidable logic with time complexity significantly lower than $O(2^{(2^n)})$. We also identify a few lower complexity fragments of FO^2 which correspond to the tractable fragments of CIS.

For the fragment with the equality predicate, we present a linear time many-one reduction to the fragment without the equality predicate. The reduction yields *equi-satisfiable* formulas and incurs a small constant blow-up in the number of predicates.

1 Introduction

Two-variable logic (FO^2) is one of the well known fragments of first-order logic that comes with decidable satisfiability problem, henceforth, denoted by $\text{SAT}(\text{FO}^2)$. The exact complexity is NEXP-complete [35, 24, 10, 21, 7] and the upper bound is usually derived from its well known *exponential size model* (ESM) property which states that every satisfiable formula has a model with size at most exponential in the length of the input formula [10]. Thus, to decide whether a formula is satisfiable, it suffices to non-deterministically “guess” a model with exponential size and verify that it is indeed a model of the formula. This is the only known algorithm for

*Affiliation: National Taiwan University. Email: r08922089@csie.ntu.edu.tw.

†Affiliation: National Taiwan University. Email: r09922064@csie.ntu.edu.tw.

‡Affiliation: National Taiwan University. Email: tonytan@csie.ntu.edu.tw.

$\text{SAT}(\text{FO}^2)$ and it is open whether it can be efficiently determinized/randomized. Enumerating all possible structures would be naïve and practically infeasible.*

In [15] a better determinization is proposed where FO^2 formulas are encoded as Boolean formulas, which can then be fed into a SAT solver for satisfiability testing. This approach stems from the observation that checking whether an FO formula has a model with a certain fixed size can be reduced to Boolean SAT, since the number of all possible ground facts becomes fixed as well. However, adopting this technique for $\text{SAT}(\text{FO}^2)$ has some major drawbacks.

First, the upper bound provided by the ESM property is not tight. While the bound *is* tight when the equality predicate is absent, it is no longer the case when the equality predicate is present. One can easily write an FO^2 formula that is satisfiable only by models of certain sizes, say, 2^n , while the upper bound provided by ESM property is of the form $3m2^n$, for some m which depends on the input formula. In such cases the FO^2 solver still has to determine the correct size, and it does so by testing various sizes from 1 up to the maximum number possible. This can lead to poor performance especially when the formula is unsatisfiable or when the formula has only “big” models. In almost all our experiments, when the models have size more than 12, the solver in [15, 20] does not terminate (within 24 hours).

Second, the conversion to Boolean formulas produces a lot of long clauses. For example, a formula of the form $\forall x \exists y \beta$ is encoded as Boolean formula of the form $\bigwedge_{i=1}^n \bigvee_{j=1}^n \ell_{i,j}$, for some n which can be exponential. Thus, there will be exponentially many clauses, each with exponentially many literals. Since most SAT solvers find such formulas challenging, this may lead to poor performance. In fact, typical benchmarks in SAT competition contain relatively few long clauses [1].

Our contribution. In this paper we propose an entirely different approach for $\text{SAT}(\text{FO}^2)$ by reducing it to a novel graph-theoretic problem that we call *Conditional Independent Set* (CIS). Briefly, an instance of CIS is a tuple of graphs (G_0, G_1, \dots, G_m) . All of them are over the same finite set of vertices where G_0 is undirected graph and the others G_1, \dots, G_m are all directed graphs. The task is to decide if there is a non-empty independent set (not necessarily maximal) Γ in G_0 such that every vertex in Γ has an outgoing edge in G_i to another vertex in Γ , for every $1 \leq i \leq m$.

We prove that CIS is NP-complete and present three simple algorithms for it: Deterministic, randomized with zero error and randomized with small one-sided error, with run time $O(1.4423^n)$, $O(1.6181^n)$ and $O(1.3661^n)$, respectively, where n is the number of vertices in the graph. Note that since CIS is NP-complete, it is unlikely that it has a polynomial time randomized algorithm.

We then show that without the equality predicate $\text{SAT}(\text{FO}^2)$ is equivalent to CIS in succinct representation in the sense that an FO^2 formula with n unary predicates can be reduced to an instance of CIS with at most 2^n vertices. The converse is also true that an instance of CIS with n vertices can be encoded as an FO^2 formula with $\log n$ unary predicates.

By applying the algorithms for CIS, we obtain the same type of algorithms for $\text{SAT}(\text{FO}^2)$ without the equality predicate with run time $O(1.4423^{(2^n)})$, $O(1.6181^{(2^n)})$ and $O(1.3661^{(2^n)})$, respectively, where n is the number of unary predicates in the input formula. To the best of our knowledge, these are the first deterministic/randomized algorithms for an NEXP-complete

*To illustrate this point, it will take almost 60 years for a 10 GHz computer to complete a program with run time 2^{2^n} when n is only as small as 6.

decidable logic with time complexity significantly lower than $O(2^{(2^n)})$. It should be noted that even when the equality predicate is absent, $\text{SAT}(\text{FO}^2)$ is already NEXP-hard [7, 21]. Thus, it is also unlikely that it has an exponential time deterministic/randomized algorithm.

For the fragment with the equality predicate, we present a linear time many-one reduction to the fragment without the equality predicate. The reduction yields *equi-satisfiable* formulas and incurs a small constant blow-up in the number of unary predicates. Note that due to the ESM property, with the equality predicate $\text{SAT}(\text{FO}^2)$ is in NEXP. Since $\text{SAT}(\text{FO}^2)$ is already NEXP-hard when the equality predicate is absent, it is implicit that there *exists* such a polynomial time many-one reduction. However, there is no known explicit reduction so far, and as far as we know, our reduction is the first one.

Finally, we perform some initial experiments comparing the performance of our algorithms with the one in [15, 20] and with Z3 solver [2]. In general our algorithms works better than the existing ones. It is also worth noting that recasting $\text{SAT}(\text{FO}^2)$ as a graph theoretic problem not only gives us more efficient and practical algorithms, but also fresh ideas on designing interesting benchmarks. These include *random* FO^2 formulas, which are formulas obtained by first generating random graph systems (in the sense of the Erdős-Rényi model) and then constructing the corresponding FO^2 formulas.

Related works. Scott [35] was the first to prove the decidability of $\text{SAT}(\text{FO}^2)$ by reducing it to the so called Gödel class formula, though his proof only works for the fragment without the equality predicate. The decidability of the general $\text{SAT}(\text{FO}^2)$ was first proved by Mortimer [24] by showing that every satisfiable FO^2 formula has double-exponential size model. The bound was later improved to single exponential by Grädel, Kolaitis and Vardi [10], which immediately implies that $\text{SAT}(\text{FO}^2)$ is in NEXP. Matching lower bound was established by Fürer [7], based on the work of Lewis [21]. In fact, the lower bound already holds for the fragment of formulas with prefix $\forall\forall\wedge\forall\exists$ using only unary predicates and without the equality predicate.

De Neville and Pratt-Hartman [6] proposed a resolution based algorithm for $\text{SAT}(\text{FO}^2)$, but their algorithm does not come with any guaranteed time complexity. From the work of Kieronski, Otto and Pratt-Hartmann [18, 30, 29], satisfiability of extensions of FO^2 with counting quantifier and one equivalence relation remains in NEXP. However, these algorithms are all non-deterministic, and it is not immediately clear to what extent they can be implemented efficiently.

There has been considerable research effort to establish efficient (deterministic/randomized) algorithms for Boolean k -SAT [22, 28, 27, 12]. For $k = 3$, well known algorithms such as by Monien and Speckenmeyer [22], Rodosek [33], Schönig [34] and PPZ/PPSZ/biased PPSZ [28, 27, 12], just to name a few, come with run time such as $O(1.619^n)$, $O(1.476^n)$, $O((4/3)^n)$, $O(1.364^n)$ and some tiny improvement on them. As far as we know, there is no similar line of work on NEXP-complete decidable logic, or in fact, any decidable logic beyond NP.

Galperin and Wigderson [8] proposed and studied a notion of succinctly represented graphs, in which graphs are represented as Boolean circuits, instead of, as lists of their edges. Papadimitriou and Yannakakis [26] showed that in such representation many NP-complete graph-theoretic problems become NEXP-complete. This notion of succinct representation is quite clearly different from the way $\text{SAT}(\text{FO}^2)$ is a succinct representation of CIS.

Finally, we note that there have been work where various extensions of two-variable logic are reduced to graph theoretic problems. See, e.g., [17, 19, 3]. However, the techniques are different from ours and they do not imply any efficient (deterministic/randomized) algorithms

for $\text{SAT}(\text{FO}^2)$.

Organization. This paper is organized as follows. We present the formal definition of CIS and its algorithms in Section 2 where we also identify a few tractable fragments of CIS. In Section 3 we consider $\text{SAT}(\text{FO}^2)$ when the equality predicate is absent, and identify fragments of FO^2 that parallel the tractable fragments of CIS. In Section 4 we present the reduction from $\text{SAT}(\text{FO}^2)$ with the equality predicate to the fragment without the equality predicate. We present some of our initial experimental results in Section 5. Finally, we conclude in Section 6. Missing details can be found in the appendix.

2 Conditional independent set

We divide this section into four subsections. Subsection 2.1 contains the formal definition of CIS and terminology that we will use in this paper. We show that CIS is NP-complete. Then, in Subsection 2.2 and 2.3 we describe our deterministic and randomized algorithms for CIS. Finally, we present some tractable fragments of CIS in Subsection 2.4.

2.1 Definition and terminology

A *graph system* is a tuple $\mathcal{G} = (G_0, G_1, \dots, G_m)$, with $m \geq 1$, where G_0, G_1, \dots, G_m are all graphs over the same (finite) set of vertices, denoted by $V(\mathcal{G})$, but G_0 is an undirected graph and all the others G_1, \dots, G_m are directed graphs. We denote by $E_i(\mathcal{G})$ the set of edges in graph G_i , for each $0 \leq i \leq m$.

We follow the convention of writing an edge as a pair (u, v) of vertices. However, an edge $(u, v) \in E_i(\mathcal{G})$, where $1 \leq i \leq m$, is understood to be a directed edge that goes from u to v . For technical reason that will become apparent later, we assume G_0 does not contain self-loops, whereas G_1, \dots, G_m may contain self-loops. To avoid clutter, we write V and E_i , instead of $V(\mathcal{G})$ and $E_i(\mathcal{G})$, when \mathcal{G} is already clear from the context.

We call G_0 the *conflict graph* in \mathcal{G} . Two adjacent vertices in G_0 are called *conflicting* vertices. An independent set in \mathcal{G} is an independent set in G_0 , i.e., a set that does not contain two conflicting vertices.

For a set $\Gamma \subseteq V$, for $1 \leq i \leq m$, we say that a vertex $u \in \Gamma$ is a G_i -good vertex in Γ , if there is $v \in \Gamma$ such that $(u, v) \in E_i$. Note that E_i may contain self-loops, thus, if $(u, u) \in E_i$, u is G_i -good in Γ . If u is G_i -good in Γ for every $1 \leq i \leq m$, we call it a *good* vertex in Γ . Otherwise, it is called a *bad* vertex in Γ . A set Γ is a *good independent set* (GIS) in \mathcal{G} , if it is a non-empty independent set and each of its vertices is a good vertex in Γ .

Intuitively, one may view the graph G_i , where $1 \leq i \leq m$, as a kind of dependency graph depicting a condition that “a vertex may be picked only if one of its outgoing neighbours in G_i is picked.” Thus, a set Γ is a GIS, if it does not contain any two conflicting vertices and for every vertex in Γ , at least one of its outgoing neighbours in G_i is also in Γ , for every $1 \leq i \leq m$.

We define the problem *Conditional Independent Set* (CIS) as given a graph system \mathcal{G} , decide if it has a GIS. In language theoretic term, CIS is the set $\{\mathcal{G} \mid \mathcal{G} \text{ has a GIS}\}$. We show that it is NP-complete, as stated below.

Theorem 1 *CIS is NP-complete and it is already NP-hard when $m = 1$.*

2.2 Deterministic algorithm

It is pretty routine to design a deterministic algorithm for CIS with run time $\tilde{O}(2^n)$. In this subsection we will present our deterministic algorithm for CIS with significantly lower complexity, i.e., $O(\delta_0^n)$, where $\delta_0 = \sqrt[3]{3}$.

We start with a simple Procedure 1 below that on input a graph system \mathcal{G} and an independent set Y in \mathcal{G} , decides if \mathcal{G} has a GIS Γ where $\Gamma \subseteq Y$.

| Procedure 1 |
|--|
| Input: A graph system \mathcal{G} and an independent set Y in \mathcal{G} . Task: Return true if and only if \mathcal{G} has a GIS $\Gamma \subseteq Y$. 1: $\Gamma := Y$. 2: while there is a bad vertex u in Γ 3: $\Gamma := \Gamma \setminus \{u\}$. 4: return true if and only if $\Gamma \neq \emptyset$. |

Procedure 1 runs in polynomial time, since checking whether a set contains a bad vertex can be done in polynomial time. To prove correctness, let Y be an independent set. Note that if there are two GIS Γ_1 and Γ_2 in Y , their union $\Gamma_1 \cup \Gamma_2$ is also a GIS in Y . This means that if there is a GIS in Y , then there is a unique maximal GIS in Y . Hence, the **while**-loop will iterate until Γ becomes the maximal GIS, and Procedure 1 returns true. On the other hand, if there is no GIS in Y , every subset of Y contains a bad vertex. Hence, the **while**-loop will iterate until Γ becomes \emptyset , and Procedure 1 returns false.

We use Procedure 1 to design a deterministic algorithm for CIS, presented as ALGORITHM-A below.

| ALGORITHM-A |
|---|
| Input: A graph system \mathcal{G} . Task: Return true if and only if \mathcal{G} has a GIS. 1: for every maximal independent set Y in \mathcal{G} do 2: Using Procedure 1, decide if there is a GIS $\Gamma \subseteq Y$. 3: if there is such a GIS, return true. 4: return false. |

Its correctness is immediate from Procedure 1 since any GIS is contained inside some maximal independent set. A well known result of Moon and Moser [23] states that there are at most $O(3^{n/3})$ maximal independent sets in a graph of n vertices. Moreover, there are algorithms [4, 38, 37] that lists all those sets in $O(3^{n/3})$ time. Hence, ALGORITHM-A runs in $O(3^{n/3})$ time. We state this formally below.

Theorem 2 ALGORITHM-A decides CIS in $O(\delta_0^n)$ time, where n is the number of vertices and $\delta_0 = \sqrt[3]{3} \leq 1.4423$.

2.3 Randomized algorithms

Next, we present our randomized algorithms. We start with the following remark.

Remark 3 Let \mathcal{G} be a graph system and let $X \subseteq Y \subseteq V(\mathcal{G})$. Let u and v be two vertices in Y that are conflicting. Suppose there is GIS Γ such that $X \subseteq \Gamma \subseteq Y$. Then, at least one of the following holds:

$$X \subseteq \Gamma \subseteq Y \setminus \{u\} \quad \text{or} \quad X \cup \{u\} \subseteq \Gamma \subseteq Y \setminus \{v\}.$$

Indeed, if $u \in \Gamma$, $v \notin \Gamma$, implying the second case. If $u \notin \Gamma$, the first case holds trivially.

Remark 3 immediately gives us a recursive procedure, presented as LAS-VEGAS below, for checking if there is a GIS in between two sets X and Y .

| LAS-VEGAS | |
|--|--|
| <p>Input: A graph system \mathcal{G} and two sets $X \subseteq Y$ where X and Y are not necessarily independent.</p> <p>Task: Return true if and only if \mathcal{G} has a GIS in between X and Y.</p> <ol style="list-style-type: none"> 1: if X is not independent set, return false. 2: Remove all vertices in Y that are conflicting with some vertex in X. 3: Remove all bad vertices in Y (like in line 2 in Procedure 1). 4: if $X \not\subseteq Y$ or $Y = \emptyset$, return false. 5: if Y is an independent set, return true. <p>(Note that if this line is reached, Y contains at least two conflicting vertices and both are not in X.)</p> <ol style="list-style-type: none"> 6: Let u, v be two conflicting vertices in Y and (u', v') be a uniform random permutation of (u, v). 7: if LAS-VEGAS($\mathcal{G}, X, Y \setminus \{u'\}$) 8: return true. 9: else 10: return LAS-VEGAS($\mathcal{G}, X \cup \{u'\}, Y \setminus \{v'\}$). | |

Note that LAS-VEGAS is a randomized algorithm with zero error. The random bit is used only to determine which one of the two conflicting vertices to be omitted first, i.e., in Line 6. Our second algorithm for CIS, presented as ALGORITHM-B below, simply runs LAS-VEGAS with $X = \emptyset$ and $Y = V(\mathcal{G})$.

| ALGORITHM-B | |
|---|--|
| <p>Input: A graph system \mathcal{G}.</p> <p>Task: Return true if and only if \mathcal{G} has a GIS.</p> <ol style="list-style-type: none"> 1: return LAS-VEGAS($\mathcal{G}, \emptyset, V(\mathcal{G})$). | |

Its correctness and analysis is stated in Theorem 4 below.

Theorem 4 ALGORITHM-B decides CIS with zero error and runs in $O(\delta_1^n)$ time, where n is the number of vertices and $\delta_1 = (\sqrt{5} + 1)/2 \leq 1.6181$.

On positive instances, i.e., on graph system with GIS, the expected run time is $O(\delta_2^n)$, where $\delta_2 = (\sqrt{3} + 1)/2 \leq 1.3661$.

Proof. Correctness is immediate from LAS-VEGAS and Remark 3. For the time complexity, it suffices to analyze LAS-VEGAS. Let $T_{j,n}$ denote the run time of LAS-VEGAS, where $j = |Y| - |X|$ and n is the number of vertices. Each $T_{j,n}$ can be defined by the following recurrence relation:

$$\begin{aligned} T_{0,n} &= T_{1,n} = O(n^2) \\ T_{j,n} &\leq T_{j-1,n} + T_{j-2,n} + O(n^2) \quad (\text{for } j \geq 2) \end{aligned}$$

Note that $T_{j,n} = O(F_j \cdot n^2)$, where F_j is the j -th Fibonacci number. By the generating function method [11, 39], $T_{j,n} = O(\delta_1^j)$, for every $j \geq 0$. Thus, the run time of ALGORITHM-B is $T_{n,n} = O(\delta_1^n)$.

To analyze its expected run time on positive instances, let $\mu_{j,n}$ denote the worst case expected run time of LAS-VEGAS where $j = |Y| - |X|$ and n is the number of vertices. Suppose there is a GIS Γ such that $X \subseteq \Gamma \subseteq Y$. Suppose u and v are two conflicting vertices in Y . There are two cases:

- Both u and v are not in Γ , in which case, we have:

$$\mu_{j,n} \leq \mu_{j-1,n} + O(n^2)$$

- One of u or v is in Γ , in which case, we have:

$$\mu_{j,n} \leq \frac{\mu_{j-1,n}}{2} + \frac{\mu_{j-1,n} + \mu_{j-2,n}}{2} + O(n^2)$$

Thus,

$$\begin{aligned} \mu_{0,n} &= \mu_{1,n} = O(n^2) \\ \mu_{j,n} &\leq \mu_{j-1,n} + \frac{\mu_{j-2,n}}{2} + O(n^2) \quad (\text{for } j \geq 2) \end{aligned}$$

By the generating function method, we have $\mu_{n,n} = O(\delta_2^n)$. ■

ALGORITHM-B can be easily converted into another randomized algorithm (with one-sided error) that runs in $O(\delta_2^n)$ time. The technique is standard. See, e.g., [25]. For the sake of completeness, we present it formally as ALGORITHM-C below, where λ is the constant factor hidden in $O(\delta_2^n)$ in Theorem 4.

| ALGORITHM-C |
|---|
| <p>Input: A graph system \mathcal{G}.</p> <p>Task: Return true if and only if \mathcal{G} has a GIS.</p> <ol style="list-style-type: none"> 1: repeat the following n times where $n = V(\mathcal{G})$. 2: Run ALGORITHM-B on \mathcal{G} until it stops, or the run time exceeds $2\lambda\delta_2^n$. 3: if ALGORITHM-B stops, return whatever it returns. 4: return false. |

Obviously ALGORITHM-C runs in $O(\delta_2^n)$ time. By Markov's inequality [25], on positive instances, the probability that the run time of ALGORITHM-B exceeds $2\lambda\delta_2^n$ is at most $1/2$. Thus, on positive instances, the probability of error is at most $(1/2)^n$. ALGORITHM-C is always correct on negative instances. We state this formally as Theorem 5 below.

Theorem 5 ALGORITHM-C decides CIS with one-sided ε -error, where $\varepsilon \leq (1/2)^n$, and runs in $O(\delta_2^n)$ time.

For implementation purpose, instead of bounding the run time of ALGORITHM-B, we can bound the number of recursive calls of LAS-VEGAS. Let $r_{j,n}$ denote the expected number of recursive calls of LAS-VEGAS on positive instances, where $j = |Y| - |X|$ and n is the number of vertices. Each $r_{j,n}$ is defined as follows.

$$\begin{aligned} r_{0,n} &= r_{1,n} = 0 \\ r_{j,n} &\leq r_{j-1,n} + \frac{r_{j-2,n}}{2} + (3/2) \quad (\text{for } j \geq 2) \end{aligned}$$

Similar to above, we can obtain $r_{j,n} = O(\delta_2^j)$, for every $j \geq 0$. Thus, the expected number of recursive calls of LAS-VEGAS on positive instances is $O(\delta_2^n)$. Using similar reasoning as above, ALGORITHM-C can be implemented by bounding the number of recursive calls to be at most $\lambda \delta_2^n$, for some appropriate constant λ , and the probability of error is at most $(1/2)^n$. This is the bound that we use in our implementation in the experiment section.

2.4 Tractable fragments

In this subsection we will present a few tractable fragments of CIS with complexity PTIME-complete, NLOG-complete and DLOG, respectively. We start with the following terminology.

Definition 6 Let $\mathcal{G} = (G_0, G_1, \dots, G_m)$.

- We say that \mathcal{G} is *conflict-free*, if G_0 does not contain any edge, i.e., it does not contain any conflicting vertices.
- We say that \mathcal{G} is *uniquely-outgoing*, if for every $1 \leq i \leq m$, every vertex has at most one outgoing edge in G_i .

Ptime-complete fragment. We first show that CIS drops to PTIME-complete on conflict-free graph systems. Note that PTIME membership is immediate, since both ALGORITHM-A and -B run in polynomial time on conflict-free graph systems. In the case of ALGORITHM-A, there is only one maximal independent set, while in the case of ALGORITHM-B, there is no recursive call.

Hardness is obtained by log-space reduction from the reachability problem for alternating graphs, which is known to be PTIME-complete [14, Theorem 3.26]. It is as follows. Let $G = (V, E, A)$ be an alternating graph, where $A \subseteq V$ is the set of universal vertices and let $s, t \in V$. Without loss of generality, we may assume that every universal vertex has exactly 2 outgoing edges and that t is an existential vertex without outgoing edge.

We construct the following graph system $\mathcal{G} = (G_0, G_1, G_2)$. The set of vertices is $V(\mathcal{G}) = V \times \{1, \dots, n\}$, where $n = |V|$. Graph G_0 does not contain any edge. The set of edges in G_1 and G_2 are as follows.

- $((t, n), (s, 1))$ and $((t, i), (t, i + 1))$ are edges in both G_1 and G_2 , for every $1 \leq i \leq n - 1$.

- For every edge (u, v) in G , where u is an existential vertex, i.e., $u \notin A$, $((u, i), (v, i + 1))$ is an edge in G_1 , for every $1 \leq i \leq n - 1$.
- For every vertex $u \in V \setminus A$, the self loop $((u, i), (u, i))$ is an edge in G_2 , for every $1 \leq i \leq n$.
- For every universal vertex $u \in A$ with outgoing edges (u, v_1) and (u, v_2) in G , $((u, i), (v_1, i + 1))$ is an edge in G_1 and $((u, i), (v_2, i + 1))$ is an edge in G_2 , for every $1 \leq i \leq n - 1$.

It is pretty straightforward that the reduction can be done in log space. The correctness follows from Lemma 7 below.

Lemma 7 *\mathcal{G} has a GIS if and only if vertex t is reachable in G from vertex s .*

NLog-complete fragment. Next, we show that CIS drops to NLOG-complete when restricted to conflict-free graph systems $\mathcal{G} = (G_0, G_1)$, i.e., $m = 1$. Indeed in this case CIS is equivalent to checking the existence of a cycle in a directed graph, as stated in Lemma 8 below. Note that it is a folklore that checking the existence of a cycle in a directed graph is NLOG-complete.

Lemma 8 *For every conflict-free graph system $\mathcal{G} = (G_0, G_1)$, \mathcal{G} has a GIS if and only if G_1 contains a cycle.*

Proof. (if) A cycle in G_1 does not contain any bad vertex, and hence, it is a GIS. (only if) Let Γ be a GIS. By definition, every vertex in Γ has an outgoing neighbor in Γ . Since Γ is finite, this implies that there is a cycle in Γ . ■

Uniquely-outgoing graph systems. Finally, we show that restricted to uniquely outgoing graph systems with $m \geq 2$, CIS is NLOG-complete. When $m = 1$, it drops to deterministic log-space.

Let $\mathcal{G} = (G_0, G_1, \dots, G_m)$ be uniquely-outgoing. For a vertex u , let $\mathcal{R}(u)$ denote the set of vertices v such that there is a path w_1, \dots, w_k , where $u = w_1$, $v = w_k$ and each $(w_j, w_{j+1}) \in \bigcup_{i=1}^m E_i(\mathcal{G})$. Intuitively, $\mathcal{R}(u)$ is the set of vertices reachable from vertex u using edges in $\bigcup_{i=1}^m E_i(\mathcal{G})$.

We have the following characterization of the existence of GIS for uniquely-outgoing graph systems.

Lemma 9 *If a uniquely-outgoing graph system \mathcal{G} has a GIS, then there is a vertex $u \in V(\mathcal{G})$ such that $\mathcal{R}(u)$ is GIS in \mathcal{G} .*

Proof. Let Γ be a GIS in \mathcal{G} and $u \in \Gamma$. Since \mathcal{G} is uniquely outgoing, $\mathcal{R}(u) \subseteq \Gamma$. Otherwise, Γ contains a bad vertex. This also implies that $\mathcal{R}(u)$ does not contain conflicting vertices, since $\mathcal{R}(u) \subseteq \Gamma$. Therefore, $\mathcal{R}(u)$ is also a GIS. ■

We will now show that CIS is NLOG-complete, when restricted to uniquely-outgoing graph systems. For the upper bound, since NLOG is closed under complement [13, 36], we describe a non-deterministic log-space algorithm that decides whether a uniquely outgoing graph system does *not* have a GIS. It works as follows. Let $\mathcal{G} = (G_0, G_1, \dots, G_m)$ be the input. Iterate through every vertex $u \in V(\mathcal{G})$ and check if one of the following hold.

- (a) There is $1 \leq i \leq m$ and a vertex $v \in \mathcal{R}(u)$ that does not have outgoing edge in G_i .
- (b) $\mathcal{R}(u)$ contains conflicting vertices.

To check (a), guess a path from u to some vertex v that does not have an outgoing edge in G_i for some $1 \leq i \leq m$. To check (b), guess two paths from u to two vertices v_1 and v_2 and verify that v_1, v_2 are conflicting vertices. It is straightforward that this algorithm uses only logarithmic space.

For the lower bound, note that non-reachability problem for standard directed graphs can be expressed as non-reachability problem for alternating graphs where all vertices are universal vertices. In this case, our reduction (for PTIME-hardness above) will yield uniquely-outgoing graph systems. (See the fourth bullet in the construction of $\mathcal{G} = (G_0, G_1, G_2)$ in our reduction for PTIME-hardness above.) Since NLOG is closed under complement, NLOG-hardness follows immediately.

Here we remark that for uniquely-outgoing $\mathcal{G} = (G_0, G_1)$, i.e., $m = 1$, every vertex $v \in \mathcal{R}(u)$ is reachable from u by a unique path. Thus, our algorithm for checking (a) above works deterministically, since every vertex has at most one outgoing edge. To check (b), we can do the following: For every pair (v_1, v_2) of conflicting vertices, we check whether v_1 and v_2 are both reachable from u . Therefore, restricted to uniquely-outgoing graph systems with $m = 1$, CIS is decidable in logarithmic space.

We summarize formally the results in this subsection as Theorem 10 below.

Theorem 10

- CIS is PTIME-complete on conflict-free graph systems with $m \geq 2$.
- CIS is NLOG-complete on conflict-free graph systems with $m = 1$.
- CIS is NLOG-complete on uniquely-outgoing graph systems with $m \geq 2$.
- CIS is decidable in deterministic log-space on uniquely-outgoing graph systems with $m = 1$.

3 FO^2 without the equality predicate

As mentioned earlier, FO^2 denotes the fragment of relational first-order logic that uses only two variables: x and y . For convenience, we assume that only unary and binary predicates are used. In this section we focus on FO^2 without the equality predicate. We introduce some standard terminology in Subsection 3.1. In Subsection 3.2 we show how to reduce $\text{SAT}(\text{FO}^2)$ to CIS. Finally, in Subsection 3.3 we present the fragments of FO^2 that correspond exactly to the tractable fragments of CIS.

3.1 Terminology

We first recall a well known result by Scott that every FO^2 sentence can be rewritten in linear time (over an extended vocabulary) into Scott normal form [35]:

$$\Phi := \forall x \forall y \alpha(x, y) \wedge \bigwedge_{i=1}^m \forall x \exists y \beta_i(x, y) \quad (1)$$

where each $\beta_i(x, y)$ and $\alpha(x, y)$ are all quantifier free formulas. If the original sentence does not contain the equality predicate, neither do the formulas $\beta_i(x, y)$ and $\alpha(x, y)$. We also note that the transformation to Scott normal form introduces new predicates, thus, it only yields equisatisfiable formulas.

For the rest of this section, let n be the total number of predicates used in Φ . We recall a few standard terminologies. A *unary literal* (in short, 1-literal) is an atomic predicate or its negation using only variable x , and a *binary literal* (in short, 2-literal) is an atomic predicate or its negation using both variables x and y . A *literal* is either a 1- or 2-literal. Note that atom like $R(x, x)$, where R is a binary relation, is 1-literal. A 2-literal is always of the form $R(x, y)$ or $R(y, x)$ or their negations.

A *unary type* (in short, 1-type) is defined as a maximally consistent set of unary literals and a *binary type* (in short, 2-type) is a maximally consistent set of binary literals. A *type* is either a 1- or 2-type. Note that a type can be viewed as a quantifier-free formula that is the conjunction of its elements. Alternatively, it can also be viewed as partial Boolean assignment to atomic predicates. The number of 1- and 2-types are at most 2^n and 2^{4k} , respectively, where k is the number of binary predicates. We will use the symbols π and η (possibly indexed) to denote 1-type and 2-type, respectively. When viewed as formula, we write $\pi(x)$ and $\eta(x, y)$, respectively. We write $\pi(y)$ to denote the formula $\pi(x)$ with x being substituted with y .

For a structure \mathcal{A} , the *type of an element* $a \in A$ is the unique 1-type π that a satisfies in \mathcal{A} . Similarly, the type of a pair $(a, b) \in A \times A$ is the unique 2-type that (a, b) satisfies in \mathcal{A} . We say that a 1- or 2-type is *realized* in \mathcal{A} if there is an element/a pair of elements that satisfies it.

3.2 Reduction to CIS

Let Φ be as in Eq. (1). For a 1-type π , we say that π is *compatible with* $\alpha(x, y)$, if there is a 2-type η such that $\pi(x) \wedge \eta(x, y) \wedge \pi(y) \models \alpha(x, y)$. Likewise, for two 1-types π_1 and π_2 (not necessarily different), we say that (π_1, π_2) is *compatible with* α , if there is a 2-type η such that $\pi_1(x) \wedge \eta(x, y) \wedge \pi_2(y) \models \alpha(x, y)$. Otherwise, we say that (π_1, π_2) is *incompatible with* α . For some $1 \leq i \leq m$, we say that (π_1, π_2) is β_i -*compatible with* α , if there is 2-type η such that $\pi_1(x) \wedge \eta(x, y) \wedge \pi_2(y) \models \alpha(x, y) \wedge \beta_i(x, y)$.

Now, let $\mathcal{G}_\Phi = (G_0, G_1, \dots, G_m)$ be a graph system defined as follows.

- The set $V(\mathcal{G})$ of vertices is the set of 1-types that are compatible with α .
- The graph G_0 consists of the following edges: (π_1, π_2) is an edge if and only if (π_1, π_2) is incompatible with α .
- For each $1 \leq i \leq m$, the graph G_i consists of the following edges: (π_1, π_2) is an edge if and only if (π_1, π_2) is β_i -compatible with α .

Note that $\pi_1(x) \wedge \eta(x, y) \wedge \pi_2(y)$ is a full Boolean assignment to the atomic predicates in $\alpha(x, y)$. Thus, deciding whether π or (π_1, π_2) is compatible or β_i -compatible with α is essentially Boolean SAT problem. Therefore, the construction of the graph system \mathcal{G}_Φ takes exponential time in n .

Next, we prove Lemma 11 that links $\text{SAT}(\text{FO}^2)$ with CIS.

Lemma 11 *Φ is satisfiable if and only if \mathcal{G}_Φ has a GIS.*

Proof. (only if) Let $\mathcal{A} \models \Phi$. Obviously, every 1-type realized in \mathcal{A} is compatible with α . Consider the set $\Gamma := \{\pi \mid \pi \text{ is realized in } \mathcal{A}\}$, which is a subset of $V(\mathcal{G}_\Phi)$. We claim that Γ is a GIS in \mathcal{G}_Φ . First, we show that Γ is an independent set. Let $\pi_1, \pi_2 \in \Gamma$. By definition, there is $a, b \in A$ whose 1-types are π_1 and π_2 , respectively. Since $\mathcal{A} \models \forall x \forall y \alpha(x, y)$, we have:

$$\mathcal{A}, x/a, y/b \models \alpha(x, y) \quad \text{and} \quad \mathcal{A}, x/b, y/a \models \alpha(x, y)$$

This means that (π_1, π_2) are not incompatible with α . Thus, π_1 and π_2 are not adjacent in the graph G_0 .

Next, we show that for every $1 \leq i \leq m$, Γ is G_i -good. Let $\pi \in \Gamma$, and let $a \in A$ be such that its 1-type is π . Let us fix an index i where $1 \leq i \leq m$. Since $\mathcal{A} \models \forall x \exists y \beta_i(x, y)$, there is an element $b \in A$ such that $\mathcal{A}, x/a, y/b \models \beta_i(x, y)$. Moreover, $\mathcal{A}, x/a, y/b \models \alpha(x, y)$. Thus, $\mathcal{A}, x/a, y/b \models \alpha(x, y) \wedge \beta_i(x, y)$. This means that $\pi(x) \wedge \eta(x, y) \wedge \pi'(y) \models \alpha(x, y) \wedge \beta_i(x, y)$, where π' and η are the 1- and 2-type of b and (a, b) , respectively. This also implies that π' is realized in \mathcal{A} , and by definition, $\pi' \in \Gamma$. Thus, (π, π') is β_i -compatible with α , which means (π, π') is an edge in G_i . This concludes that Γ is a GIS in \mathcal{G}_Φ .

(if) Let Γ be a GIS in \mathcal{G}_Φ . We will construct a structure $\mathcal{A} \models \Phi$. First, for each $\pi \in \Gamma$, we fix a countable infinite set A_π such that for different $\pi, \pi' \in \Gamma$, the sets A_π and $A_{\pi'}$ are disjoint. The universe of \mathcal{A} is the set $A = \bigcup_{\pi \in \Gamma} A_\pi$. Next, we will define the interpretation of each relation symbol by defining the 1-types of each $a \in A$ and the 2-type of each pair $(a, b) \in A \times A$. For 1-types, we set the 1-type of each element in A_π to be π itself.

The 2-type of every pair $(a, b) \in A \times A$ is defined as follows. We first enumerate all the elements in A as a_1, a_2, \dots . The assignment of the 2-types is done by iterating the following process starting from $i = 1$ to $i \rightarrow \infty$. In the i -th iteration, we partially assign 2-types of pairs involving a_i as follows.

- Let π be the 1-type of a_i , i.e., $a_i \in A_\pi$. We pick m elements a_{i_1}, \dots, a_{i_m} such that for each $1 \leq j \leq m$:
 - $i_j > i$, and
 - each $a_{i_j} \in A_{\pi_j}$, where (π, π_j) is an edge in $E_j(\mathcal{G}_\Phi)$.

Such elements exist due to the fact that Γ is a GIS and that each A_π is an infinite set.

By definition, for every $1 \leq j \leq m$, (π, π_j) is β_j -compatible with α . Hence, there is 2-type η_j such that:

$$\pi(x) \wedge \eta_j(x, y) \wedge \pi_j(y) \models \alpha(x, y) \wedge \beta_j(x, y) \quad (2)$$

We set the 2-type of (a_i, a_{i_j}) to be η_j , for each $1 \leq j \leq m$.

- For every $j < i$, whenever the 2-type of (a_i, a_j) is not yet defined, it is defined as follows. Let $\pi' \in \Gamma$ be the 1-type of a_j . Then, we pick 2-type η such that:

$$\pi(x) \wedge \eta(x, y) \wedge \pi'(y) \models \alpha(x, y) \quad (3)$$

Such η exists, since Γ is GIS, and hence, (π, π') is compatible with α . We set the 2-type of (a_i, a_j) to be η .

As $i \rightarrow \infty$, the 2-type of every pair is well-defined.

We now show that $\mathcal{A} \models \Phi$. First, $\mathcal{A} \models \forall x \forall y \alpha(x, y)$, since for every pair $(a, b) \in A \times A$, we only assign 2-type η where either Eq. (2) or Eq. (3) hold, i.e., when $\alpha(x, y)$ is respected. Moreover, the first bullet ensures that for every element a_i , the element a_{i_j} is chosen such that $\mathcal{A}, x/a_i, y/a_{i_j} \models \alpha(x, y) \wedge \beta_j(x, y)$. Thus, $\mathcal{A} \models \forall x \exists y \beta_j(x, y)$, for every $1 \leq j \leq m$. This concludes the proof of Lemma 11. \blacksquare

Let $\delta_0, \delta_1, \delta_2$ be the constants defined in Theorems 2, 4 and 5, respectively. Combining Lemma 11 with the results in Section 2, we obtain the following.

Theorem 12 *On the fragment of FO^2 without the equality predicate, the following holds.*

- *There is a non-deterministic algorithm for $\text{SAT}(\text{FO}^2)$ that runs in $O(2^n)$ time.*
- *There is a deterministic algorithm for $\text{SAT}(\text{FO}^2)$ that runs in $\tilde{O}(\delta_0^{(2^n)})$ time.*
- *There is a randomized algorithm with zero error for $\text{SAT}(\text{FO}^2)$ that runs in $\tilde{O}(\delta_1^{(2^n)})$ time.*
- *There is a randomized algorithm with small one-sided error for $\text{SAT}(\text{FO}^2)$ that runs in $\tilde{O}(\delta_2^{(2^n)})$ time.*

Here the input formula is in Scott normal form (1) and n is the number of predicates.

It is worth comparing the run time stated in Theorem 12 with the one obtained via the ESM property. Let Φ be input formula in the form of Eq. (1), where k_1 and k_2 be the number of unary and binary predicate symbols, respectively. The ESM property gives us domain size $3m2^n$, where $n = k_1 + k_2$. This means that the Boolean formula constructed contains $k_1 \cdot 3m2^n + k_2 \cdot 9m^2 2^{2n}$ Boolean variables, and yields worst case run time $2^{(k_1 \cdot 3m2^n + k_2 \cdot 9m^2 2^{2n})}$, a much higher complexity than ours which is of the form δ^{2^n} , for various $\delta < 1.7$.

To end this section, we remark that for every graph system \mathcal{G} , one can easily construct an FO^2 formula φ (without the equality predicate and using only unary predicates) such that $\mathcal{G}_\varphi = \mathcal{G}$, and the number of unary predicates used is logarithmic in the number of vertices in \mathcal{G} . This way one can view $\text{SAT}(\text{FO}^2)$ as succinct representation of CIS where the complexity is measured in the number of bits needed for naming the vertices in \mathcal{G} . Naturally, for some graph systems, the formulas that encode them have length proportional to the number of vertices, even if they use only logarithmic number of unary predicates.

3.3 Fragments with lower complexity

In this subsection we present a few fragments of FO^2 whose satisfiability problem has complexity lower than NEXP-complete. The results in this subsection parallel those in Subsection 2.4. We start with the following definitions.

Definition 13 A formula Φ in Scott normal form (1) is called a *conflict-free/uniquely-outgoing* formula, if its graph system \mathcal{G}_Φ is conflict-free/uniquely-outgoing, respectively.

Note that we can decide in polynomial space if a formula is conflict-free or uniquely-outgoing.

EXP-complete fragment. First, we show that restricted on conflict-free formulas, $\text{SAT}(\text{FO}^2)$ drops to EXP-complete. The upper bound follows directly from Theorem 10 and the fact that the constructed graph system may have exponentially many vertices.

For The lower bound, we use the fact that alternating polynomial space Turing machines are equivalent to exponential time deterministic Turing machines [5]. Let M be an alternating 1-tape Turing machine that uses cn space, for some $c \geq 1$. Let Q be the set of its states and Δ the tape alphabet. Without loss of generality, we assume that there are exactly two transitions that can be applied on every universal state. Moreover, every configuration always leads to a halting configuration, since we can assume that M has a "counter" that counts the number of steps taken so far and M rejects immediately if the counter reaches $O(2^{cn})$.

For an input word $w = a_1 \cdots a_n \in \{0, 1\}^*$, we construct a formula of the form:

$$\Phi' := \forall x \phi_0(x) \wedge \forall x \exists y \phi_1(x, y) \wedge \forall x \exists y \phi_2(x, y)$$

such that M accepts w if and only if Φ' is satisfiable. Note that since Φ' does not have the conjunct $\forall \forall$, its graph system will not have any conflicting vertices.

The vocabulary of Φ' consists of only unary predicates $U_{b,i}$, where $b \in Q \cup \Delta$ and $1 \leq i \leq cn$. Intuitively, a configuration $b_1 \cdots b_{k-1}(q, b_k)b_{k+1} \cdots b_{cn}$ (i.e., the head is in cell k and in state q and the content of the tape is $b_1 \cdots b_{cn}$) is represented as 1-type whose positive literals are $U_{q,k}(x)$ and $U_{b_i,i}(x)$ for every $1 \leq i \leq cn$. All other literals are negative.

The formulas $\phi_0(x), \phi_1(x, y), \phi_2(x, y)$ are as follows.

- $\phi_0(x)$ states that 1-type of x represents a configuration.
- $\phi_1(x, y)$ and $\phi_2(x, y)$ state that the configuration represented by the 1-type of y is the next step of the configuration represented by the 1-type of x .

Here we define the "next" step of an accepting configuration to be the initial configuration.

Two formulas $\phi_1(x, y)$ and $\phi_2(x, y)$ are required to define the next step of a configuration with universal state. Intuitively, if w is accepted by M , then Φ has a model that represents its accepting run. On the other hand, if w is rejected by M , the run will go to a rejecting configuration, for which there is no "next" step and Φ does not have any model. The construction of $\phi_0(x), \phi_1(x, y), \phi_2(x, y)$ is routine.

PSPACE-complete fragment. Next, we show that on conflict-free formulas with $m = 1$, $\text{SAT}(\text{FO}^2)$ drops further to PSPACE-complete. Here m is as in the Scott normal form (1).

For the upper bound, note that by Lemma 8, it suffices to check the existence of a cycle in the constructed graph system, which can be done by guessing a vertex and a cycle that contains it. Each vertex is a 1-type, hence requires linear space. Verifying whether there is an edge between two vertices is essentially a Boolean SAT problem, hence, can also be done in linear space.

For the lower bound, the reduction is similar to the one for EXP-hardness, except that M is now a deterministic polynomial space Turing machine. On input word $w = a_1 \cdots a_n \in \{0, 1\}^*$, we construct a formula of the form:

$$\Phi'' := \forall x \phi_0(x) \wedge \forall x \exists y \phi_1(x, y)$$

where $\phi_0(x)$ and $\phi_1(x, y)$ are as in Φ' . Note that since M is a deterministic Turing machine, i.e., there is no universal state, we only require one formula $\phi_1(x, y)$ to express the next step of a configuration.

Uniquely-outgoing formulas. Finally, we show that on uniquely-outgoing formulas, $\text{SAT}(\text{FO}^2)$ is PSPACE-complete. The upper bound is obtained by Lemma 9, i.e., by checking if there is a vertex u such that $\mathcal{R}(u)$ is a GIS. Note that the (non-deterministic) algorithm that checks whether $\mathcal{R}(u)$ is not a GIS can be employed here. Again, since each vertex is a 1-type that requires linear space, and verifying whether there is an edge between two vertices can be done in linear space, overall the algorithm uses only polynomial space.

For the lower bound, note that the constructed Φ'' above yields a uniquely-outgoing graph system. Recall that $\phi_1(x, y)$ states that the 1-type of y represents the next step of the 1-type of x . Since M is deterministic, the 1-type of y must be unique. Thus, the constructed graph system of Φ'' is uniquely-outgoing.

We summarize formally the results in this subsection as Theorem 14 below.

Theorem 14

- $\text{SAT}(\text{FO}^2)$ is EXP-complete on conflict-free formulas with $m \geq 2$.
- $\text{SAT}(\text{FO}^2)$ is PSPACE-complete on conflict-free formulas with $m = 1$.
- $\text{SAT}(\text{FO}^2)$ is PSPACE-complete on uniquely-outgoing formulas.

To end this section, it is worth noting that two-variable guarded fragment formulas (without the equality predicate) fall into the category of conflict-free formulas. Indeed, the normal form of guarded fragment formulas with two variables are of the form:

$$\forall x \gamma(x) \wedge \bigwedge_{i=1}^n \forall x \forall y e_i(x, y) \rightarrow \alpha_i(x, y) \wedge \bigwedge_{i=1}^m \forall x \exists y f_i(x, y)$$

where $e_i(x, y), f_i(x, y)$ are atomic predicates and $\gamma(x), \alpha_i(x, y)$ are quantifier free. See, e.g., [16, 31]. Due to the guard $e_i(x, y)$, every realized 1-types π, π' are compatible. Thus, its graph system is conflict-free and our algorithms for $\text{SAT}(\text{FO}^2)$ runs in exponential time. This recovers a special case of Grädel's result [9] that the satisfiability of guarded fragment (with the equality predicate) with a fixed number of variables/arity of the predicates is EXP-complete. Combining this with our proof above, we obtain the following.

- The satisfiability problem for the formulas of the form $\varphi \wedge \psi$, where φ is a guarded fragment formula with two variables and ψ has prefix $\forall \wedge (\forall \exists)^*$ is EXP-complete.
- The satisfiability problem for the formulas with prefix $\forall \wedge \forall \exists$ is PSPACE-complete.

4 FO^2 with the equality predicate

In general FO^2 with the equality predicate is more expressive than the fragment without. For example, the formula $\forall x \forall y U(x) \wedge U(y) \rightarrow x = y$, which semantically states that the predicate U can only hold on at most one element, cannot be expressed even in full first-order logic without using the equality predicate.

In this section we will show how to reduce in linear time a formula with the equality predicate to an equi-satisfiable formula without the equality predicate. We start by observing that every

FO² formula with the equality predicate in Scott normal form can be further rewritten into the following form:

$$\Psi := \forall x \gamma(x) \quad \wedge \quad \forall x \forall y (x \neq y \rightarrow \alpha(x, y)) \quad \wedge \quad \bigwedge_{i=1}^m \forall x \exists y \beta_i(x, y) \wedge x \neq y \quad (4)$$

where each $\beta_i(x, y)$ is an atomic predicate and $\alpha(x, y)$ does not use the equality predicate.

Note that in Eq. (4) β_i is an atomic predicate, whereas in Eq. (1) β_i is a quantifier free formula. This is for technical convenience. Without loss of generality, we assume Ψ does not use constant symbols, since they can be represented with unary predicates. We also assume that there is no atom of the form $R(x, x)$ or $R(y, y)$, where R is binary predicate, since they can be treated like unary predicates.

Let $\tau = \{U_1, \dots, U_n, \beta_1, \dots, \beta_k\}$ be the vocabulary of Ψ , where $k \geq m$, and U_i and β_i are unary and binary predicates used in Ψ , respectively. We will construct a formula Ψ^* equisatisfiable to Ψ . The intuition is as follows. The formula Ψ^* is defined so that if it is satisfiable, then the universe of any of its models can be seen as built of the pairs $(a, b) \in \bigcup_{i=1}^m \beta_i^{\mathcal{A}}$, for some structure \mathcal{A} that satisfies Ψ . The main challenge is to describe the properties of those pairs without using the equality predicate.

The formula Ψ^* will be defined over vocabulary τ^* which consists of the following predicates.

- Binary predicates: β_1, \dots, β_k (as in τ).
- Unary predicates: $K_s, K_t, S_1, \dots, S_n, T_1, \dots, T_n, P_1, \dots, P_k, \overline{P}_1, \dots, \overline{P}_k, Z_1, \dots, Z_{n+\log_2(3m)}$.

For the rest of this section, we denote by $p = n + \log_2 3m$.

The terminology of *king type* and *king element* from [10] will be crucial. So we recall them formally here.

Definition 15 Suppose \mathcal{A} is a structure over vocabulary τ . We say that a 1-type π is a *king type* in \mathcal{A} , if there is only one element $a \in A$ with 1-type π . In this case, the element a is called a *king element* in \mathcal{A} .

The notion of *edge representation structure* below is our formalism on how structures over τ^* correspond to pairs of elements in the predicates $\bigcup_{i=1}^m \beta_i$ in structures over τ .

Definition 16 Let \mathcal{A} be a structure over τ . The *edge representation structure of \mathcal{A}* is a structure \mathcal{B} over vocabulary τ^* defined as follows.

- The universe B consists of all the pairs $(a, b) \in A \times A$, where (a, b) or (b, a) is in $\beta_i^{\mathcal{A}}$ for some $1 \leq i \leq m$.
- $(a, b) \in K_s^{\mathcal{B}}$ if and only if a is a king element in \mathcal{A} .
- $(a, b) \in K_t^{\mathcal{B}}$ if and only if b is a king element in \mathcal{A} .
- For each $1 \leq i \leq n$,
 - $(a, b) \in S_i^{\mathcal{B}}$ if and only if $a \in U_i^{\mathcal{A}}$, and
 - $(a, b) \in T_i^{\mathcal{B}}$ if and only if $b \in U_i^{\mathcal{A}}$.

- For each $1 \leq i \leq k$,
 - $(a, b) \in P_i^{\mathcal{B}}$ if and only if $(a, b) \in \beta_i^{\mathcal{A}}$, and
 - $(a, b) \in \overline{P}_i^{\mathcal{B}}$ if and only if $(b, a) \in \beta_i^{\mathcal{A}}$.

As mentioned earlier, since atoms such as $\beta_i(x, x)$ and $\beta_i(y, y)$ are treated as unary predicates, it is implicit that $a \neq b$, for every $(a, b) \in \beta_i^{\mathcal{A}}$, and hence, for every $(a, b) \in B$. Note also that the definition of \mathcal{B} does not contain the interpretation of β_i 's and Z_i 's. Predicates β_i 's will be used to describe the property of “non-pairs” in \mathcal{A} , whereas Z_i 's will be used to describe the property of pairs (a, b) , where a is non-king and b is king in \mathcal{A} .

We need some more terminology. For a structure \mathcal{A} over τ , the terms *source* and *target* of a pair $(a, b) \in A \times A$ refer to the elements in the first and second components in the pair (a, b) , respectively, i.e., a and b . For the rest of this section, we use \mathcal{A} to denote a structure over τ and \mathcal{B} a structure over τ^* . As usual, A and B denote the universe of \mathcal{A} and \mathcal{B} , respectively.

Let π be a 1-type over τ^* . The *S-type* of π is defined to be the intersection of π and the set $\{S_i(x), \neg S_i(x) \mid 1 \leq i \leq n\}$. In other words, *S-types* are maximally consistent sets involving only the predicates S_1, \dots, S_n . The notions of *T-type* and *Z-type* can be defined in similar manner, i.e., involving only the predicates T_1, \dots, T_n and Z_1, \dots, Z_p , respectively. The *P-type* of π is the intersection between π and the set $\{P_i(x), \overline{P}_i(x), \neg P_i(x), \neg \overline{P}_i(x) \mid 1 \leq i \leq k\}$. Note that *P-type* involves both the predicates $P_i(x)$'s and $\overline{P}_i(x)$'s.

The *S-type*, *T-type*, *Z-type* and *P-type* of an element $b \in B$ is defined as the *S-type*, *T-type*, *Z-type* and *P-type* of π , respectively, where π is the 1-type of b . Intuitively, if an element $b \in B$ represents a pair (a_1, a_2) in \mathcal{A} , the *S-type* and *T-type* of b represent the 1-types of the source and target of pair (a_1, a_2) , respectively. The *P-type* of b represents the 2-type of pair (a_1, a_2) . We will use *Z-type* to denote the “id” of the sources of pairs. Intuitively, if two elements in B have the same *Z-types*, it means that their sources have the same id, which means that they can be regarded as the same element. This will be useful to enforce certain properties on the 2-type of pairs (a_1, a_2) , where the source is not a king element, but the target is.

We define some formulas that will be useful later on.

$$\xi_S(x, y) \quad := \quad \bigwedge_{i=1}^n S_i(x) \leftrightarrow S_i(y)$$

Intuitively, it states that the *S-types* of x and y are the same. We can define $\xi_T(x, y)$, $\xi_Z(x, y)$ and $\xi_P(x, y)$ in similar manner which state that x and y have the same *T-type*, *Z-type* and *P-type*, respectively.

$$\begin{aligned} \xi_{\text{rev}}(x, y) &:= \bigwedge_{i=1}^k (P_i(x) \leftrightarrow \overline{P}_i(y)) \wedge (\overline{P}_i(x) \leftrightarrow P_i(y)) \\ \xi_{S,T}(x, y) &:= \bigwedge_{i=1}^n S_i(x) \leftrightarrow T_i(y) \end{aligned}$$

Intuitively, these two formulas state that x is the “reverse” of y and that the *S-type* of x is the “same” as the *T-type* of y , respectively.

Next, the formula $\xi_{\neq}(x, y)$ below states that the sources of x and y represent different elements.

$$\xi_{\neq}(x, y) := (K_s(x) \wedge K_s(y)) \rightarrow \neg \xi_s(x, y)$$

The intuition is as follows. Suppose $b_1, b_2 \in B$ represent the pairs (a_1, a'_1) and (a_2, a'_2) in \mathcal{A} , respectively. Recall that b_1 and b_2 belong to K_s means that the sources of b_1 and b_2 are king elements in \mathcal{A} . Thus, if a_1 and a_2 are kings and they have different types, then they are different. Later on, we will see that if at least one of a_1 or a_2 is not king, we can assume that they are different elements, even when they have the same type.

Finally, we have the formula $\xi_{\text{id}}(x, y)$ defined below.

$$\xi_{\text{id}}(x, y) := \xi_s(x, y) \wedge ((\neg K_s(y) \wedge K_t(y)) \rightarrow \xi_z(x, y))$$

Intuitively, it states that x and y have the same S -type and if the source of y is not king, but the target is a king, then x and y have the same Z -type.

Now we are ready to define Ψ^* :

$$\Psi^* := \bigwedge_{i=1}^{10} \psi_i$$

Below is the definition of each ψ_i together with their intuitive meaning.

$$\begin{aligned} \psi_1 &:= \forall x \bigvee_{i=1}^m P_i(x) \vee \overline{P}_i(x) \\ \psi_2 &:= \forall x \forall y \quad \xi_s(x, y) \rightarrow (K_s(x) \leftrightarrow K_s(y)) \\ \psi_3 &:= \forall x \forall y \quad \xi_t(x, y) \rightarrow (K_t(x) \leftrightarrow K_t(y)) \\ \psi_4 &:= \forall x \exists y \left(\begin{array}{l} \xi_{s,t}(x, y) \wedge \xi_{s,t}(y, x) \wedge \xi_{\text{rev}}(x, y) \\ \wedge (K_s(x) \leftrightarrow K_t(y)) \\ \wedge (K_t(x) \leftrightarrow K_s(y)) \end{array} \right) \\ \psi_5 &:= \forall x \quad \xi_{s,t}(x, x) \rightarrow (\neg K_s(x) \wedge \neg K_t(x)) \\ \psi_6 &:= \forall x \forall y \left(\begin{array}{l} \xi_s(x, y) \wedge \xi_t(x, y) \\ \wedge \xi_z(x, y) \wedge K_t(x) \end{array} \right) \rightarrow \xi_P(x, y) \\ \psi_7 &:= \forall x \quad K_s(x) \rightarrow \bigwedge_{i=1}^p \neg Z_i(x) \end{aligned}$$

The intention of the formulas ψ_1 – ψ_7 is to capture the natural properties of pairs in \mathcal{A} and do not depend on the original formula Ψ . Their intuitive meaning is as follows. Suppose $\mathcal{A} \models \Psi$. Formula ψ_1 states that we are only interested in pairs $(a_1, a_2) \in \beta_i^{\mathcal{A}}$, where $1 \leq i \leq m$ and it is essential since $\mathcal{A} \models \forall x \exists y \beta_i(x, y) \wedge x \neq y$, for every $1 \leq i \leq m$. Formula ψ_2 states that if the sources of two pairs have the same 1-type, then either both sources are kings or both are not kings. Formula ψ_3 states likewise regarding the targets. Formula ψ_4 states that every pair must have its “inverse.” Formula ψ_5 states that if the source and target of a pair have the same

1-type, then both are not king elements. Formula ψ_6 states that, for every two pairs, if their sources and targets have the same 1-type and the target is a king and the sources have the same id, then 2-type of the pairs are the same. Finally, formula ψ_7 states that all king elements have fixed id.

The rest of the formulas ψ_8 – ψ_{10} are defined according to the formula Ψ . Recall that Ψ is in the form of Eq. (4). Formula ψ_8 is defined as follows.

$$\psi_8 := \forall x \gamma_1(x)$$

where $\gamma_1(x)$ is the formula obtained from $\gamma(x)$ by replacing every atom $U_i(x)$ with $S_i(x)$, for every $1 \leq i \leq n$. The intention of ψ_8 is to represent $\forall x \gamma(x)$.

Formula ψ_9 is defined as follows.

$$\psi_9 := \forall x \alpha_1(x) \quad \wedge \quad \forall x \forall y \xi_{\neq}(x, y) \rightarrow \alpha_2(x, y)$$

where $\alpha_1(x)$ and $\alpha_2(x, y)$ are as follows.

- $\alpha_1(x)$ is obtained from $\alpha(x, y)$ by replacing every atom $U_i(x)$ with $S_i(x)$, every $U_i(y)$ with $T_i(x)$, every $\beta_i(x, y)$ with $P_i(x)$ and every $\beta_i(y, x)$ with $\bar{P}_i(x)$.
- $\alpha_2(x, y)$ is the formula obtained from $\alpha(x, y)$ by replacing every atom $U_i(x)$ with $S_i(x)$ and every $U_i(y)$ with $S_i(y)$. The binary literals stay the same.

The intention is to represent the part $\forall x \forall y (x \neq y \rightarrow \alpha(x, y))$.

Finally, formula ψ_{10} is defined as follows.

$$\psi_{10} := \bigwedge_{i=1}^m \forall x \exists y \xi_{\text{id}}(x, y) \wedge P_i(y)$$

It is intended to represent the part $\bigwedge_{i=1}^m \forall x \exists y \beta_i(x, y) \wedge x \neq y$.

Remark 17 It is obvious that Ψ^* can be constructed in linear time in the length of Ψ . Note that the number of unary and binary predicates in Ψ^* are $3n + 2k + 2 + \log_2 m$ and k , respectively, where n and k are the number of unary and binary predicates in Ψ .

Moreover, the conjunction of ψ_1 – ψ_9 , except ψ_4 , can be combined into one $\forall\forall$ conjunct. Thus, the constructed Ψ^* is in the form: $\forall x \forall y \alpha'(x, y) \wedge \bigwedge_{i=1}^{m+1} \forall x \exists y \beta'_i(x, y)$, which is Scott normal form. Note that the number of $\forall\exists$ conjuncts increases by 1 due to ψ_4 .

The rest of this section is devoted to the proof that Ψ and Ψ^* are indeed equi-satisfiable, stated as Lemmas 18 and 19 below.

Lemma 18 *If Ψ is satisfiable, then Ψ^* is satisfiable.*

Proof. Let $\mathcal{A} \models \Psi$. By the ESM property of FO^2 [10, Theorem 4.3], we assume that $|A| \leq 3m2^n$. Let \mathcal{B} be the edge representation structure of \mathcal{A} . We first need to define the interpretation of each predicate β_i and Z_i in \mathcal{B} . For β_i , it suffices to define the 2-types of every pair $((a_1, b_1), (a_2, b_2))$ in \mathcal{B} . There are three cases.

- Case 1: $a_1 = a_2 = a$ and a is a king element in \mathcal{A} .

In this case, both $((a, b_1), (a, b_2))$ and $((a, b_2), (a, b_1))$ are not in $\beta_i^{\mathcal{B}}$, for every $1 \leq i \leq k$. In other words, the 2-type of $((a_1, b_1), (a_2, b_2))$ contains the literals $\neg\beta_i(x, y)$ and $\neg\beta_i(y, x)$, for every $1 \leq i \leq k$.

- Case 2: $a_1 = a_2 = a$ and a is not a king element in \mathcal{A} .

In this case, let $a' \in A$ be an element such that $a' \neq a$, but a' has the same 1-type as a . We define the 2-type of $((a, b_1), (a, b_2))$ in \mathcal{B} as the 2-type of (a, a') in \mathcal{A} .

Note that since 2-type of (a, a') uniquely determine its “inverse” (a', a) , it is implicit that the 2-type of $((a, b_2), (a, b_1))$ in \mathcal{B} is the 2-type of (a', a) in \mathcal{A} .

- Case 3: $a_1 \neq a_2$.

In this case, we define the 2-type of $((a_1, b_1), (a_2, b_2))$ in \mathcal{B} as the 2-type of (a_1, a_2) in \mathcal{A} .

Now we define the interpretation of each predicate Z_i in \mathcal{B} .

- For each king element $a \in A$, we define (a, b) *not* to be in $Z_i^{\mathcal{B}}$, for every Z_i .
- For each non-king element $a \in A$, we choose a subset $Q_a \subseteq \{Z_1, \dots, Z_p\}$ such that for different $a, a' \in A$, the sets Q_a and $Q_{a'}$ are different. Note that since $|A| \leq 3m2^n$, such sets Q_a 's exist.

Then, for every $(a, b) \in B$, we define (a, b) to be in $Z_i^{\mathcal{B}}$ if and only if $Z_i \in Q_a$.

Now, we will show that $\mathcal{B} \models \Psi^*$. It is routine to verify that \mathcal{B} satisfies the formulas $\psi_1 - \psi_7$. So we will only show that $\mathcal{B} \models \psi_8 \wedge \psi_9 \wedge \psi_{10}$. We first show that for every $(a, b), (a_1, b_1), (a_2, b_2) \in B$, the following holds.

- (a) $\mathcal{B}, x/(a, b) \models \gamma_1(x)$.
- (b) $\mathcal{B}, x/(a, b) \models \alpha_1(x)$.
- (c) $\mathcal{B}, x/(a_1, b_1), y/(a_2, b_2) \models \xi_{\neq}(x, y) \rightarrow \alpha_2(x, y)$.

Recall that for $(a, b) \in B$, we have $a \neq b$. To prove (a), note that by definition, $\gamma_1(x)$ is exactly the same formula as $\gamma(x)$, except that each $U_i(x)$ is replaced by $S_i(x)$. Now since the S -type of (a, b) is exactly the 1-type of a . and that $\mathcal{A}, x/a \models \gamma(x)$, we have $\mathcal{B}, x/(a, b) \models \gamma_1(x)$. The proof of (b) is similar.

To prove (c), suppose $\mathcal{B}, x/(a_1, b_1), y/(a_2, b_2) \models \xi_{\neq}(x, y)$. There are two cases.

- Case 1: $a_1 \neq a_2$.

By construction of \mathcal{B} , the 2-type of $((a_1, b_1), (a_2, b_2))$ in \mathcal{B} is exactly the 2-type of (a_1, a_2) in \mathcal{A} . Moreover, the S -types of (a_1, b_1) and (a_2, b_2) in \mathcal{B} are exactly the 1-type of a_1 and a_2 in \mathcal{A} , respectively. Since $\alpha_2(x, y)$ is exactly the same formula as $\alpha(x, y)$, except that each $U_i(x)$ and $U_i(y)$ are replaced by $S_i(x)$ and $S_i(y)$, respectively, it follows that $\mathcal{B}, x/(a_1, b_1), y/(a_2, b_2) \models \alpha_2(x, y)$.

- Case 2: $a_1 = a_2 = a$ and it is not a king element.

By construction of \mathcal{B} , the 2-type of $((a, b_1), (a, b_2))$ in \mathcal{B} is the 2-type of (a, a') in \mathcal{A} , for some $a' \neq a$ with the same 1-type as a . With the same reasoning as above, it follows that $\mathcal{B}, x/(a, b_1), y/(a, b_2) \models \alpha_2(x, y)$.

(a)–(c) above immediately implies that $\mathcal{B} \models \psi_8 \wedge \psi_9$. To show that $\mathcal{B} \models \psi_{10}$, let $(a, b) \in B$ and $1 \leq i \leq m$. Since $\mathcal{A} \models \Psi$, there is b' such that $\mathcal{A}, x/a, y/b' \models \beta_i(x, y) \wedge x \neq y$. By the construction of \mathcal{B} , we have:

$$\mathcal{B}, x/(a, b), y/(a, b') \models \xi_s(x, y) \wedge P_i(y) \quad \text{and} \quad \mathcal{B}, x/(a, b), y/(a, b') \models \xi_z(x, y)$$

Thus, $\mathcal{B}, x/(a, b), y/(a, b') \models \xi_{\text{id}}(x, y) \wedge P_i(y)$. This completes our proof of Lemma 18. \blacksquare

Lemma 19 *If Ψ^* is satisfiable, then Ψ is satisfiable.*

Proof. Suppose $\mathcal{B} \models \Psi^*$. We say that an S -type π is *realized* in \mathcal{B} , if there is an element $b \in B$ with S -type π . Moreover, it is a *king S -type* in \mathcal{B} , if there is an element $b \in B$ with S -type π and $b \in K_s^{\mathcal{B}}$. Note that since $\mathcal{B} \models \psi_2$, if there is such an element b , every element with the same S -type as b belongs to the predicate $K_s^{\mathcal{B}}$.

In the following, since S -types and T -types correspond to 1-types over vocabulary τ , when there is no confusion, abusing the notation, we will often use the terms S -type/ T -type interchangeably with 1-type over τ . For example, suppose π is an S -type. We say that “*the 1-type of an element a in \mathcal{A} is π* ” when we mean that its 1-type contains atom $U_i(x)$ if and only if $S_i(x)$ is in π , for every $1 \leq i \leq n$. Likewise, for T -types.

We construct a structure \mathcal{A} over vocabulary τ as follows. For every realized S -type π in \mathcal{B} , we pick a set A_π such that the following holds.

- If π is a king S -type, then A_π is a singleton.
- Otherwise, A_π is an infinite countable set.

Note that A_π is well-defined, since $\mathcal{B} \models \psi_2$. Moreover, we pick two disjoint sets A_π and $A_{\pi'}$, for every two different S -types π and π' . The universe of \mathcal{A} is defined as the union of all A_π , where π ranges over the S -types realized in \mathcal{B} .

We now define the interpretation of the predicate U_i and β_i in \mathcal{A} . To define the interpretation of each U_i in \mathcal{A} , it suffices to define the 1-type of each element. For every π , the 1-type of each element in A_π is π . Note that since $\mathcal{B} \models \forall x \gamma_1(x)$, it follows immediately that $\mathcal{A} \models \forall x \gamma(x)$.

Next, we define the interpretation of each β_i in \mathcal{A} . In the following we will refer to the element in a singleton A_π as a king element, whereas elements in infinite set A_π as non-king elements. Let \mathcal{K} be the set of all king elements, and let \mathcal{P} be the set of all non-king elements. We will define the 2-type of every pair $(a_1, a_2) \in A \times A$, where $a_1 \neq a_2$ according to the three steps outlined below.

(Step 1) The goal of this step is to assign 2-types involving elements in \mathcal{K} so that $\mathcal{A}, x/a \models \bigwedge_{i=1}^m \exists y \beta_i(x, y) \wedge x \neq y$, for every element $a \in \mathcal{K}$.

Let $a \in \mathcal{K}$. Let π be the S -type where $a \in A_\pi$. Since $\mathcal{B} \models \psi_{10}$, for every $1 \leq i \leq m$, there is an element $b_i \in B$ such that $\mathcal{B}, x/a, y/b_i \models \xi_{\text{id}}(x, y)$ and $b_i \in P_i^{\mathcal{B}}$. Since $\mathcal{B} \models \psi_4$, for every such

b_i , there is $b'_i \in B$ whose S -type is the same as the T -type of b_i . Note also that if the T -type of b_i is king T -type, then since $\mathcal{B} \models \psi_3$, the S -type of b'_i is a king S -type. Let a_i be an element in \mathcal{A} whose 1-type is the S -type of b'_i .

Let $W(a)$ denote the set $\{a_1, \dots, a_m\}$. For each a_i , we define the 2-type of (a, a_i) as follows. For every $1 \leq j \leq k$, the following holds.

- If $b_i \in P_j^{\mathcal{B}}$, it contains $\beta_j(x, y)$, else it contains $\neg\beta_j(x, y)$;
- If $b_i \in \overline{P}_j^{\mathcal{B}}$, it contains $\beta_j(y, x)$, else it contains $\neg\beta_j(y, x)$.

We do this process for every element $a \in \mathcal{K}$. Note that if $a_i \in W(a)$ is also a king element, the 2-type defined on the pair (a, a_i) is well-defined. This is because $\mathcal{B} \models \psi_4 \wedge \psi_5 \wedge \psi_6 \wedge \psi_7$, thus, there is only one 2-type between a and a_i . Note also that since $\mathcal{B} \models \psi_8$, and in particular, $\mathcal{B} \models \forall x \alpha_1(x)$, we have that $\mathcal{A}, x/a, y/a_i \models x \neq y \rightarrow \alpha(x, y)$.

For convenience, we also assume the sets $W(a)$'s are defined so that for different $a, a' \in \mathcal{K}$, the sets $W(a) \cap \mathcal{P}$ and $W(a') \cap \mathcal{P}$ are disjoint. Note that this assumption can always hold, since for every non-king S -type π , we define A_π to be infinite and there are only finitely many king elements.

(Step 2) The goal of this step is to assign 2-types so that $\mathcal{A}, x/a \models \bigwedge_{i=1}^m \exists y \beta_i(x, y) \wedge x \neq y$, for every element $a \in \mathcal{P}$.

Let \mathcal{W} denote the set $\bigcup_{a \in \mathcal{K}} W(a) \cap \mathcal{P}$. We first achieve the goal for the elements from the set \mathcal{W} . Note that the set \mathcal{W} contains all the elements a such that there is exactly one $a' \in \mathcal{K}$ where the 2-type of (a, a') in \mathcal{A} is already defined in Step 1.

Let $a \in \mathcal{W}$ and let $a' \in \mathcal{K}$ be the element such that 2-type of (a, a') is already defined. Let $b \in B$ be an element such that its S -type, T -type and P -type are the 1-type of a , 1-type of a' and 2-type of (a, a') in \mathcal{A} , respectively. Since $\mathcal{B} \models \psi_{10}$, for every $1 \leq i \leq m$, there is $b_i \in B$ such that: $\mathcal{B}, x/b, y/b_i \models \xi_{\text{id}}(x, y)$ and $b_i \in P_i^{\mathcal{B}}$. Now, for each b_i , let $a_i \in A$ be an element with 1-type the same as the T -type of b_i . If a_i is not a king element, we can assume that a_i is an element in \mathcal{P} such that 2-types involving a_i is not defined yet. Recall that for non-king type π , A_π is infinite. Thus, such a_i always exists. Then, the 2-type of (a, a_i) can then be defined in similar manner as in Step 1. That it is well defined is established by similar reasoning.

For every element in $a \in \mathcal{P} \setminus \mathcal{W}$, 1-types involving a can also be defined similarly, except that we require from $b \in B$ that its S -type is the same as 1-type of a . Note that after this step, $\mathcal{A} \models \bigwedge_{i=1}^m \forall x \exists y \beta_i(x, y) \wedge x \neq y$. Moreover, since $\mathcal{B} \models \forall x \alpha_1(x)$, for every pair $(a, a') \in A \times A$, $\mathcal{A}, x/a, y/a' \models x \neq y \rightarrow \alpha(x, y)$.

(Step 3) The last step is to define the 2-type of the remaining pairs $(a, a') \in A \times A$. Let $(a, a') \in A \times A$ be a pair whose 2-type yet to be defined. Let π and π' be the 1-types of a and a' , respectively. Then, the 2-type of (a, a') is defined as $\eta(x, y)$ where:

$$\pi(x) \wedge \eta(x, y) \wedge \pi'(y) \wedge x \neq y \models \alpha(x, y)$$

Since $\mathcal{B} \models \forall x \forall y \xi_{\neq}(x, y) \rightarrow \alpha_2(x, y)$, such $\eta(x, y)$ always exists for every π and π' . This concludes our proof of Lemma 19. \blacksquare

5 Preliminary experimental results

In this section we present some of our preliminary experimental results. We implement our FO^2 solver that works as follows. The input formula is in Scott normal form (1) or (4), i.e., without or with the equality predicate, respectively. When it is in form (1), it constructs the graph system first and then runs one of the CIS algorithms. When it is in form (4), it first performs the reduction in Section 4 before proceeds as in the case of (1).

Our solver is implemented in C++ with gcc 9.3.0 (Ubuntu 9.3.0-10ubuntu2) and perform the experiments on i7-6700 CPU @ 3.40GHz with 4 Cores and 8 CPUs and 8GB memory. The OS is Ubuntu 20.04 LTS. We use Z3 [2] version 4.8.7 for solving Boolean SAT when constructing the graph systems.

Below are some snapshots of our experiments where our solvers are compared with the one developed in [15, 20] and Z3 solver. For discussion on how Z3 can be used directly for $\text{SAT}(\text{FO}^2)$, see [15]. The time taken by our solver includes the construction of the graph systems. TO stands for “time out” (24 hours), OM for “out of memory,” UN for “unknown” (when Z3 gives up analyzing and declares “unknown”), and ER for “Exceeds maximal Recursion depth.” We record – when an experiment is not performed since the result should be apparent. In all our experiments ALGORITHM-C outputs correctly. For more experiments and their detailed commentary, we refer the reader to the appendix.

Experiment A (without the equality predicate). The formula is taken from [15] where it is called *2col*. The vocabulary is $\{U_1, \dots, U_n, V, E\}$, where U_i, V are unary and E binary. Formula \mathcal{E}_n^A states that: (i) The sets U_1, \dots, U_n are pairwise disjoint, and each of them is not empty. (ii) For every $1 \leq i \leq n$, for every $j \neq i + 1$, there is no $(a, b) \in E$ such that $a \in U_i$ and $b \in U_j$. (Here $n + 1$ is defined to be 1.) (iii) The set U_1 is a subset of V . (iv) For every a , there is b such that $(a, b) \in E$ and $b \in U_1 \cup \dots \cup U_n$. (v) For every $(a, b) \in E$, $a \in V$ if and only if $b \notin V$.

\mathcal{E}_n^A is satisfiable if and only if n is even. If satisfiable, its smallest model has cardinality n .

| | n | Run time on \mathcal{E}_n^A | | | | |
|--------|-----|-------------------------------|---------|---------|---------|-------|
| | | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| sat. | 6 | 5.74s | 0.1s | 0.1s | 0.1s | UN |
| | 14 | 10h 37m 4s | 0.4s | 0.4s | 0.4s | OM |
| | 30 | OM | 2.5s | 2.5s | 2.5s | OM |
| | 200 | – | 37m 29s | 37m 25s | 37m 32s | – |
| | 500 | – | TO | TO | TO | – |
| unsat. | 3 | TO | 0.04s | 0.04s | 0.04s | 0.02s |
| | 13 | – | 0.32s | 0.32s | 0.33s | 2m 9s |
| | 31 | – | 2.55s | 2.55s | 2.54s | OM |
| | 201 | – | 36m 55s | 36m 50s | 37m 3s | – |

Experiment B (without the equality predicate). The vocabulary is $\{U_1, \dots, U_{2n}\}$, where each U_i is unary predicate. The formula \mathcal{E}_n^B is defined so that its graph system $\mathcal{G} = (G_0, G_1)$ has 2^{2n} vertices and the conflict graph G_0 is the Moon-Moser graph [23]. That is, G_0 has around

$\delta_0^{(2^{2n})}$ number of maximal independent sets, the maximum number possible. Moreover, G_1 is defined such that \mathcal{G} has only one GIS. This formula is satisfiable for every $n \geq 2$, and the smallest model has cardinality $\lfloor 2^{2n}/3 \rfloor$.

In this instance ALG-A is sensitive towards the “ordering” of the maximal independent set and the unique GIS is contained inside one of the “last” few maximal independent sets. Thus, ALG-A performs rather poorly.

| n | Run time on \mathcal{E}_n^B | | | | |
|-----|-------------------------------|-------|-----------|------------|-------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 3.84s | 0.09s | 0.09s | 0.09s | 0.02s |
| 3 | TO | TO | 0.49s | 0.49s | UN |
| 6 | – | – | 18m 44s | 18m 43s | UN |
| 7 | – | – | 4h 56m 3s | 4h 56m 50s | UN |
| 8 | – | – | TO | TO | UN |

Experiment C (with the equality predicate). The vocabulary is $\{U_1, \dots, U_n\}$. Formula \mathcal{E}_n^C states that:

- For every element, there is another element whose 1-type is the successor of the 1-type of the former.
- Every 1-type is realizable only on one element.

This formula uses equality predicate and it is satisfiable only by a model with size 2^n .

| n | Run time on \mathcal{E}_n^C | | | | |
|-----|-------------------------------|------------|------------|------------|-------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 3 | 3.82s | 0.73s | 0.73s | 0.73s | 0.03s |
| 4 | TO | 3.04s | 3.05s | 3.06s | 0.1s |
| 5 | – | 13s | 13.07s | 13.09s | UN |
| 7 | – | 8m 21s | 8m 20s | 8m 19s | UN |
| 8 | – | 2h 29m 17s | 2h 30m 49s | 2h 30m 48s | UN |

Experiment D (random formulas without the equality predicate). In this experiment we use random FO^2 formulas, which are obtained by first generating random graph systems $\mathcal{G} = (G_0, G_1)$, and then constructing the corresponding FO^2 formulas. Both G_0 and G_1 are generated independently using the Erdős-Rényi model where the probability of an edge is $1/2$.

The constructed formula is of the form: $\forall x \forall y \alpha(x, y) \wedge \forall x \exists y \beta(x, y)$, where $\alpha(x, y)$ and $\beta(x, y)$ are both in CNF. This is to avoid “explicit” listing of the edges inside $\alpha(x, y)$ and $\beta(x, y)$. This way, our solver does not get the edges in \mathcal{G} for free, since it still needs to solve Boolean SAT to obtain them. Unfortunately, the constructed formulas are pretty huge, and in some cases, take more than 20MB. Thus, we cannot display them here due to space constraints.

In the following n denotes the number of unary predicates. The number of vertices in \mathcal{G} is always 2^n .

| n | sat./unsat. | Run time on \mathcal{E}_n^D | | | | |
|-----|-------------|-------------------------------|--------|--------|--------|---------|
| | | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | unsat. | 3.51s | 0.03s | 0.03s | 0.03s | 0.01s |
| | sat. | 2.19s | 0.03s | 0.03s | 0.03s | UN |
| 5 | sat. | 38.00s | 0.36s | 0.35s | 0.35s | UN |
| | sat. | 50.54s | 0.36s | 0.36s | 0.36s | UN |
| 7 | sat. | ER | 18.38s | 18.43s | 18.38s | UN |
| | sat. | ER | 20.16s | 20.09s | 20.07s | 0.51s |
| 8 | sat. | ER | 6m 53s | 6m 53s | 6m 54s | 15m 52s |
| | sat. | ER | 7m 47s | 7m 48s | 7m 47s | 5.12s |

6 Concluding remarks

In this paper we present a novel graph-theoretic approach to SAT(FO^2), which yield more efficient algorithms and new intuitions on FO^2 . Our approach also gives us a few fragments of SAT(FO^2) that come with lower complexity, as well as fresh ideas on how to design interesting benchmarks. In the future we plan to design more benchmarks.

While experimental results seem to validate our algorithms, we also note that our current implementation is a rather naïve one, where the whole graph system is explicitly constructed before running our CIS algorithms. This construction is the main bottleneck of our current implementation, since the CIS algorithms actually run pretty fast. It should be possible that the graph system is constructed only as needed, similar to the tableaux style algorithms employed for other logics [32]. We leave this for future work.

Acknowledgement. We would like to thank Bartosz Bednarczyk, Michael Benedikt, Filip Murlak and Anthony Widjaja Lin for their constructive suggestions and feedback, as well as the anonymous reviewers for their thorough review. We also acknowledge the generous financial support from Taiwan Ministry of Science and Technology grant no. 109-2221-E-002-143-MY3 and National Taiwan University grant no. 109L891808.

References

- [1] *SAT competition*, 2020.
- [2] *Z3 solver*, 2020.
- [3] M. Benedikt, E. Kostylev, and T. Tan. Two variable logic with ultimately periodic counting. In *ICALP*, 2020.
- [4] C. Bron and J. Kerbosch. Finding all cliques of an undirected graph (algorithm 457). *Commun. ACM*, 16(9):575–576, 1973.
- [5] A. Chandra, D. Kozen, and L. Stockmeyer. Alternation. *J. ACM*, 28(1):114–133, 1981.

- [6] H. de Nivelle and I. Pratt-Hartmann. A resolution-based decision procedure for the two-variable fragment with equality. In *IJCAR*, pages 211–225, 2001.
- [7] M. Fürer. The computational complexity of the unconstrained limited domino problem (with implications for logical decision problems). In *Logic and Machines: Decision Problems and Complexity*, pages 312–319, 1983.
- [8] H. Galperin and A. Wigderson. Succinct representations of graphs. *Inf. Control.*, 56(3):183–198, 1983.
- [9] E. Grädel. On the restraining power of guards. *J. Symb. Log.*, 64(4):1719–1742, 1999.
- [10] E. Grädel, P. Kolaitis, and M. Vardi. On the decision problem for two-variable first-order logic. *Bull. Symbolic Logic*, 3(1):53–69, 03 1997.
- [11] R. Graham, D. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science, 2nd Ed.* Addison-Wesley, 1994.
- [12] T. Hansen, H. Kaplan, O. Zamir, and U. Zwick. Faster k -sat algorithms using biased-ppsz. In *STOC*, 2019.
- [13] Neil Immerman. Nondeterministic space is closed under complementation. *SIAM J. Comput.*, 17(5):935–938, 1988.
- [14] Neil Immerman. *Descriptive complexity*. Graduate texts in computer science. Springer, 1999.
- [15] S. Itzhaky, T. Kotek, N. Rinetzky, M. Sagiv, O. Tamir, H. Veith, and F. Zuleger. On the automated verification of web applications with embedded SQL. In *ICDT*, pages 16:1–16:18, 2017.
- [16] Yevgeny Kazakov. A polynomial translation from the two-variable guarded fragment with number restrictions to the guarded fragment. In *JELIA*, volume 3229 of *LNCS*, 2004.
- [17] E. Kieronski, J. Michaliszyn, I. Pratt-Hartmann, and L. Tendera. Two-variable first-order logic with equivalence closure. *SIAM J. Comput.*, 43(3):1012–1063, 2014.
- [18] E. Kieronski and M. Otto. Small substructures and decidability issues for first-order logic with two variables. *J. Symb. Log.*, 77(3):729–765, 2012.
- [19] E. Kopczynski and T. Tan. Regular graphs and the spectra of two-variable logic with counting. *SIAM J. Comput.*, 44(3):786–818, 2015.
- [20] T. Kotek. *FO²-Solver*, 2017 (accessed April 10, 2020).
- [21] H. Lewis. Complexity results for classes of quantificational formulas. *J. Comput. Syst. Sci.*, 21(3):317–353, 1980.
- [22] B. Monien and E. Speckenmeyer. Solving satisfiability in less than 2^n steps. *Discret. Appl. Math.*, 10(3):287–295, 1985.

- [23] J. Moon and L. Moser. On cliques in graphs. *Israel J. Math.*, 3:23–28, 1965.
- [24] M. Mortimer. On language with two variables. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 21:135–140, 1975.
- [25] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.
- [26] C. Papadimitriou and M. Yannakakis. A note on succinct representations of graphs. *Inf. Control.*, 71(3):181–185, 1986.
- [27] R. Paturi, P. Pudlák, M. Saks, and F. Zane. An improved exponential-time algorithm for k -sat. *J. ACM*, 52(3):337–364, 2005.
- [28] R. Paturi, P. Pudlák, and F. Zane. Satisfiability coding lemma. *Chicago J. Theor. Comput. Sci.*, 1999, 1999.
- [29] I. Pratt-Hartmann. Complexity of the two-variable fragment with counting quantifiers. *Journal of Logic, Language and Information*, 14(3):369–395, 2005.
- [30] I. Pratt-Hartmann. The two-variable fragment with counting and equivalence. *Math. Log.*, 61(6):474–515, 2015.
- [31] Ian Pratt-Hartmann. Complexity of the guarded two-variable fragment with counting quantifiers. *J. Log. Comput.*, 17(1):133–155, 2007.
- [32] John Alan Robinson and Andrei Voronkov, editors. *Handbook of Automated Reasoning (in 2 volumes)*. Elsevier and MIT Press, 2001.
- [33] R. Rodosek. A new approach on solving 3-satisfiability. In *Artificial Intelligence and Symbolic Mathematical Computation (AISMC)*, 1996.
- [34] U. Schöning. A probabilistic algorithm for k -sat and constraint satisfaction problems. In *FOCS*, 1999.
- [35] D. Scott. A decision method for validity of sentences in two variables. *The Journal of Symbolic Logic*, page 377, 1962.
- [36] R. Szelepcsényi. The method of forced enumeration for nondeterministic automata. *Acta Inf.*, 26(3):279–284, 1988.
- [37] E. Tomita, A. Tanaka, and H. Takahashi. The worst-case time complexity for generating all maximal cliques and computational experiments. *Theor. Comput. Sci.*, 363(1):28–42, 2006.
- [38] S. Tsukiyama, M. Ide, H. Ariyoshi, and I. Shirakawa. A new algorithm for generating all the maximal independent sets. *SIAM J. Comput.*, 6(3):505–517, 1977.
- [39] H. Wilf. *Generatingfunctionology*. A K Peters, Ltd., 3 edition, 2006. Available in <https://www.math.upenn.edu/~wilf/DownldGF.html>.

APPENDIX

A NP-completeness of CIS

NP membership is straightforward. The hardness is obtained by reduction from the independent set problem: Given an undirected graph $G = (V, E)$ and an integer $k \leq |V|$, decide if G has an independent set of size k .

We construct a graph system $\mathcal{G} = (G_0, G_1)$ where the set of vertices is $V(\mathcal{G}) = V \times \{0, \dots, k-1\}$. The set of edges in G_0 is as follows.

- For every $0 \leq i \leq k-1$, the set $V \times \{i\}$ forms a clique in G_0 .
- For every vertex $u \in V$, the set $\{u\} \times \{0, \dots, k-1\}$ forms a clique in G_0 .
- For every $0 \leq i \neq j \leq k-1$, for every edge $(u, v) \in E$, $((u, i), (v, j))$ is an edge in G_0 .

The set of edges in G_1 is as follows. For every $0 \leq i \leq k-1$, there are directed edges from every vertex in $V \times \{i\}$ to every vertex in $V \times \{i+1 \pmod{k}\}$.

It is routine to show that G has an independent set of size k if and only if \mathcal{G} has a GIS.

(only if) Let $I = \{v_0, \dots, v_{k-1}\}$ be an independent set of size k . It is not difficult to see that $\{(v_0, 0), \dots, (v_{k-1}, k-1)\}$ is a GIS in \mathcal{G} .

(if) Let Γ be a GIS in \mathcal{G} . For $0 \leq i \leq k-1$, let $V_i = V \times \{i\}$. By definition of the edges in G_1 , if there is $0 \leq i \leq k-1$ such that $\Gamma \cap V_i \neq \emptyset$, then $\Gamma \cap V_{i+1 \pmod{k}} \neq \emptyset$. Thus, for every $0 \leq i \leq k-1$, $\Gamma \cap V_i \neq \emptyset$. Moreover, since each V_i forms a clique in G_0 , Γ contain exactly one vertex from each V_i . Let $(v_0, 0), \dots, (v_{k-1}, k-1)$ be the vertices in Γ . By the definition of edges in G_0 , $\{v_0, \dots, v_{k-1}\}$ is an independent set in G .

B The recurrence relations used in Section 2

In this appendix, using generating function method, we present the closed form of the value a_j defined by the following recurrence relation, where c is some constant.

$$a_0 = c, \quad a_1 = c, \quad \text{and} \quad a_j = a_{j-1} + \frac{a_{j-2}}{2} + c \quad (\text{for } j \geq 2)$$

Define $p_j = a_j + 2c$. Then, we obtain the following recurrence relation.

$$p_0 = 3c, \quad p_1 = 3c, \quad \text{and} \quad p_j = p_{j-1} + \frac{p_{j-2}}{2} \quad (\text{for } j \geq 2)$$

Its characteristic equation is $x^2 = x + \frac{1}{2}$ whose solutions are $x = \frac{1 \pm \sqrt{3}}{2}$. Then, we can obtain that for each $j \geq 0$:

$$\begin{aligned} p_j &= \left(\frac{3 + \sqrt{3}}{2} \left(\frac{1 + \sqrt{3}}{2} \right)^j + \frac{3 - \sqrt{3}}{2} \left(\frac{1 - \sqrt{3}}{2} \right)^j \right) c \\ a_j &= \left(\frac{3 + \sqrt{3}}{2} \left(\frac{1 + \sqrt{3}}{2} \right)^j + \frac{3 - \sqrt{3}}{2} \left(\frac{1 - \sqrt{3}}{2} \right)^j - 2 \right) c = \mathcal{O}(c\delta_2^j) \end{aligned}$$

Similar calculation can also be performed for the following Fibonacci like relation, where c is some constant.

$$b_0 = c, \quad b_1 = c, \quad \text{and} \quad b_j = b_{j-1} + b_{j-2} + c \quad (\text{for } j \geq 2)$$

Let $q_j = b_j + c$. Then, we obtain the following recurrence relation.

$$q_0 = 2c, \quad q_1 = 2c, \quad \text{and} \quad q_j = q_{j-1} + q_{j-2} \quad (\text{for } j \geq 2)$$

Its characteristic equation is $x^2 = x + 1$ whose solutions are $x = \frac{1 \pm \sqrt{5}}{2}$. Then, we obtain that for each $j \geq 0$:

$$\begin{aligned} q_j &= \left(\frac{5 + \sqrt{5}}{5} \left(\frac{1 + \sqrt{5}}{2} \right)^j + \frac{5 - \sqrt{5}}{5} \left(\frac{1 - \sqrt{5}}{2} \right)^j \right) c \\ b_j &= \left(\frac{5 + \sqrt{5}}{5} \left(\frac{1 + \sqrt{5}}{2} \right)^j + \frac{5 - \sqrt{5}}{5} \left(\frac{1 - \sqrt{5}}{2} \right)^j - 1 \right) c = \mathcal{O}(c\delta_1^j) \end{aligned}$$

C Missing details in Subsection 2.4

C.1 Formal definition of alternating graphs and its reachability problem

We review some of the definitions from [14]. An *alternating* graph is a directed graph $G = (V, E, A)$, where $A \subseteq V$. Vertices in A and $V \setminus A$ are called *universal* and *existential* vertices, respectively. We assume that all universal vertices have at least one outgoing edge. For two vertices $s, t \in V$, we say that t is *reachable from s in G* , if there is a rooted tree T whose nodes are labeled with vertices from G such that the following holds.

- The root node is labeled with s .
- All the leaf nodes in T are labeled with t .
- For every non-leaf node x in T , the following holds.
 - If x is labeled with an existential vertex $u \in V \setminus A$, then x has only one child labeled with vertex v where $(u, v) \in E$.
 - If x is labeled with a universal vertex $u \in A$ with k outgoing edges $(u, v_1), \dots, (u, v_k)$, then x has k children z_1, \dots, z_k labeled with v_1, \dots, v_k , respectively.

The reachability problem for alternating graphs is defined as follows. On input alternating graph G and two vertices s, t , decide if t is reachable from s in G . It is not difficult to see that this problem is just a graph theoretic formulation of alternating logarithmic space Turing machines, which are equivalent with deterministic polynomial time Turing machines [5].

C.2 Proof of Lemma 7

(if) Suppose t is reachable from s in G . Let T be the witness tree and let d be the depth. We define the depth of the root node of T as 1. Obviously we can also assume that $d \leq n$.

Let $\Gamma \subseteq V(\mathcal{G})$ be the following set.

$$\Gamma := \{(t, 1), \dots, (t, n)\} \cup \{(u, i) \mid \text{there is a non-leaf node in } T \text{ with label } u \text{ and depth } i\}$$

We claim that every vertex $(u, i) \in \Gamma$ is a good vertex in Γ , and hence, Γ is a GIS in \mathcal{G} .

First, (t, n) is not a bad vertex, since by definition, $(s, 1) \in \Gamma$ and $((t, n), (s, 1))$ is an edge in both G_1 and G_2 . Similarly, all vertices $(t, 1), \dots, (t, n-1)$ are not bad vertices.

Let $(u, i) \in \Gamma$, where $u \neq t$. Thus, there is a non-leaf node x in T with label u . There are two cases.

- u is an existential vertex in G .

By construction of \mathcal{G} , there is a self-loop $((u, i), (u, i))$ in G_2 . Therefore, (u, i) is G_2 -good.

We now show that (u, i) is G_1 -good. By definition, x has only one child. Let v be the label of this child, thus, $(u, v) \in E$. If v is a leaf node, then $v = t$. Since $(t, i+1) \in \Gamma$ and $((u, i), (t, i+1))$ is an edge in G_1 , the vertex (u, i) is G_1 -good. Similarly, if v is not a leaf node, v has depth $i+1$. By definition of Γ , $((u, i), (v, i+1))$ is an edge in G_1 . Thus, (u, i) is G_1 -good.

- u is a universal vertex in G .

By definition, u has two outgoing edges in G , denoted by (u, v_1) and (u, v_2) . By the construction of \mathcal{G} , we can assume that G_1 contains the edges $((u, i), (v_1, i+1))$ and G_2 contains the edges $((u, i), (v_2, i+1))$.

Now, x has two children y_1 and y_2 , labeled with v_1 and v_2 , respectively. If y_1 is a leaf node, then $v_1 = t$. Since $(t, i+1) \in \Gamma$ and $((u, i), (t, i+1))$ is an edge in G_1 , (u, i) is G_1 -good. If y_1 is not a leaf node, then $(v_1, i+1) \in \Gamma$. Since $((u, i), (v_1, i+1))$ is an edge in G_1 , (u, i) is G_1 -good. The proof that (u, i) is G_2 -good is similar, thus, omitted.

(only if) Suppose \mathcal{G} has a GIS Γ . Note that edges in G_1 are only one directional, i.e., they can only go from (u, i) to $(v, i+1)$. The only exception is the edge $((t, n), (s, 1))$. So, for every vertex in Γ to be G_1 -good, the set Γ must contain both (t, n) and $(s, 1)$. We have the following claim that immediately implies that t is reachable from s in G .

Claim 1 *For every vertex $(u, i) \in \Gamma$, t is reachable from u in G .*

Proof. The proof is by “backward” induction on i . The base case is $i = n$. Since (u, n) is G_1 -good in Γ , vertex (u, n) must be (t, n) . The claim holds trivially, since t is reachable from t in G .

For the induction hypothesis, we assume that for every $(u, i) \in \Gamma$, t is reachable from u in G . The induction step is as follows. Let $(u, i-1) \in \Gamma$. There are two cases.

- u is an existential vertex in G .

Since $(u, i-1)$ is G_1 -good, there is some $(v, i) \in \Gamma$ such that $((u, i-1), (v, i))$ is an edge in G_1 . Applying induction hypothesis on (v, i) , t is reachable from v in G . Let T be the witness tree.

Since $((u, i-1), (v, i))$ is an edge in G_1 , (u, v) is an edge in G . Thus, t is reachable from u in G , where the witness tree T' is obtained by inserting a new root node x with label u , and the child of x is the root node of T .

- u is a universal vertex in G .

Since $(u, i-1)$ is G_1 - and G_2 -good, there is some $(v_1, i), (v_2, i) \in \Gamma$ and $((u, i-1), (v_1, i))$ and $((u, i-1), (v_2, i))$ are edges in G_1 and G_2 , respectively. Applying induction hypothesis on (v_1, i) and (v_2, i) , t is reachable from both v_1 and v_2 in G . Let T_1 and T_2 be the respective witness trees.

Since $((u, i-1), (v_1, i))$ and $((u, i-1), (v_2, i))$ are edges in G_1 and G_2 , respectively, (u, v_1) and (u, v_2) are edges in G . Now, t is reachable from u in G , where the witness tree T' is obtained by inserting a new root node x with label u , and x has two children which are the root nodes of T_1 and T_2 . ■

This completes our proof of Lemma 7.

C.3 Missing details in Subsection 3.3

Let L be the language accepted by an alternating polynomial space 1-tape Turing machine M . We assume that M uses cn space, for some $c \geq 1$. Let Q be the set of its states and its tape alphabet be $\{0, 1, \#\}$, where $\#$ denotes the blank symbol. Let q_0 , q_{acc} and q_{rej} be its initial, accepting and rejecting states, respectively.

Without loss of generality, we assume that there are exactly two transitions that can be applied on every universal state. Modifying M , if necessary, we also assume that every configuration always leads to a halting configuration, i.e., either an accepting or rejecting configuration. This does not effect the generality of our reduction, since we can assume that M has a "counter" that counts the number of steps taken so far. When the counter reaches $O(2^{cn})$, M rejects immediately.

As mentioned in the main body, on input word $w = a_1 \cdots a_n \in \{0, 1\}^*$, we construct a formula of the form, where $[cn] = \{1, \dots, cn\}$.

$$\Phi' := \forall x \phi_0(x) \wedge \forall x \exists y \phi_1(x, y) \wedge \forall x \exists y \phi_2(x, y)$$

where the vocabulary consists of only unary predicates $U_{b,i}$. Here $b \in Q \cup \{0, 1, \#\}$ and $1 \leq i \leq cn$.

The formulas $\phi_0(x), \phi_1(x, y), \phi_2(x, y)$ are defined as follows.

- $\phi_0(x)$ states that 1-type of x represents a configuration. Formally,

$$\begin{aligned} & \left(\bigvee_{(q,i) \in Q \times [cn]} \left(U_{q,i} \wedge \bigwedge_{(p,j) \neq (q,i) \text{ and } (p,j) \in Q \times [cn]} \neg U_{p,j}(x) \right) \right) \\ & \wedge \bigwedge_{i \in [cn]} \bigvee_{b \in \{0,1,\#\}} \left(U_{b,i}(x) \wedge \bigwedge_{b' \neq b \text{ and } b' \in \{0,1,\#\}} \neg U_{b',i}(x) \right) \end{aligned}$$

- $\phi_1(x, y)$ and $\phi_2(x, y)$ state that the configuration represented by the 1-type of y is the next step of the configuration represented by the 1-type of x .

Here $\phi_1(x, y)$ handles the next step for the existential configurations. Formally,

$$\bigwedge_{(q,i) \in Q \times [cn] \text{ where } q \text{ is existential and } b \in \{0,1,\#\}} \left((U_{q,i}(x) \wedge U_{b,i}(x)) \rightarrow \varphi_{\text{next}(q,b,i)}(y) \right)$$

where $\varphi_{\text{next}(q,b,i)}(y)$ denotes that certain unary predicates must hold on y according to the transitions in M when the state is in q and the head is reading symbol b .

Here we define that if $q = q_{acc}$, then $\varphi_{\text{next}(q,b,i)}(y)$ states that 1-type of y must represent the initial configuration.

To handle universal states, we add the following conjunct in $\phi_1(x, y)$:

$$\bigwedge_{(q,i) \in Q \times [cn] \text{ where } q \text{ is universal and } b \in \{0,1,\#\}} \left((U_{q,i}(x) \wedge U_{b,i}(x)) \rightarrow \varphi_{\text{next-1}(q,b,i)}(y) \right)$$

and define $\phi_2(x, y)$ as:

$$\bigwedge_{(q,i) \in Q \times [cn] \text{ where } q \text{ is universal and } b \in \{0,1,\#\}} \left((U_{q,i}(x) \wedge U_{b,i}(x)) \rightarrow \varphi_{\text{next-2}(q,b,i)}(y) \right)$$

Here $\varphi_{\text{next-1}(q,b,i)}(y)$ and $\varphi_{\text{next-2}(q,b,i)}(y)$ denotes the “next” step when applying the first and second transitions when applied on the universal state q .

It is routine to verify that w is accepted by M if and only if Φ has a model.

D Some preliminary experimental results

Terminology. We use capital letters U , V and W (possibly indexed) to denote unary predicates and E (possibly indexed) to denote binary predicates. Unary predicates are often viewed as sets, so we use set-theoretic terminologies when referring to them. Binary predicates are viewed as directed edges, so we use graph-theoretic terms such as “outgoing” and “incoming” edges when referring to them.

Experiment 1 (without the equality predicate). This experiment is taken from [15] where it is called *path-unsat*. The vocabulary is $\{U_1, \dots, U_n, V, E\}$ and formula \mathcal{E}_n^1 states the following.

- The sets U_1, \dots, U_n are pairwise disjoint, and each of them is not empty.
- For every $1 \leq i \leq n$, for every $j \neq i + 1$, there is no edge from any element in U_i to any element in U_j .
(Here we define $n + 1$ to be 1.)
- The set U_1 is subset of V and the set U_n is subset of the complement of V .
- For every edge $(a, b) \in E$, if a belongs to V , so does b .
- Every element must have an outgoing edge to one of the elements in the sets $U_1 \cup \dots \cup U_n$.

Formally,

$$\begin{aligned} \mathcal{E}_n^1 := & \forall x \left(\bigwedge_{1 \leq i \neq j \leq n} U_i(x) \rightarrow \neg U_j(x) \right) \wedge \bigwedge_{i=1}^n \exists x U_i(x) \\ & \wedge \bigwedge_{1 \leq i \leq n} \bigwedge_{j \neq i+1} \forall x \forall y \left((U_i(x) \wedge U_j(y)) \rightarrow \neg E(x, y) \right) \\ & \wedge \forall x (U_1(x) \rightarrow V(x)) \wedge \forall x (U_n(x) \rightarrow \neg V(x)) \wedge \forall x \forall y (E(x, y) \rightarrow (V(x) \rightarrow V(y))) \\ & \wedge \forall x \exists y (E(x, y) \wedge (U_1(y) \vee \dots \vee U_n(y))) \end{aligned}$$

This formula is unsatisfiable for every $n \geq 1$. Its graph system has around $2n$ vertices and there are no conflicting vertices. The majority of run time is spent on building the graph system which takes time exponential in n . Once the graph system has been built, since there are no conflicting vertices, all our algorithms run in time polynomial in n . This explains why the performance of all our algorithms is pretty good and rather similar.

| n | Run time on \mathcal{E}_n^1 | | | | |
|-----|-------------------------------|---------|---------|---------|--------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 3 | 1m 41s | 0.04s | 0.04s | 0.04s | 0.01s |
| 6 | TO | 0.09s | 0.09s | 0.09s | 0.35s |
| 16 | – | 0.49s | 0.49s | 0.49s | 5m 39s |
| 30 | – | 2.32s | 2.30s | 2.31s | OM |
| 100 | ER | 1m 32s | 1m 32s | 1m 33s | OM |
| 150 | – | 21m 2s | 20m 58s | 21m 2s | – |
| 200 | – | 39m 16s | 39m 15s | 39m 13s | – |

Experiment 2 (without the equality predicate). This experiment is taken from [15] where it is called *2col*, and it is the same as Experiment A in the main body. The vocabulary is $\{U_1, \dots, U_n, V, E\}$ and formula \mathcal{E}_n^2 states the following.

- The sets U_1, \dots, U_n are pairwise disjoint, and each of them is not empty.
- For every $1 \leq i \leq n$, for every $j \neq i + 1$, there is no edge from any element in U_i to any element in U_j .
(Here we define $n + 1$ to be 1.)
- The set U_1 is a subset of V .
- Every element must have an outgoing edge to one of the elements in the sets $U_1 \cup \dots \cup U_n$.
- For every edge $(a, b) \in E$, a belongs to V if and only if b does not belong to V .

Formally,

$$\begin{aligned} \mathcal{E}_n^2 := & \forall x \left(\bigwedge_{1 \leq i \neq j \leq n} U_i(x) \rightarrow \neg U_j(x) \right) \wedge \bigwedge_{1 \leq i \leq n} \exists x U_i(x) \\ & \wedge \bigwedge_{1 \leq i \leq n} \bigwedge_{j \neq i+1} \forall x \forall y \left((U_i(x) \wedge U_j(y)) \rightarrow \neg E(x, y) \right) \\ & \wedge \forall x (U_1(x) \rightarrow V(x)) \wedge \forall x \forall y (E(x, y) \rightarrow (V(x) \rightarrow V(y))) \\ & \wedge \forall x \forall y (E(x, y) \rightarrow (V(x) \rightarrow \neg V(y)) \wedge (\neg V(x) \rightarrow V(y))) \end{aligned}$$

It is satisfiable if and only if n is even, and the smallest model has cardinality n . Its graph system has around $2n$ vertices with no conflicting vertices. Hence, our algorithms performs rather similarly in this instance as in Experiment 1.

| | n | Run time on \mathcal{E}_n^2 | | | | |
|--------|-----|-------------------------------|------------|-----------|------------|-------|
| | | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| sat. | 6 | 5.74s | 0.1s | 0.1s | 0.1s | UN |
| | 12 | 19s | 0.3s | 0.3s | 0.3s | OM |
| | 14 | 10h 37m 4s | 0.4s | 0.4s | 0.4s | OM |
| | 30 | OM | 2.5s | 2.5s | 2.5s | OM |
| | 100 | — | 1m 46s | 1m 46s | 1m 46s | — |
| | 150 | — | 21m 35s | 21m 33s | 21m 37s | — |
| | 200 | — | 37m 29s | 37m 25s | 37m 32s | — |
| | 300 | — | 9h 24m 32s | 9h 24m 1s | 9h 33m 18s | — |
| | 500 | — | TO | TO | TO | — |
| unsat. | 3 | TO | 0.04s | 0.04s | 0.04s | 0.02s |
| | 5 | — | 0.09s | 0.07s | 0.07s | 0.08s |
| | 13 | — | 0.32s | 0.32s | 0.33s | 2m 9s |
| | 31 | — | 2.55s | 2.55s | 2.54s | OM |
| | 101 | — | 1m 45s | 1m 45s | 1m 45s | — |
| | 151 | — | 22m 1s | 22m 2s | 22m 3s | — |
| | 201 | — | 36m 55s | 36m 50s | 37m 3s | — |

Experiment 3 (without the equality predicate). This experiment is taken from [15] where it is called *exponential*. The vocabulary is $\{U_1, \dots, U_n\}$ and formula \mathcal{E}_n^3 states the following.

- The set U_1 is not empty.
- For every $1 \leq i \leq n$, for every element a , there is an element b whose 1-type differs exactly on the atom $U_i(x)$.

Formally,

$$\mathcal{E}_n^3 := \exists x U_1(x) \wedge \bigwedge_{i=1}^n \forall x \exists y \left((U_i(x) \leftrightarrow \neg U_i(y)) \wedge \bigwedge_{j \neq i} U_j(x) \leftrightarrow U_j(y) \right)$$

It is satisfiable for every $n \geq 1$ and the smallest model has cardinality 2^n . Its graph system has 2^n vertices and there are no conflicting vertices. However, all vertices are all good vertices. Thus, once the graph system has been built, our algorithms run pretty fast.

| n | Run time on \mathcal{E}_n^3 | | | | |
|-----|-------------------------------|-----------|------------|-----------|----|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 3 | 3.12s | 0.12s | 0.11s | 0.11s | OM |
| 4 | 1m 29s | 0.31s | 0.27s | 0.27s | OM |
| 5 | TO | 0.75s | 0.75s | 0.75s | OM |
| 8 | – | 36.6s | 36.3s | 36.7s | OM |
| 9 | – | 2m 43s | 2m 43s | 2m 42s | OM |
| 12 | – | 6h 43m 5s | 6h 44m 20s | 6h 45m 3s | OM |
| 13 | – | TO | TO | TO | OM |

Experiment 4 (without the equality predicate). The vocabulary is $\{U_1, \dots, U_m, V_1, \dots, V_n\}$. In this example, we denote by U -type the maximal consistent set of atoms $U_i(x)$ or their negations. V -types can be defined similarly. We view U -type π as a string $b_1 \dots b_m \in \{0, 1\}^m$, where $b_i = 1$ if and only if $U_i(x) \in \pi$. Obviously, it can also be viewed as a number between 0 and $2^m - 1$.

Formula $\mathcal{E}_{m,n}^4$ states the following

- For every two elements, if they have the same U -types, then their V -types must be the same.
- Every element a has an outgoing edge to another element b such that U -type of b is the “successor” of U -type a .

Here “successor” is the standard successor operation where U -types are viewed as numbers and the successor of 1^m is 0^m . This formula is satisfiable for every $m, n \geq 1$ and the smallest model has cardinality 2^m .

Formally,

$$\mathcal{E}_{m,n}^4 := \forall x \forall y \left(\left(\bigwedge_{i=1}^m U_i(x) \leftrightarrow U_i(y) \right) \rightarrow \left(\bigwedge_{i=1}^n V_i(x) \leftrightarrow V_i(y) \right) \right) \wedge \forall x \exists y \text{SUC}_U(x, y)$$

where $\text{SUC}_U(x, y)$ denotes that the U -type of y is the successor of the U -type of x , defined as follows.

$$\begin{aligned} \text{SUC}_U(x, y) := & \left(\bigwedge_{i=1}^m U_i(x) \rightarrow \bigwedge_{i=1}^m \neg U_i(y) \right) \\ & \wedge \left(\neg \bigwedge_{i=1}^m U_i(x) \rightarrow \bigvee_{i=1}^m \left(\neg U_i(x) \wedge U_i(y) \wedge \bigwedge_{j=i+1}^m (U_j(x) \leftrightarrow U_j(y)) \right. \right. \\ & \left. \left. \wedge \bigwedge_{j=1}^{i-1} (U_j(x) \wedge \neg U_j(y)) \right) \right) \end{aligned}$$

It is satisfiable for every $m, n \geq 1$. Its graph system has 2^{m+n} vertices, which are partitioned into 2^m sets, and each set forms a clique in the conflict graph. Every GIS must contain exactly one vertex from each set. In this case, every maximal independent set is in fact a GIS.

| m | n | Run time on $\mathcal{E}_{m,n}^4$ | | | | |
|-----|-----|-----------------------------------|-----------|------------|------------|----|
| | | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 3 | 3.34s | 0.21s | 0.21s | 0.21s | UN |
| 2 | 5 | 3.23s | 1.74s | 1.74s | 1.75s | UN |
| 3 | 3 | 3.46s | 0.52s | 0.52s | 0.52s | UN |
| 4 | 1 | TO | 0.19s | 0.19s | 0.19s | UN |
| 4 | 3 | — | 1.41s | 1.40s | 1.41s | UN |
| 5 | 2 | — | 1.34s | 1.35s | 1.34s | UN |
| 8 | 5 | — | 1h 8m 50s | 1h 10m 15s | 1h 10m 6s | UN |
| 9 | 5 | — | 5h 9m 22s | 5h 8m 49s | 5h 10m 12s | UN |

Experiment 5 (without the equality predicate). The vocabulary is $\{U_1, \dots, U_m, V_1, \dots, V_n, W\}$. In this experiment we write $(u, v, w) \in \{0, 1\}^m \times \{0, 1\}^n \times \{0, 1\}$ to represent 1-types, where bits in u , v and w correspond to the atoms $U_i(x)$'s, $V_i(x)$'s and $W(x)$, respectively.

Formula $\mathcal{E}_{m,n}^5$ states that there are no two elements whose 1-types differ only on the atom $W(x)$, and for every element with 1-type (u, v, w) , there is an element with 1-type $(0^m, v', w')$ where v' is as follows.

- If $u = 0^m$, then $w' = 0$ and v' is the successor of v .
- Otherwise, $w' = 1$ and v' is arbitrary.

Formally,

$$\begin{aligned} \mathcal{E}_{m,n}^5 := & \forall x \forall y \left(\left(W(x) \leftrightarrow \neg W(y) \right) \rightarrow \neg \left(\bigwedge_{i=0}^m U_i(x) \leftrightarrow U_i(y) \wedge \bigwedge_{i=0}^n V_i(x) \leftrightarrow V_i(y) \right) \right) \\ & \wedge \forall x \exists y \left(\left(\bigwedge_{i=1}^m \neg U_i(y) \right) \wedge \left(\left(\bigwedge_{i=1}^m \neg U_i(x) \right) \rightarrow \neg W(y) \wedge \text{SUC}_V(x, y) \right) \right. \\ & \quad \left. \wedge \left(\left(\neg \bigwedge_{i=1}^m \neg U_i(x) \right) \rightarrow W(y) \right) \right) \end{aligned}$$

Formula $\text{SUC}_V(x, y)$ denotes that the V -type of y is the successor of the V -type of x , defined similarly as in Experiment 4.

This formula is satisfiable for every $m, n \geq 1$ and the smallest model has cardinality 2^n . Its graph system has 2^{m+n+1} vertices. The conflict graph contains exactly 2^{m+n} disjoint edges. So there are $2^{(2^{m+n})}$ maximal independent sets, but there is only one GIS, which contains 2^m vertices.

| m | n | Run time on $\mathcal{E}_{m,n}^5$ | | | | |
|-----|-----|-----------------------------------|------------|------------|------------|----|
| | | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 1 | 2 | 3s | 0.14s | 0.09s | 0.09s | UN |
| 3 | 2 | 3.24s | 0.49s | 0.49s | 0.49s | UN |
| 5 | 2 | 3.24s | 4.48s | 4.45s | 4.49s | UN |
| 1 | 4 | TO | 27.81s | 0.52s | 0.52s | UN |
| 3 | 4 | — | 36.28s | 4.62s | 4.64s | UN |
| 5 | 4 | — | 1m 54s | 1m 1s | 1m 1s | UN |
| 8 | 4 | — | 1h 16m 57s | 1h 12m 41s | 1h 12m 13s | UN |
| 9 | 4 | — | 5h 13m 3s | 5h 3m 19s | 5h 1m 46s | UN |
| 1 | 5 | — | TO | 1.56s | 1.58s | UN |
| 2 | 5 | — | TO | 4.94s | 4.95s | UN |

We should note that in this instance ALG-A is very sensitive towards the “ordering” of the maximal independent set. In our case here, the unique GIS is contained inside one of the “first” few maximal independent sets, so ALG-A manages to detect it early. It should be possible to rename the unary predicates so that the GIS is contained in one of the “last” maximal independent sets, in which case ALG-A will perform poorly. On the other hand, the performance of ALG-B and ALG-C is not effected by such ordering.

Experiment 6 (without the equality predicate). This is the same as Experiment B in the main body. The vocabulary is $\{U_1, \dots, U_{2n}\}$. Formula \mathcal{E}_n^6 is of the form:

$$\mathcal{E}_n^6 := \forall x \forall y \alpha(x, y) \wedge \forall x \exists y \beta(x, y)$$

Before we describe its detailed definition, which is a bit technical, we will first describe intuitively its graph system \mathcal{G} . It has 2^{2n} vertices, where each vertex correspond to 1-type. The conflict graph G_0 is the Moon-Moser graph [23], i.e., it has around $\delta_0^{(2^{2n})}$ number of maximal independent sets, the maximum number possible. Graph G_1 is defined so that \mathcal{G} has only one GIS. To avoid early detection of bad vertices, we make sure that there is no bad vertex in $V(\mathcal{G})$.

To present the exact description of the graph system, we write a 1-type π as a binary string $b_1 \cdots b_{2n} \in \{0, 1\}^{2n}$, where for each $1 \leq i \leq n$, $b_i = 1$ if and only if $U_i(x) \in \pi$. By the i -th bit of a 1-type π , we mean bit b_i . Abusing the notation, we will say that “the i -th bit of x ” to refer to the i -th bit of 1-type of x .

The graph system $\mathcal{G} = (G_0, G_1)$ is as follows. The set of vertices is $\{0, 1\}^{2n}$. The edges in graph G_0 is defined as follows. $(b_1 \cdots b_{2n}, c_1 \cdots c_{2n})$ is an edge if and only if either one of (a) or (b) below holds.

- (a) $b_3 \cdots b_{2n} = c_3 \cdots c_{2n} = 1^{2n-2}$ and $b_1 b_2 \neq c_1 c_2$.
- (b) There is an odd k where $3 \leq k \leq 2n - 1$ and the following holds.

- $b_k b_{k+1} \neq 11$ and $c_k c_{k+1} \neq 11$.
- $b_k b_{k+1} \neq c_k c_{k+1}$.
- $b_{k+2} \cdots b_{2n} = c_{k+2} \cdots c_{2n} = 1^{2n-k-1}$.
- $b_1 \cdots b_{k-1} = c_1 \cdots c_{k-1}$.

The edges in graph G_1 is defined as follows. $(b_1 \cdots b_{2n}, c_1 \cdots c_{2n})$ is an edge if and only if either one of (a) or (b) below holds.

- (a) $b_3 \cdots b_{2n} = 1^{2n-2}$ and $c_1 c_2 c_3 \cdots c_{2n} = 0^{2n}$.
- (b) There is an odd k where $3 \leq k \leq 2n - 1$ and $b_k b_{k+1} \neq 11$ and $b_{k+2} \cdots b_{2n} = 1^{2n-k-1}$ and the following holds.

- If $b_1 \cdots b_{k-1} = 1^{k-1}$, the following holds.
 - $c_1 \cdots c_{k-3} = 0^{k-3}$.
 - $c_{k-2} c_{k-1} \neq 11$.
 - $c_k \cdots c_{2n} = 1^{2n-k+1}$.
 - $b_k b_{k+1} = 00$ if and only if $c_{k-2} c_{k-1} = 00$.
- If $b_1 \cdots b_{k-1} \neq 1^{k-1}$, the following holds.
 - $c_1 \cdots c_{k-1}$ is the successor of $b_1 \cdots b_{k-1}$.
 - $c_k c_{k+1} \neq 11$.
 - $c_{k+2} \cdots c_{2n} = 1^{2n-k-1}$.
 - $b_k b_{k+1} = 00$ if and only if $c_k c_{k+1} = 00$.

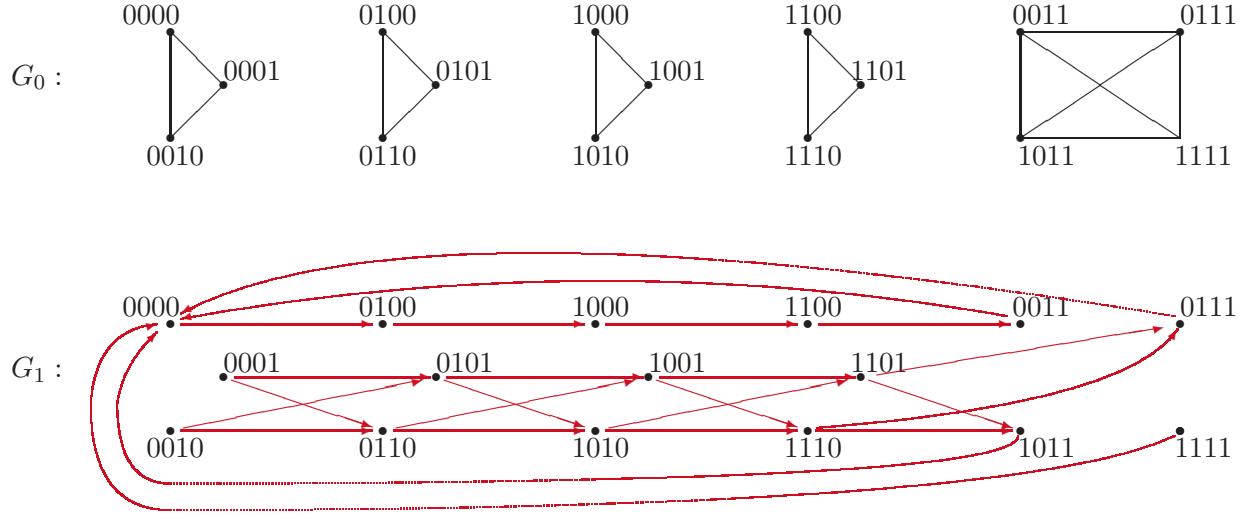


Figure 1: The graph system $\mathcal{G} = (G_0, G_1)$ of formula \mathcal{E}_n^6 with $n = 2$. The number of vertices is $2^4 = 16$. The upper graph is the conflict graph G_0 which is the Moon-Moser graph of $2^4 = 16$ vertices. The number of maximal independent sets is $3^4 \times 4 = 324$. The lower graph is graph G_1 . Note that it has only one GIS $\{0000, 0100, 1000, 1100, 0011\}$

See Figure 1 for an example when $n = 2$.

To present the formal description of our formulas, we will use the following auxiliary formulas, where $1 \leq k \leq l \leq 2n$.

$$\begin{aligned}
\text{ZERO}_{k,l}(x) &:= \bigwedge_{i=k}^l \neg U_i(x) & \text{ONE}_{k,l}(x) &:= \bigwedge_{i=k}^l U_i(x) \\
\text{EQ}_{k,l}(x, y) &:= \bigwedge_{i=k}^l U_i(x) \leftrightarrow U_i(y) & \text{EQ}'_{k,l}(x, y) &:= \bigwedge_{i=k}^l U_i(x) \leftrightarrow U_{i-2}(y) \\
\text{SUC}_k(x, y) &:= \bigvee_{i=1}^k \left(\text{EQ}_{1,(i-1)}(x, y) \wedge \neg U_i(x) \wedge U_i(y) \wedge \text{ONE}_{i+1,k}(x) \wedge \text{ZERO}_{i+1,k}(y) \right) \\
\psi_k(x, y) &:= \left(\text{EQ}_{1,k-1}(x, y) \wedge \neg \text{ONE}_{k,k+1}(x) \wedge \neg \text{ONE}_{k,k+1}(y) \right. \\
&\quad \left. \wedge \text{ONE}_{k+2,2n}(x) \wedge \text{ONE}_{k+2,2n}(y) \right) \rightarrow \text{EQ}_{k,k+1}(x, y) \\
\xi_k(x, y) &:= \left(\neg \text{ONE}_{k,k+1}(x) \wedge \text{ONE}_{k+2,2n}(x) \right) \\
&\rightarrow \left(\text{ONE}_{1,k-1}(x) \rightarrow \left(\text{ZERO}_{1,k-3}(y) \wedge \text{EQ}'_{k,k+1}(x, y) \wedge \text{ONE}_{k,2n}(y) \right) \right. \\
&\quad \left. \wedge \neg \text{ONE}_{1,k-1}(x) \rightarrow \left(\text{SUC}_k(x, y) \wedge \text{EQ}_{k,k+1}(x, y) \wedge \text{ONE}_{k+2,2n}(y) \right) \right)
\end{aligned}$$

The intuitive meaning of the first four formulas are pretty obvious. Note that in $\text{EQ}'_{k,l}(x, y)$ the

bits for y are shifted below by 2. The intuitive meaning of the last three are as follows.

- $\text{SUC}_k(x, y)$ states that the first k bits of 1-type of y is the successor of the first k bits of 1-type of x .

Note that it can only hold when the first k bits of x contain some 0.

- $\psi_k(x, y)$ is the formulation of Condition (b) for edges in G_0 for a specific k .
- $\xi_k(x, y)$ is the formulation of Condition (b) for edges in G_1 for a specific k .

Now, formulas $\alpha(x, y)$ and $\beta(x, y)$ are defined as follows.

$$\begin{aligned}\alpha(x, y) &:= \left(\left(\text{ONE}_{3,2n}(x) \wedge \text{ONE}_{3,2n}(y) \right) \rightarrow \text{EQ}_{1,2}(x, y) \right) \wedge \left(\bigwedge_{3 \leq k \leq 2n \text{ and } k \text{ is odd}} \psi_k(x, y) \right) \\ \beta(x, y) &:= \left(\text{ONE}_{3,2n}(x) \rightarrow \text{ZERO}_{1,2n}(y) \right) \wedge \left(\bigwedge_{3 \leq k \leq 2n \text{ and } k \text{ is odd}} \xi_k(x, y) \right)\end{aligned}$$

This formula is satisfiable for every $n \geq 2$. The smallest model has cardinality $\lfloor 2^{2n}/3 \rfloor$. Again, in this instance ALG-A is sensitive towards the “ordering” of the maximal independent set. In this case the unique GIS is contained inside one of the “last” few maximal independent sets, hence, ALG-A performs rather poorly.

| n | Run time on \mathcal{E}_n^6 | | | | |
|-----|-------------------------------|-------|-----------|------------|-------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 3.84s | 0.09s | 0.09s | 0.09s | 0.02s |
| 3 | TO | TO | 0.49s | 0.49s | UN |
| 4 | – | – | 4.40s | 4.37s | UN |
| 5 | – | – | 1m 2s | 1m 2s | UN |
| 6 | – | – | 18m 44s | 18m 43s | UN |
| 7 | – | – | 4h 56m 3s | 4h 56m 50s | UN |
| 8 | – | – | TO | TO | UN |

Experiment 8 (without the equality predicate). The formula \mathcal{E}_n^8 in this experiment is similar to \mathcal{E}_n^6 . The only difference is that the graph system now has many GIS, though still significantly less than the number of maximal independent sets. This formula is satisfiable for every $n \geq 2$ and the smallest model has cardinality $\lfloor 2^{2n}/3 \rfloor$. Again, the performance of our algorithms in this instance is similar to the one in Experiment 6.

Formally, we change condition (b) in \mathcal{E}_n^6 into the following.

(b') There is an odd k where $3 \leq k \leq n-1$ and $b_k b_{k+1} \neq 11$ and $b_{k+2} \cdots b_{2n} = 1^{2n-k-1}$ and the following holds.

- If $b_1 \cdots b_{k-1} = 1^{k-1}$, the following holds.
 - $c_1 \cdots c_{k-3} = 0^{k-3}$.
 - $c_{k-2} c_{k-1} \neq 11$.
 - $c_k \cdots c_{2n} = 1^{2n-k+1}$.
 - If $b_k b_{k+1} = 00$, then $c_{k-2} c_{k-1} = 00$ or $c_{k-2} c_{k-1} = 01$.
 - If $b_k b_{k+1} \neq 00$, then $c_{k-2} c_{k-1} \neq 00$.
- If $b_1 \cdots b_{k-1} \neq 1^{k-1}$, the following holds.
 - $c_1 \cdots c_{k-1}$ is the successor of $b_1 \cdots b_{k-1}$.
 - $c_k c_{k+1} \neq 11$.
 - $c_{k+2} \cdots c_{2n} = 1^{2n-k-1}$.
 - If $b_k b_{k+1} = 00$, then $c_k c_{k+1} = 00$ or $c_k c_{k+1} = 01$.
 - If $b_k b_{k+1} \neq 00$, then $c_k c_{k+1} \neq 00$.

Figure 3 below shows graph G_1 with $n = 2$. Formula \mathcal{E}_n^8 is satisfiable. The graph system has many GIS, though still significantly less than the number of maximal independent sets.

| n | Run time on \mathcal{E}_n^8 | | | | |
|-----|-------------------------------|-------|-----------|-----------|-------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 3.64s | 0.09s | 0.09s | 0.09s | 0.03s |
| 3 | TO | TO | 0.49s | 0.49s | UN |
| 4 | – | – | 4.39s | 4.39s | UN |
| 5 | – | – | 1m 3s | 1m 3s | UN |
| 6 | – | – | 17m 0s | 17m 1s | UN |
| 7 | – | – | 5h 3m 25s | 5h 2m 22s | UN |
| 8 | – | – | TO | TO | UN |

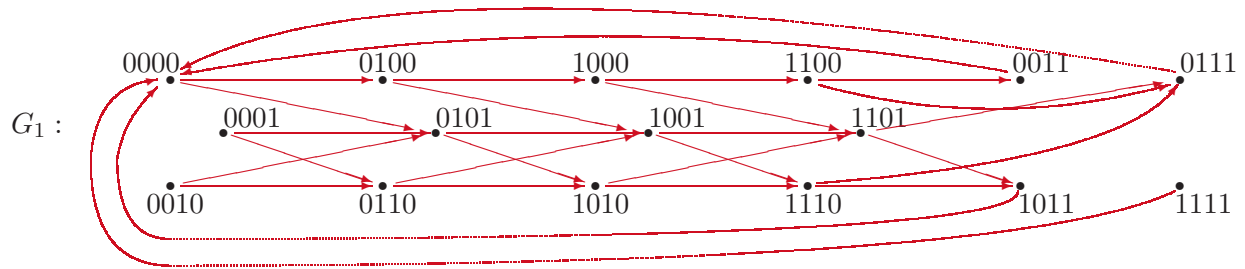


Figure 3: Graph G_1 for formula \mathcal{E}_n^8 with $n = 2$. The only difference with the one in Figure 1 is that it has additional edges $(0000, 0101)$, $(0100, 1001)$, $(1000, 1101)$ and $(1100, 0111)$. Note that the graph system has many GIS, though still significantly less than the number of maximal independent sets.

Experiment 9 (with the equality predicate). This is the same as Experiment C in the main body. The vocabulary is $\{U_1, \dots, U_n\}$. Formula \mathcal{E}_n^9 states the following.

- For every element, there is another element whose 1-type is the successor of the 1-type of the former.
- Every 1-type is realizable only on one element.

Formally,

$$\mathcal{E}_n^9 := \forall x \forall y \left(x \neq y \rightarrow \neg \bigwedge_{i=1}^n U_i(x) \leftrightarrow U_i(y) \right) \wedge \forall x \exists y \text{SUC}(x, y),$$

where $\text{SUC}(x, y)$ denotes that the 1-type of y is the successor of the 1-type of x , defined similarly as in Experiment 4. This formula is satisfiable and every model has size 2^n .

| n | Run time on \mathcal{E}_n^9 | | | | |
|-----|-------------------------------|------------|------------|------------|-------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 3.09s | 0.18s | 0.18s | 0.18s | 0.01s |
| 3 | 3.82s | 0.73s | 0.73s | 0.73s | 0.03s |
| 4 | TO | 3.04s | 3.05s | 3.06s | 0.1s |
| 5 | – | 13s | 13.07s | 13.09s | UN |
| 6 | – | 59.81s | 59.82s | 1m 0s | UN |
| 7 | – | 8m 21s | 8m 20s | 8m 19s | UN |
| 8 | – | 2h 29m 17s | 2h 30m 49s | 2h 30m 48s | UN |

Experiment 10 (with the equality predicate). The vocabulary is $\{U_1, \dots, U_n, E_1\}$. Formula \mathcal{E}_n^{10} states the following.

- Every element has an outgoing E -edge to another (different) element.
- There is at most one element that has 1-type $\bigwedge_{i=1}^n U_i(x)$, i.e., the maximal 1-type.
- Every E -edge (a, b) satisfies the following.
 - If the 1-types of a is the maximal 1-type, then the 1-type of b is also the maximal 1-type.
 - If the 1-types of a is not the maximal 1-type, then 1-type of b is the successor of 1-type of a .

Formally, it is defined as follows.

$$\begin{aligned} \mathcal{E}_n^{10} \quad := \quad & \forall x \forall y \, x \neq y \rightarrow \left((\neg \text{MAX}(x) \vee \neg \text{MAX}(y)) \wedge (E(x, y) \rightarrow \widetilde{\text{SUC}}(x, y)) \right) \\ & \wedge \quad \forall x \exists y \, E_1(x, y) \wedge x \neq y \end{aligned}$$

where:

$$\begin{aligned} \text{MAX}(x) \quad &:= \bigwedge_{i=1}^n U_i(x) \\ \widetilde{\text{SUC}}(x, y) \quad &:= \text{MAX}(x) \rightarrow \text{MAX}(y) \\ &\wedge \neg \text{MAX}(x) \rightarrow \bigvee_{i=1}^n \left(\neg U_i(x) \wedge U_i(y) \wedge \bigwedge_{j=i+1}^n (U_j(x) \leftrightarrow U_j(y)) \right. \\ &\quad \left. \wedge \bigwedge_{j=1}^{i-1} (U_j(x) \wedge \neg U_j(y)) \right) \end{aligned}$$

This formula is not satisfiable.

| n | Run time on \mathcal{E}_n^{10} | | | | |
|-----|----------------------------------|------------|------------|------------|-------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 4m 19s | 2.03s | 2.03s | 2.02s | 0.29s |
| 3 | TO | 23.01s | 23.05s | 22.99s | 0.03s |
| 4 | – | 5m 54s | 5m 55s | 5m 55s | 0.06s |
| 5 | – | 3h 30m 55s | 3h 31m 20s | 3h 31m 36s | UN |
| 6 | – | TO | TO | TO | UN |

Experiment 11 (with the equality predicate). The vocabulary is $\{U_1, \dots, U_n, E_1\}$. Formula \mathcal{E}_n^{11} is defined similarly as \mathcal{E}_n^{10} , with the difference in the definition of E which is defined according to U_n as follows. (a, b) is an edge if the following holds.

- If a belongs to U_n , then formula $\widetilde{\text{SUC}}(a, b)$ must hold, where $\widetilde{\text{SUC}}$ is as defined in Experiment 10.
- If a does not belongs to U_n , then b is the successor of a , where the successor is defined only on the predicates U_1, \dots, U_{n-1} .

Formally, it is defined as follows.

$$\begin{aligned} \mathcal{E}_n^{11} \quad := \quad & \forall x \forall y \, x \neq y \rightarrow \left((\neg \text{MAX}(x) \vee \neg \text{MAX}(y)) \wedge (E(x, y) \rightarrow \psi(x, y)) \right) \\ & \wedge \quad \forall x \exists y \, E_1(x, y) \wedge x \neq y \end{aligned}$$

where:

$$\begin{aligned} \psi(x, y) \quad := \quad & \left(U_n(x) \rightarrow \widetilde{\text{SUC}}(x, y) \right) \\ & \wedge \quad \left((\neg U_n(x) \rightarrow \text{SUC}_{n-1}(x, y)) \wedge (\text{ONE}_{1,n-1}(x) \rightarrow \text{ZERO}_{1,n-1}(y)) \right) \end{aligned}$$

Formulas MAX and $\widetilde{\text{SUC}}$ are as defined in Experiment 10, whereas $\text{ONE}_{1,n-1}$, $\text{ZERO}_{1,n-1}$ SUC_{n-1} are as in Experiment 6. This formula is satisfiable and the smallest model has size 2^{n-1} .

| n | Run time on \mathcal{E}_n^{11} | | | | |
|-----|----------------------------------|-----------|-----------|-----------|----|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 2.20s | 3.62s | 3.61s | 3.61s | UN |
| 3 | 3.43s | 39.49s | 39.63s | 39.58s | UN |
| 4 | 4.34s | 10m 8s | 10m 11s | 10m 10s | UN |
| 5 | TO | 5h 0m 29s | 5h 0m 30s | 5h 0m 21s | UN |
| 6 | – | TO | TO | TO | UN |

Experiment 12 (with the equality predicate)

The vocabulary is $\{U_1, \dots, U_n, V, E\}$. Formula \mathcal{E}_n^{12} is defined as follows.

- Every 1-type is realized only on at most one element.
- Every element has an outgoing edge to a different element, i.e., there is no self-loop.
- E -edges are anti-symmetric.
- For every E -edge (a, b) , the following holds.
 - The unary predicates U_1, \dots, U_n that hold on a are the same as those that hold on b .
 - V holds on a if and only if V does not hold on b .

Formally,

$$\mathcal{E}_n^{12} := \forall x \forall y \ x \neq y \rightarrow \left(\begin{array}{l} \neg \xi(x, y) \wedge (E(x, y) \rightarrow \neg E(y, x)) \\ \wedge \left(E(x, y) \rightarrow \left((V(x) \leftrightarrow \neg V(y)) \wedge \xi_U(x, y) \right) \right) \end{array} \right) \\ \wedge \forall x \exists y \ E(x, y) \wedge x \neq y$$

where:

$$\xi_U(x, y) := \bigwedge_{i=1}^n U_i(x) \leftrightarrow U_i(y) \quad \text{and} \quad \xi(x, y) := (V(x) \leftrightarrow V(y)) \wedge \xi_U(x, y)$$

This formula is not satisfiable.

| n | Run time on \mathcal{E}_n^{12} | | | | |
|-----|----------------------------------|------------|------------|------------|-------|
| | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 12m 57s | 0.65s | 0.65s | 0.65s | 0.28s |
| 3 | TO | 2.70s | 2.69s | 2.70s | 0.02s |
| 4 | – | 11.53s | 11.55s | 11.56s | 0.02s |
| 5 | – | 54.71s | 54.62s | 54.63s | 0.01s |
| 6 | – | 8m 16s | 8m 15s | 8m 17s | 0.01s |
| 7 | – | 2h 33m 49s | 2h 32m 52s | 2h 32m 52s | 0.02s |
| 8 | – | TO | TO | TO | 0.01s |

Here we should note that \mathcal{E}_n^{12} is of the form:

$$\forall x \forall y \ \alpha(x, y) \wedge \forall x \exists y \ \beta(x, y) \quad \text{where } \alpha(x, y) \wedge \beta(x, y) \text{ is a contradiction.}$$

In this case, our solver still tries to construct the graph system, instead of simply checking $\alpha(x, y) \wedge \beta(x, y)$ yields a contradiction. This explains why our solver performs much worse than Z3 here.

Experiment 13 (random formulas). This is the same as Experiment D in the main body. In this experiment we use random FO^2 formulas, which are obtained by first generating random graph systems $\mathcal{G} = (G_0, G_1)$, and then constructing the corresponding FO^2 formulas. Both G_0 and G_1 are generated using Erdős-Rényi model where the probability of an edge is $1/2$.

The constructed formula is of the form:

$$\forall x \forall y \alpha(x, y) \wedge \forall x \exists y \beta(x, y),$$

where $\alpha(x, y)$ and $\beta(x, y)$ are both in CNF. This is to avoid “explicit” listing of the edges as a formula. This way, our solver does not get the edges in \mathcal{G} for free, since it still needs to solve Boolean SAT to obtain them. Unfortunately, the constructed formulas are pretty huge, and in some cases, take even more than 20MB. Thus, we cannot display them here due to space constraints.

In the following n denotes the number of unary predicates and the number of vertices in the graph system is always 2^n .

| n | the number of vertices in \mathcal{G} | sat./unsat. | Run time on \mathcal{E}_n^{13} | | | | |
|-----|---|-------------|----------------------------------|--------|--------|--------|---------|
| | | | [15, 20] | ALG-A | ALG-B | ALG-C | Z3 |
| 2 | 4 | sat. | 2.02s | 0.04s | 0.04s | 0.03s | 0.01s |
| | | sat. | 2.16s | 0.03s | 0.03s | 0.03s | 0.01s |
| | | unsat. | 3.51s | 0.03s | 0.03s | 0.03s | 0.01s |
| | | unsat. | 3.35s | 0.03s | 0.02s | 0.02s | 0.01s |
| | | sat. | 2.19s | 0.03s | 0.03s | 0.03s | UN |
| | | sat. | 2.14s | 0.04s | 0.03s | 0.03s | UN |
| 3 | 8 | sat. | 2.48s | 0.05s | 0.05s | 0.05s | 0.01s |
| | | sat. | 2.48s | 0.06s | 0.05s | 0.05s | UN |
| | | sat. | 2.73s | 0.06s | 0.05s | 0.05s | 0.01s |
| | | sat. | 2.60s | 0.05s | 0.05s | 0.05s | 0.01s |
| | | sat. | 2.50s | 0.05s | 0.05s | 0.05s | 0.01s |
| | | sat. | 2.51s | 0.05s | 0.05s | 0.05s | 0.01s |
| 4 | 16 | sat. | 4.70s | 0.12s | 0.12s | 0.12s | 0.01s |
| | | sat. | 4.95s | 0.13s | 0.12s | 0.12s | UN |
| | | sat. | 4.44s | 0.13s | 0.12s | 0.12s | UN |
| | | sat. | 4.98s | 0.12s | 0.12s | 0.12s | 0.23s |
| | | sat. | 5.07s | 0.12s | 0.12s | 0.12s | 0.21s |
| 5 | 32 | sat. | 55.90s | 0.35s | 0.35s | 0.35s | 0.02s |
| | | sat. | 36.30s | 0.36s | 0.36s | 0.35s | 0.02s |
| | | sat. | 38.00s | 0.36s | 0.35s | 0.35s | UN |
| | | sat. | 50.54s | 0.36s | 0.36s | 0.36s | UN |
| | | sat. | 36.79s | 0.36s | 0.36s | 0.35s | 1.24s |
| 6 | 64 | sat. | 17m 25s | 1.84s | 1.83s | 1.83s | 8.19s |
| | | sat. | 23m 59s | 1.94s | 1.95s | 1.94s | 11.84s |
| | | sat. | 30m 23s | 1.90s | 1.90s | 1.90s | 0.09s |
| | | sat. | 41m 8s | 1.84s | 1.84s | 1.84s | 0.09s |
| | | sat. | 18m 21s | 1.92s | 1.92s | 1.92s | 0.09s |
| 7 | 128 | sat. | ER | 18.94s | 18.96s | 18.97s | 2m 50s |
| | | sat. | ER | 18.38s | 18.43s | 18.38s | UN |
| | | sat. | ER | 20.16s | 20.09s | 20.07s | 0.51s |
| 8 | 256 | sat. | ER | 6m 44s | 6m 43s | 6m 44s | 5.28s |
| | | sat. | ER | 6m 53s | 6m 53s | 6m 54s | 15m 52s |
| | | sat. | ER | 7m 47s | 7m 48s | 7m 47s | 5.12s |

Brief remarks on the experimental results. Our experiments show that the solver in [15, 20] performs badly when the smallest model of the input formula has cardinality bigger than 12 or when the formula is unsatisfiable. This is because the sizes of the generated clauses are propotional to the size of the model, and SAT solvers in general perform rather poorly on long clauses. The performance becomes significantly poorer for unsatisfiable formulas, as seen in Experiments 2, 7, 10 and 12, in which cases it has to iterate the SAT solver multiple times, with the model size increasing on each iteration.

Note also that the performance of [20] is pretty similar regardless of the “difficulty” of the input formula. For example, its performance on \mathcal{E}_n^1 and \mathcal{E}_n^6 are rather similar, but one can argue that \mathcal{E}_n^1 should be “easier” than \mathcal{E}_n^6 . One reason is that the graph system of \mathcal{E}_n^1 is conflict-free, whereas \mathcal{E}_n^6 is not. Moreover, the size of the graph system of \mathcal{E}_n^1 is linear in n , whereas for \mathcal{E}_n^6 it becomes exponential in n . That the graph system of \mathcal{E}_n^6 is the Moon-Moser graph obviously increases the difficulty.

On the other hand, one can see our algorithms performs very well on \mathcal{E}_n^1 where it can handle up to around 200 unary predicates. On \mathcal{E}_n^6 , it does not perform that well as it already hits “time out” when handling 16 unary predicates.

The performance of our algorithms drops significantly on formulas with the equality predicate. Nevertheless, they still outperform the others on most instances, except on Experiment 12 where Z3 is the clear winner.