Two-Variable Universal Logic with Transitive Closure*

Emanuel Kieroński and Jakub Michaliszyn

Institute of Computer Science, University of Wrocław, Poland
{kiero,jmi}@cs.uni.wroc.pl

Abstract

We prove that the satisfiability problem for the two-variable, universal fragment of first-order logic with constants (or, alternatively phrased, for the Bernays-Schönfinkel class with two universally quantified variables) remains decidable after augmenting the fragment by the transitive closure of a single binary relation. We give a 2-NExpTime-upper bound and a 2-ExpTime-lower bound for the complexity of the problem. We also study the cases in which the number of constants is restricted. It appears that with two constants the considered fragment has the finite model property and NExpTime-complete satisfiability problem. Adding a third constant does not change the complexity but allows to construct infinity axioms. A fourth constant lifts the lower complexity bound to 2-ExpTime. Finally, we observe that we are close to the border between decidability and undecidability: adding a third variable or the transitive closure of a second binary relation lead to undecidability.

1998 ACM Subject Classification F.4 Mathematical Logic and Formal Languages

Keywords and phrases two-variable logic, transitive closure, decidability

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

Classical papers from the 1930s showed that the satisfiability problem for first-order logic, FO, is undecidable. This raised the question which natural fragments of FO are decidable. A large research program led to a complete characterization, with respect to the decidability, of the so-called quantifier prefix classes. In particular, the Bernays-Schönfinkel class, i.e. the class of all formulas starting from a quantifier prefix of the form $\exists^*\forall^*$ followed by a quantifier free formula, appeared to be decidable. Note that, as existential quantifiers can be simulated by constants, the Bernays-Schönfinkel class may be alternatively viewed as the universal fragment of FO (i.e. the class of universal prenex-normal form FO formulas) with constants.

Another interesting decidable fragment of FO is the two-variable fragment, FO^2. With respect to the number of variables it appears to be the maximal fragment whose satisfiability problem is decidable, as undecidability of FO^3 follows from [8]. Decidability of FO^2 was shown in [15] by establishing a finite model property, namely, that every satisfiable formula has a finite model of size at most doubly exponential with respect to its length. This bound on the size of models was later improved in [5] to singly exponential, which implied a NExpTime-upper bound on the complexity of the satisfiability problem. A corresponding lower bound follows from [4, 13], so the satisfiability problem for FO^2 is NExpTime-complete.

The importance of FO^2 can be justified by the fact that it or its natural extensions and variants embed many formalisms used in computer science, such as modal, temporal

* Partially supported by Polish Ministry of Science and Higher Education grant N N206 371339.

Conference title on which this volume is based on.
Editors: Billy Editor, Bill Editors; pp. 1–18
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

or description logics. Unfortunately, FO^2 has a drawback, which becomes significant when one thinks about practical applications: it cannot express transitivity of a binary relation. Moreover, in contrast to modal logic or to some variants of the guarded fragment [16, 12], extending FO^2 by transitivity statements leads to undecidability [6, 10].

Actually, in applications for program verification or knowledge representation it would be even more desirable to have a transitive closure operator. While in the world of modal logics there exist decidable variants equipped with transitive closure operators, with a notable example of propositional dynamic logic, PDL [3], not too many natural decidable fragments of first-order logic with transitive closure are known. One exception is an extension of the two-variable guarded fragment with a transitive closure operator applied to binary symbols appearing only in guards. This is shown to be decidable and 2-ExpTime-complete in [14]. In a recent paper [11], FO^2 with the *equivalence* closure (i.e. reflexive, symmetric and transitive closure) operator is show to be decidable, and 2-NExpTime-complete, if the closure operator is applied to two distinguished binary symbols.

In [7] the universal fragment of first-order logic with constants is shown to be decidable when extended with the *deterministic* transitive closure operator, DTC, applied to a single, distinguished binary symbol, provided that only positive occurrences of DTC are allowed (thus we cannot say, e.g. that an element satisfying P is forbidden to be connected by a deterministic path to an element satisfying Q).

Some related results are obtained also in [2] where a logic motivated by the two-variable Bernays-Schönfinkel class extended with datalog is considered. This logic allows to state that some paths exist among constants, however, as it is actually a fragment of first-order logic, it is not able to express transitive closures.

In this paper we consider the universal, two-variable fragment of first-order logic with constants, and extend it with the transitive closure of a single, distinguished binary relation. In contrast to the mentioned fragment with DTC, we allow also for negative occurrences of transitive closures.

In [7] it is shown that if we allow to use the deterministic transitive closure or the transitive closure of a single binary relation both positively and negatively, then the universal fragment of FO becomes undecidable. The proof uses four universally quantified variables. Actually, Corollary 10 from [7] suggests that also the fragment with just two variables, two constants, and the transitive closure of one relation is undecidable. However, the statement of that corollary is not precise and there is no detailed proof. In this paper we clarify this issue by showing that in the case of two variables the satisfiability problem is decidable.

We also find quite intriguing that hardness of the investigated fragment depends on the number of constants (or, alternatively phrased, on the number of existential quantifiers in $\exists^*\forall^2$ formulas).

Our results and outline of the paper. To present our results precisely we introduce the following notation. We denote by $\forall^n_{TC}[m, k]$ the set of first-order formulas of the form $\forall x_1 \ldots x_n \varphi$, with quantifier free φ, over signatures containing m pairs of distinguished binary relation symbols: $R_1, R_1^+, \ldots, R_m, R_m^+$, k constant symbols $c_1, \ldots, c_k$, and no function symbols of arity greater than 0; the equality symbol is also allowed. We consider satisfiability of such formulas over structures in which for all $1 \leq i \leq m$ the interpretation of R_i^+ is the transitive closure of the interpretation of R_i. We define also the classes of formulas in which the number of constants is unbounded as $\forall^n_{TC}[m]=\bigcup_{i=0}^{\infty} \forall^n_{TC}[m, i]$.

We prove that the satisfiability problem for $\forall^2_{TC}[1]$ is decidable in 2-NExpTime (Section 6). In the case of $\forall^2_{TC}[1, 2]$ we show even an exponential model property, so it can be decided in NExpTime (Section 4). Slightly surprisingly, $\forall^2_{TC}[1, 3]$ lacks the finite model property

(Section 3), but we still are able to show a NExpTime-upper complexity bound (Section 7). The satisfiability problem for $\forall^2_{TC}[1,4]$ becomes 2-ExpTime-hard (Section 5). We also note some contrasting undecidability results, namely for $\forall^3_{TC}[1]$ and $\forall^2_{TC}[2]$ (Section 7).

2 Preliminaries

2.1 Conventions

We mostly work with $\forall^2_{TC}[1]$ and its fragments with bounded number of constants. In this case, we suppose without loss of generality that signatures contain only unary and binary relation symbols (cf. [5]), we denote by R the distinguished binary relation whose transitive closure is available, and use R^+ for this transitive closure. To simplify the presentation we assume that constants are not explicitly present in the signature, but rather they are simulated by means of special unary predicates $K_1, \ldots, K_k$. In this case we require that in a model of a given formula there exists exactly one element satisfying K_i, for all $1 \leq i \leq k$; we simply denote this element by c_i. We do not obey this assumption when presenting example formulas and proving lower bounds. Eliminating constants in favor of such special unary predicates can be done in a standard way.

We use a standard convention and if $\mathfrak{A}$ is a structure then we denote its universe by A. Similarly, if $V \subseteq A$ then we denote by $\mathfrak{V}$ the substructure of $\mathfrak{A}$ induced by V, i.e. $\mathfrak{A}\restriction V$.

2.2 Atomic types

An (atomic) 1-*type* (over a given signature) is a maximal satisfiable set of atoms or negated atoms with free variable x. Similarly, an (atomic) 2-*type* is a maximal satisfiable set of atoms and negated atoms with free variables x, y. We assume that literals built using our special symbol R^+ are also members of atomic types. Note that the numbers of 1-types and 2-types are bounded exponentially in the size of the signature. We often identify a type with the conjunction of all its elements.

Observe that in the case of signatures restricted to unary and binary symbols, to completely describe a structure it is enough to list the 2-types of all pairs of elements. However, we usually start our constructions by defining 1-types.

For a given σ-structure $\mathfrak{A}$, and $a \in A$ we say that a *realizes* a 1-type α if α is the unique 1-type such that $\mathfrak{A} \models \alpha[a]$. We denote by $\mathrm{tp}_{\mathfrak{A}}(a)$ the 1-type realized by a. Similarly, for distinct $a, b \in A$, we denote by $\mathrm{tp}_{\mathfrak{A}}(a, b)$ the unique 2-type *realized* by the pair a, b, i.e. the type β such that $\mathfrak{A} \models \beta[a, b]$. We denote by $\boldsymbol{\alpha}[\mathfrak{A}]$ the set of all 1-types, and by $\boldsymbol{\beta}[\mathfrak{A}]$ the set of all 2-types realized in $\mathfrak{A}$. For $S_1, S_2 \subseteq A$, we denote by $\boldsymbol{\alpha}_{\mathfrak{A}}[S_1]$ the set of all 1-types realized in S_1, by $\boldsymbol{\beta}_{\mathfrak{A}}[S_1, S_2]$ the set of all 2-types $\mathrm{tp}_{\mathfrak{A}}(a_1, a_2)$ with $a_i \in S_i$. We sometimes skip subscripts if the structure is clear from the context.

2.3 Small cliques

Let $\mathfrak{A}$ be a structure. We say that $C \subseteq A$ is an R^+-*clique*, or simply a *clique*, if C is a maximal set of elements such that for all distinct $a, b \in C$ we have $\mathfrak{A} \models aR^+b \wedge bR^+a$. In the other words an R^+-clique is a maximal strongly R-connected component in $\mathfrak{A}$. We show that we can restrict our attention to structures with cliques of a bounded size.

▶ **Lemma 1.** *Let φ be a formula in $\forall^2_{TC}[1]$ and let $\mathfrak{A} \models \varphi$. Then there exists a model of φ such that the size of every R^+-clique in this model is bounded exponentially in $|\varphi|$.*

Towards a proof of this lemma we first show how to replace a single R^+-clique C in $\mathfrak{A}$ by its small counterpart C'. In [14] the following lemma is proved.

▶ **Lemma 2.** *Let φ be an FO^2 formula and $\mathfrak{M} \models \varphi$ its strongly R-connected model (an R^+-clique). Then there exists a strongly R-connected model $\mathfrak{M}' \models \varphi$ of size bounded exponentially in $|\varphi|$ such that $\boldsymbol{\alpha}[\mathfrak{M}] = \boldsymbol{\alpha}[\mathfrak{M}']$.*

We apply the above lemma to $\mathfrak{C}$ and $\psi = \varphi \wedge \psi^c$, where $\psi^c = \forall xy \bigwedge_i (K_i(x) \wedge K_i(y) \rightarrow x = y)$, obtaining a structure $\mathfrak{C}'$. In particular $\mathfrak{C}'$ contains realizations of the same special predicates K_i as $\mathfrak{C}$, and each of them is realized at most once. It remains to connect $\mathfrak{C}'$ with $\mathfrak{A}{\restriction}A \setminus C$. For any $a \in A \setminus C$ and any $\alpha \in \boldsymbol{\alpha}[C]$, if there exists $b \in C$, of type α, such that $\mathfrak{A} \models aRb \vee bRa$ then we set $b' = b$. Otherwise we choose an arbitrary element of type α in C as b'. For every element $b'' \in C'$ of type α we set $\mathrm{tp}_{\mathfrak{A}'}(a, b'') = \mathrm{tp}_{\mathfrak{A}}(a, b')$. Let us denote by $\mathfrak{A}'$ the structure so obtained. The proof of the following claim is omitted due to page limit.

▶ **Claim 3.** $\mathfrak{A}'$ is indeed a model of φ.

In the case of a finite model we apply the above step successively to all R^+-cliques, obtaining finally a model with small cliques. For the case of an infinite model, note that $\forall^2_{TC}[1]$ satisfies downward Löwenheim-Skolem property, so we may assume that the initial model is countable, and apply our procedure to all R^+-cliques in countably many steps. The desired model with small R^+-cliques is the natural limit of the described process. This finishes our proof of Lemma 1.

For a pair of distinct elements a, b we say that they are in *free position* in $\mathfrak{A}$ if $\mathfrak{A} \models \neg aR^+b \wedge \neg bR^+a$. A clique C_1 is in free position with C_2 if every element from C_1 is in free position with every element of C_2.

2.4 Saturations

In our constructions it is sometimes convenient to have structures with many R-edges. Let $\mathfrak{A}$ be a structure and let us build $\mathfrak{A}'$ by adding to $\mathfrak{A}$ a number of R-edges, in the following way. If there is a pair of elements $a_1, a_2 \in A$ such that $\mathfrak{A} \models a_1Ra_2 \wedge \neg a_2R^+a_1$ and a pair of elements $b_1, b_2 \in A$, such that $\mathrm{tp}_{\mathfrak{A}}(a_1) = \mathrm{tp}_{\mathfrak{A}}(b_1)$, $\mathrm{tp}_{\mathfrak{A}}(a_2) = \mathrm{tp}_{\mathfrak{A}}(b_2)$ and $\mathfrak{A} \models b_1R^+b_2 \wedge \neg b_1Rb_2 \wedge \neg b_2R^+b_1$, then we modify the 2-type of b_1, b_2 by setting $\mathrm{tp}_{\mathfrak{A}'}(b_1, b_2) = \mathrm{tp}_{\mathfrak{A}}(a_1, a_2)$. We repeat this step until no further modifications are possible. We call the obtained structure an *R-saturation* of $\mathfrak{A}$. A structure which is its own R-saturation is called *R-saturated.*

Note that the R-edges added in the above process do not change the R^+-relations among the elements. As all the modified 2-types are realized in $\mathfrak{A}$, we have the following proposition.

▶ **Proposition 4.** Let φ be a $\forall^2_{TC}[1]$ formula and $\mathfrak{A}$ its model. Then an R-saturation of $\mathfrak{A}$ is an R-saturated model of φ.

3 An infinity axiom

To demonstrate the strength of the considered fragment we show in this section that there exists a $\forall^2_{TC}[1, 3]$-formula $\eta = \forall xy\eta_0$ which is satisfiable but has only infinite models.

We define η_0 as the conjunction of formulas (1)-(3) below.

(1) there exists a path from c_1 to c_2 and there are no R^+-loops.

$$c_1R^+c_2 \wedge \neg xR^+x,$$

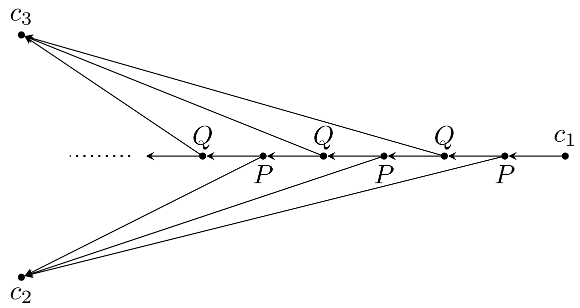

Figure 1 An infinite model of η.

(2) P and Q are disjoint, every element in P has an R^+ path to c_3, and every element in Q has an R^+-path to c_2.

$$(Px \wedge Qx \rightarrow \bot) \wedge (Px \rightarrow xR^+c_3) \wedge (Qx \rightarrow xR^+c_2),$$

(3) R-edges are allowed only between elements of specific types.

$$xRy \rightarrow ((x = c_1 \wedge Py) \vee (Px \wedge Qy) \vee (Qx \wedge Py) \vee (Px \wedge y = c_2) \vee (Qx \wedge y = c_3)).$$

It is not hard to see that η is satisfied in the infinite model depicted in Fig.1. Also any model of η must embed an infinite chain of elements, on which predicates P and Q alternate.

4 A finite model property for formulas with two constants

Now we show that the presence of three constants in the previous section was essential.

▶ **Lemma 5.** *Every satisfiable* $\forall^2_{TC}[1,2]$*-formula* φ *has a finite model of size bounded exponentially in* $|\varphi|$.

Let $\mathfrak{A} \models \varphi$ be a model with cliques bounded exponentially in $|\varphi|$, as guaranteed by Lemma 1. By Proposition 4 we may assume that $\mathfrak{A}$ is R-saturated. Let C_1 be the clique containing c_1, and C_2 be the clique containing c_2.

Note that if $C_1{=}C_2$ then $\mathfrak{A}\restriction C_1 \models \varphi$, and that if $\mathfrak{A} \models \neg c_1R^+c_2 \wedge \neg c_2R^+c_1$ then $\mathfrak{A}\restriction C_1 \cup C_2 \models \varphi$. In both cases we have finite models of φ of exponentially bounded size.

Consider the case when $\mathfrak{A} \models c_1R^+c_2 \wedge \neg c_2R^+c_1$ (the symmetric case can be treated analogously). Let us take a shortest path π from c_1 to c_2. Let us write π as $c_1 = a_{11}, a_{12}, \ldots, a_{1k_1}, a_{21}, a_{22}, \ldots, a_{2k_2}, \ldots, a_{l1}, a_{l2}, \ldots, a_{lk_l} = c_2$, where for each i the path $a_{i1}, \ldots, a_{ik_i}$ is the maximal fragment of π containing elements from the same clique. We denote by C_i the clique containing the elements a_{ij}. Observe that if π leaves a clique C_i then it never enters it again, i.e. if $1 \leq i < j \leq l$ then $C_i \neq C_j$.

We claim that $\mathfrak{A}' = \mathfrak{A}\restriction C_1 \cup \ldots \cup C_l$ is a model of φ. Indeed, if two elements belong to the same clique in $\mathfrak{A}'$ then they also belong the same clique in $\mathfrak{A}$; if a pair of elements is connected non-symmetrically by R^+ in $\mathfrak{A}'$ then they are also connected non-symmetrically by R^+ in $\mathfrak{A}$; finally, there are no elements in free position in $\mathfrak{A}'$. Thus all atomic 2-types realized in $\mathfrak{A}'$ are also realized in $\mathfrak{A}$, which implies that $\mathfrak{A}' \models \varphi$. Note that taking whole cliques of elements from π to $\mathfrak{A}'$, instead of considering just $\mathfrak{A}\restriction\pi$, is important, as φ may require some elements to lie on an R-cycle.

We claim that the size of $\mathfrak{A}'$ is bounded exponentially in $|\varphi|$. This follows from the fact that for $1 \leq i < j \leq l$ we have $\mathrm{tp}(a_{i1}) \neq \mathrm{tp}(a_{j1})$. Indeed, assume to the contrary that for some i, j we have that $\mathrm{tp}(a_{i1}) = \mathrm{tp}(a_{j1})$. Then the path π' obtained from π by removing the fragment $a_{i1}, \ldots, a_{j-1,k_{j-1}}$ is a path from c_1 to c_2, which is shorter than π. Note that π' is indeed an R-path, since $\mathfrak{A} \models a_{i-1,k_{i-1}} R a_{i1}$, and thus, by R-saturation of $\mathfrak{A}$, we have also $\mathfrak{A} \models a_{i-1,k_{i-1}} R a_{j1}$. Thus the number of cliques in $\mathfrak{A}'$ is not greater than $|\boldsymbol{\alpha}|$, the size of every clique is bounded exponentially in $|\varphi|$, and thus also $|A'|$ is bounded exponentially in $|\varphi|$.

This finishes the proof of Lemma 5. It naturally leads to the following complexity result.

▶ **Theorem 6.** *The satisfiability problem for* $\forall^2_{TC}[1, 2]$ *is decidable in* NExpTime.

A corresponding lower bound can be obtained even in the absence of constants (assuming that we consider satisfiability in non-empty structures). The idea is similar to the proof of Theorem 5 from [7]. We construct a formula whose models are grids of exponential size. Instead of using two constants to distinguish the left-upper and the right-lower corners of the grid we say that every element is R-reachable from itself but not by a direct R-edge: $xR^+x \wedge \neg xRx$. We allow edges only between elements which are neighbors on a snake-like path through the whole grid. We allow also for an R-edge from the right-lower corner to the left-upper corner. Thus models are R-cycles which have to contain all elements of the grid.

▶ **Theorem 7.** *The satisfiability problem for* $\forall^2_{TC}[1, 0]$ *is* NExpTime*-hard.*

5 Lower bound for formulas with four constants

Now we show that in the presence of four constants the lower bound for the satisfiability problem can be lifted to 2-ExpTime. To simplify the presentation we assume first that there are nine constants available, and then we present a trick which allows to get rid of five of them.

5.1 A construction involving nine constants

The proof goes by a reduction from alternating Turing machines with exponentially bounded space. The general idea of the proof and the shape of intended models are similar to the ones used in [9]. However, the lack of existential quantifiers makes the tasks of enforcing desired shapes of models and then simulating Turing machines more tricky.

Tree-like structures. To simulate a run of an alternating Turing machine it is convenient to have a structure which resembles an infinite binary tree, with each node being able to encode a single configuration, and identify its successor nodes. Let us describe how to enforce a desired structure.

We use unary predicates $P_0, \ldots, P_{n-1}$ and assume that for any element a they encode a value $0 \leq \bar{P}(a) < 2^n$ in a natural way, i.e. $P_i(a)$ is true exactly if the ith bit of the binary representation of $\bar{P}(a)$ is equal to 1. Let us abbreviate by $\bar{P}(x) = \bar{P}(y)$, $\bar{P}(x) = \bar{P}(y) + 1$, $\bar{P}(x) = k$ (for $0 \leq k < 2^n$) quantifier-free formulas with an obvious meaning. Such formulas can be constructed of size polynomial in n in a standard fashion.

We say that elements $a_0, \ldots, a_{2^n-1}$ form a *node* in a structure $\mathfrak{A}$ if $\bar{P}(a_i) = i$ and $\mathfrak{A} \models a_{i-1} R a_i$ for $0 < i < 2^n$. The purpose of a node will be to encode information about a single configuration of a Turing machine. We use unary predicates H^d_i for $0 \leq i < 4$, $d \in \{L, R\}$ to distinguish eight types of nodes. An additional predicate H^I serves for distinguishing an initial node.

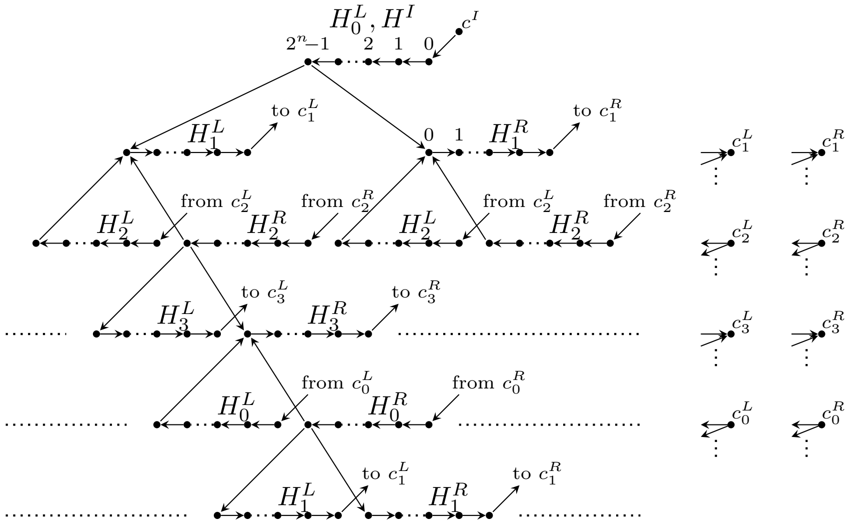

Figure 2 An initial fragment of the structure $\mathfrak{T}$ from the proof of the lower bound.

Let $\mathfrak{T}$ be the structure depicted in Fig. 2. It is drawn in a way suggesting its similarity to a binary tree, note however that actually this structure is shallow: every R-path has length not greater than $2^n + 2$.

▶ **Claim 8.** There exists a formula λ such that:

(a) $\mathfrak{T} \models \lambda$

(b) any model $\mathfrak{A} \models \lambda$ locally resembles $\mathfrak{T}$, i.e. there exists a node of type H_0^L satisfying H^I, and for every node $a_0, \ldots, a_{2^n-1}$ of type H_i^d there exists a *left successor node* $a_0^L, \ldots, a_{2^n-1}^L$ of type $H^L_{i+1 \mod 4}$ and a *right successor node* $a_0^R, \ldots, a_{2^n-1}^R$ of type $H^R_{i+1 \mod 4}$ such that if i is even then $\mathfrak{A} \models a_{2^n-1} R a_0^L \wedge a_{2^n-1} R a_0^R$ and if i is odd then $\mathfrak{A} \models a^L_{2^n-1} R a_0 \wedge a^R_{2^n-1} R a_0$.

We construct λ from five conjuncts. Conjuncts (1) and (2) say that for some elements there are paths from or to some constants. Conjuncts (3)-(5) say that R-edges are allowed only between elements of specific 1-types (actually only such types whose realizations are connected by an R-edge in $\mathfrak{T}$). Below we describe these conjuncts in more details.

(1) there is an R-path from c^I to c_1^L.

(2) every element satisfying H_0^L or H_0^R can reach (by some R^+-paths) elements c_1^L and c_1^R; every element satisfying H_2^L or H_2^R can reach elements c_3^L and c_3^R; every element satisfying H_1^L or H_1^R can be reached from elements c_2^L and c_2^R; every element satisfying H_3^L or H_3^R can be reached from elements c_0^L and c_0^R.

(3) (edges incident to constants) for $i \in \{0, 2\}$ and $d \in \{L, R\}$ element c_i^d has no incoming R-edges, and has outgoing R-edges only to elements a such that $\bar{P}(a) = 0$ and $H_i^d(a)$ holds; for $i \in \{1, 3\}$ and $d \in \{L, R\}$ element c_i^d has no outgoing R-edges, and has incoming R-edges only from elements a such that $\bar{P}(a) = 2^n - 1$ and $H_i^d(a)$ holds; c^I has no incoming R-edges and has outgoing R-edges only to elements a such that $\bar{P}(a) = 0$ and $H_0^L(a) \wedge H^I(a)$ holds.

(4) (edges inside nodes) if an element a satisfies $\bar{P}(a) < 2^n - 1 \wedge H_i^d(a)$ than it has outgoing edges only to elements b satisfying $H_i^d(b)$ such that $\bar{P}(b) = \bar{P}(a) + 1$ and $H^I(a) \leftrightarrow H^I(b)$;

if an element a satisfies $\bar{P}(a) > 0 \wedge H_i^d(a)$ than it has incoming edges only from elements b satisfying $H_i^d(b)$ such that $\bar{P}(a) = \bar{P}(b) + 1$ and $H^I(a) \leftrightarrow H^I(b)$;

(5) (edges among nodes) an element a such that $\bar{P}(a) = 2^n - 1$ and $H_i^d(a)$ for $i \in \{0, 2\}$ hold has incoming edges only from elements in H_i^d, and has outgoing edges only to elements b such that $\bar{P}(b) = 0$ and $H_{i+1}^L(b) \vee H_{i+1}^R(b) \vee H_{i-1 \bmod 4}^L(b) \vee H_{i-1 \bmod 4}^R(b)$; an element a such that $\bar{P}(a) = 0$ and $H_i^d(a)$ for $i \in \{1, 3\}$ hold has an outgoing edges only from elements in H_i^d, and has incoming edges only from elements b such that $\bar{P}(b) = 2^n - 1$ and $H_{i+1 \bmod 4}^L(b) \vee H_{i+1 \bmod 4}^R(b) \vee H_{i-1}^L(b) \vee H_{i-1}^R(b)$.

Clearly $\mathfrak{T} \models \lambda$. Consider an arbitrary model $\mathfrak{A} \models \lambda$. By (1) there must be a path from C^I to c_1^L. By (3) this path must begin with an edge to an element a_0 such that $\mathfrak{A} \models \bar{P}(a_0) = 0 \wedge H^I(a_0) \wedge H_0^L(a_0)$. Then, by (4) this path must go through a whole node of type H_0^L, satisfying also H^I. The last element of this node must have by (2) a path to c_1^L and a path to c_1^R. By (5) the first of this paths must go through an element a_1^L satisfying $\mathfrak{A} \models \bar{P}(a_1^L) = 0 \wedge H_1^L(a_1^L)$, and the other through an element a_1^R satisfying $\mathfrak{A} \models \bar{P}(a_1^R) = 0 \wedge H_1^R(a_1^R)$. Both paths must then go through whole nodes of appropriate types. Elements a_1^L and a_1^R must have by (2) paths from c_2^L and c_2^R, which again have to go through whole nodes of types H_2^L and H_2^R. This reasoning can be generalized to an inductive argument that part (b) of Claim 8 holds.

Simulating alternating Turing machines. A well-known theorem from [1] says that 2-ExpTime is equal to AExpSpace, the class of problems solvable by *alternating* Turing machines in exponentially bounded space.

For a given alternating machine M and its input w we can construct a formula κ_w^M which is satisfiable iff M accepts w. We define κ_w^M as the conjunction of λ and some formulas encoding computations of M. Every element of a model of λ corresponds to single tape cell of M, and stores information about this cell, as well as about the two neighboring cells. Thus, formulas of the form $(xR^+y \wedge \bar{P}(x) = \bar{P}(y) \wedge H_i^d(x) \wedge H_{i+1 \bmod 4}^d(y)) \rightarrow \ldots$ can be used to say that two consecutive nodes of a model describe two consecutive configurations of M. Details are omitted due to space limit.

5.2 Four constants suffice

The following lemma will be used to reduce the number of constants required in the proof of 2-ExpTime-hardness from nine to four. Actually, it has a stronger statement and allows to reduce satisfiability of $\forall^2_{TC}[1, n]$ and even $\forall^2_{TC}[1]$ in polynomial time to satisfiability of $\forall^2_{TC}[1, 4]$, assuming that we consider only structures in which relation R^+ restricted to constants is a partial order.

▶ **Lemma 9.** *For each n and each $\forall^2_{TC}[1, n]$ sentence φ there is a polynomially computable $\forall^2_{TC}[1, 4]$ formula φ' such that φ' has a model if and only if φ has a model in which for all $i < j$ there is no R-path from c_j to c_i.*

We sketch the main idea of the proof. Assume that constants $c_1, \ldots, c_4$ are available. We simulate n additional constants by n fresh unary predicates $S_1, \ldots, S_n$. We use also auxiliary unary predicates $P_1, \ldots, P_n$, $Q_1, \ldots, Q_n$. We say that each of the predicates S_i, P_i, Q_i is satisfied in at most one element. We want to enforce that each of S_i is satisfied at least once, and for $i < j$, each pair of realizations of S_i, S_j may appear either in free position or may be connected by an R^+ path from the one satisfying S_i to the one satisfying S_j. To do so we enforce first the upper and the lower horizontal chains of elements from Fig. 3. Then we say that the element satisfying P_i has an R-path to the element satisfying

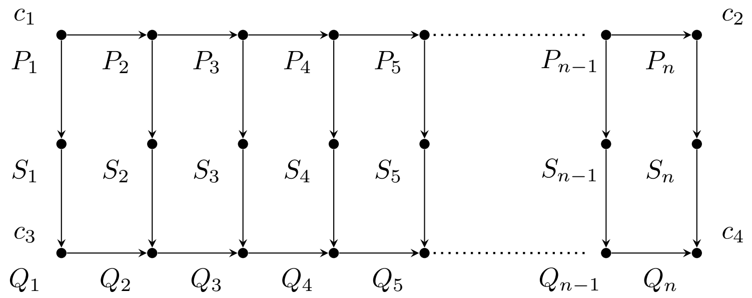

Figure 3 A model of ψ.

Q_i. By an appropriate restriction of 2-types containing R we can enforce that these paths go through elements satisfying S_i. We guarantee that all S_i are realized, by saying that there are no R-paths from P_i to Q_j for $i > j$. Here the assumption from the statement of the lemma, about admissible R^+-connections among constants is relevant. Details of the proof of Lemma 9 are omitted due to space limit.

We are now ready to formulate the following theorem.

▶ **Theorem 10.** *The satisfiability problem for* $\forall^2_{TC}[1,4]$ *is 2-*EXPTIME*-hard.*

Proof. We define λ^* by renaming the constants in λ in the following way: $c^I \to c_1$, $c_2^L \to c_2$, $c_2^R \to c_3$, $c_4^L \to c_4$, $c_4^R \to c_5$, $c_1^L \to c_6$, $c_1^R \to c_7$, $c_3^L \to c_8$, $c_3^R \to c_9$. Clearly, renaming the constants does not change the properties of formulas. Moreover, λ^* guarantees that $c_1 - c_5$ have no incoming edges and $c_6 - c_9$ have no outgoing edges, and therefore in any model of λ^* there are no paths from c_j to c_i for any $i < j$. We apply Lemma 9 to λ^* obtaining λ'. We can now replace λ by λ' when constructing κ_w^M from the previous subsection. ◀

6 Decidability of formulas with an unbounded number of constants

In this section we show that the satisfiability problem for $\forall^2_{TC}[1]$ is decidable. We use a standard approach which consists in an analysis of arbitrary models and rebuilding them to obtain a shape which admits descriptions of a bounded size. In this case we show that every formula has a model which can be divided into at most doubly exponentially many fragments, called *zones*, each of which is either a clique or an infinite, regular chain of cliques.

6.1 Clique types

Let $\mathfrak{A}$ be a structure. We say that a clique C has a *clique type* $\delta = (\mathcal{C}, \mathcal{A}, \mathcal{B})$ in $\mathfrak{A}$, if $\mathcal{C}$ is the set of atomic 1-types realized in C, $\mathcal{A}$ is the set of atomic 1-types of the elements located *above* C, i.e. the elements b such that for all $a \in C$ we have $\mathfrak{A} \models bR^+a \wedge \neg aR^+a$, and $\mathcal{B}$ is the set of atomic 1-types of the elements located *below* C, i.e. the elements b such that for all $a \in C$ we have $\mathfrak{A} \models aR^+b \wedge \neg bR^+a$. We denote by $\Delta[\mathfrak{A}]$ the set of all clique types realized in $\mathfrak{A}$. Note that $|\Delta[\mathfrak{A}]|$ is bounded doubly exponentially in the signature.

6.2 Zones

For a pair of cliques C_1, C_2 we write $C_1 \leq_c C_2$ if $C_1 = C_2$ or for all $a_1 \in C_1, a_2 \in C_2$ we have $a_1R^+a_2$. Relation $\leq_c$ naturally induces a relation $\leq_\delta$ on clique types. We define:

$\delta_1 \leq_\delta \delta_2$ iff there exist cliques $C_1, C_2 \subseteq A$, C_i of type δ_i, such that $C_1 \leq_c C_2$. Let $\leq^*_\delta$ be the transitive closure of $\leq_\delta$. Let $\delta_1 \approx \delta_2$ iff $\delta_1 \leq^* \delta_2$ and $\delta_2 \leq^* \delta_1$. Clearly, $\approx$ is an equivalence relation over $\Delta[\mathfrak{A}]$. The set of elements of $\mathfrak{A}$, belonging to the cliques realizing the extended types from the same equivalence class of $\approx$, is called a *zone*. Note that the number of zones of $\mathfrak{A}$ is bounded doubly exponentially in the signature.

We say that a zone V is *singular* if every R^+-connection inside V is symmetric. A few simple properties of zones, having straightforward proofs, are collected below.

▶ **Proposition 11.** (i) Let $\delta_1 = (\mathcal{C}_1, \mathcal{A}_1, \mathcal{B}_1)$ and $\delta_2 = (\mathcal{C}_2, \mathcal{A}_2, \mathcal{B}_2)$ be two clique types realized in a zone V. Then $\mathcal{A}_1 = \mathcal{A}_2$ and $\mathcal{B}_1 = \mathcal{B}_2$.

(ii) If a zone V is singular then V contains only realizations of a single clique type.

(iii) Let $\delta = (\mathcal{C}, \mathcal{A}, \mathcal{B})$ be a clique type realized in a non-singular zone V. Then for every $\alpha \in \mathcal{C}$ we have $\alpha \in \mathcal{A}$ and $\alpha \in \mathcal{B}$.

(iv) Let α_1 and α_2 be atomic types realized in a non-singular zone V. Then there exists a pair of elements a_1, a_2 in $\mathfrak{A}$ (but not necessarily in $\mathfrak{V}$) such that $\mathrm{tp}(a_1) = \alpha_1$, $\mathrm{tp}(a_2) = \alpha_2$, and $\mathfrak{A} \models a_1 R^+ a_2 \wedge \neg a_2 R^+ a_1$.

(v) Let π be a path connecting two elements belonging to a non-singular zone V. Then every element a on π belongs to V.

6.3 Making zones regular

Let V be a zone in a structure $\mathfrak{A}$. We show how to replace $\mathfrak{V}$ by a zone $\mathfrak{V}'$ being either a single clique or an infinite, regular chain of cliques, in such a way that the resulting structure $\mathfrak{A}'$ satisfies all $\forall^2_{TC}[1]$ formulas satisfied in $\mathfrak{A}$.

Building a singular zone. If V is singular then it consists of some number of cliques in free position, and, by Proposition 11 (ii), all of them have the same clique type δ. In this case $\mathfrak{V}'$ is a single realization of δ.

Building a non-singular zone. Consider a non-singular zone V. By Proposition 11 (i) there are $\mathcal{A}$, $\mathcal{B}$ such that every clique type realized in $\mathfrak{V}$ has the form $(\mathcal{C}, \mathcal{A}, \mathcal{B})$ for some set $\mathcal{C}$. The construction of a regular version of a $\mathfrak{V}$ relies on the following proposition.

▶ **Proposition 12.** There exists a sequence of (not necessarily distinct) clique types $\delta_0, \ldots, \delta_{l-1}$, where $\delta_i = (\mathcal{C}_i, \mathcal{A}, \mathcal{B})$, and atomic types $\alpha_0^{in}, \alpha_0^{out}, \ldots, \alpha_{l-1}^{in}, \alpha_{l-1}^{out}$ such that:

(a) l is bounded exponentially in the size of the signature,
(b) for every $\alpha \in \boldsymbol{\alpha}[V]$ there exists i such that $\alpha \in \mathcal{C}_i$,
(c) for every i, δ_i is a clique type of a clique in $\mathfrak{V}$,
(d) for every i we have $\alpha_i^{in}, \alpha_i^{out} \in \mathcal{C}_i$,
(e) for every i there exists in $\mathfrak{A}$ a realization a of α_i^{out} and a realization b of $\alpha_{i+1 \bmod l}^{in}$ such that $\mathfrak{A} \models aRb \wedge \neg bR^+a$.

Proof. Let $\delta'_0, \ldots, \delta'_{s-1}$ be an enumeration of all clique types from $\Delta[\mathfrak{V}]$. By the definition of a zone and the relation $\leq_\delta$ there is a $\leq_\delta$-path from δ'_i to $\delta'_{i+1 \bmod s}$, for every $0 \leq i < s$. By concatenating such paths we obtain a sequence $\delta_0, \ldots, \delta_{t-1}$ meeting conditions (b)-(e) (assuming a natural choice of α_i^{in} and α_i^{out}). In this path we choose for every $\alpha \in \boldsymbol{\alpha}[\mathfrak{V}]$ a clique type $\delta_\alpha = (\mathcal{C}_\alpha, \mathcal{A}_\alpha, \mathcal{B}_\alpha)$ such that $\alpha \in \mathcal{C}_\alpha$. Observe that if $\alpha_i^{in} = \alpha_j^{in}$ for some $0 \leq i < j < t$ such that $\delta_i, \ldots, \delta_{j-1}$ does not contain any δ_α then we can remove $\delta_i, \ldots, \delta_{j-1}$ from the sequence without violating conditions (b)-(e). This observation allows to easily shorten the sequence to a required length. ◀

Let $\delta_0, \ldots, \delta_{l-1}$ be a sequence of clique types guaranteed by Proposition 12. We construct $\mathfrak{V}'$ as an infinite chain of cliques $\ldots C_{-2}, C_{-1}, C_0, C_1, C_2, \ldots$ such that the clique C_i has type $\delta_{i \mod l}$. For every pair $\alpha_1, \alpha_2 \in \boldsymbol{\alpha}[V]$ we choose a 2-type $\beta_{1\to 2} \models xR^+y \wedge \neg yR^+x \wedge \alpha_1(x) \wedge \alpha_2(y)$ realized in $\mathfrak{A}$. An appropriate $\beta_{1\to 2}$ exists in $\boldsymbol{\beta}[\mathfrak{A}]$ by Proposition 11 (iv). If it is possible we choose $\beta_{1\to 2}$ containing xRy. For all $a_1 \in C_i, a_2 \in C_j$, $i < j$, such that $\text{tp}(a_1) = \alpha_1$, $\text{tp}(a_2) = \alpha_2$ we set $\text{tp}(a_1, a_2) := \beta_{1\to 2}$. This finishes the construction of $\mathfrak{V}'$. Note that by our choice of atomic 2-types and condition (e) from Proposition 12, we have that for all $i < j$ there exists an R-path from each element of C_i to each element of C_j.

Connecting a rebuilt zone to the remaining part of the model. Consider an element $a \in A \setminus V$. Let $\alpha = \text{tp}_{\mathfrak{A}}(a)$. We distinguish three cases.

Case 1: In $\mathfrak{A}$ element a is in free position with all elements in V. For any 1-type $\alpha' \in \boldsymbol{\alpha}[\mathfrak{V}']$ we find an element $b \in V$ of type α' (such an element exists as our construction ensures that $\boldsymbol{\alpha}[\mathfrak{V}'] = \boldsymbol{\alpha}[\mathfrak{V}]$), and for any $b' \in V'$ of type α' we set $\text{tp}_{\mathfrak{A}'}(a, b') = \text{tp}_{\mathfrak{A}}(a, b)$. Clearly this ensures that a is in free position with all elements from V'.

Case 2: In $\mathfrak{A}$ there is an R-path from a to an element of V. For any 1-type $\alpha' \in \boldsymbol{\alpha}[V]$:

- if there exists a realization $b \in V$ of α' such that $\mathfrak{A} \models aRb$ then for all $b' \in V'$ of type α' we set $\text{tp}_{\mathfrak{A}'}(a, b') = \text{tp}_{\mathfrak{A}}(a, b)$.
- otherwise find a realization $b \in A$ of α' such that $\mathfrak{A} \models aR^+b$ and for all $b' \in V'$ of type α' we set $\text{tp}_{\mathfrak{A}'}(a, b') = \text{tp}_{\mathfrak{A}}(a, b)$. Note that in this subcase the existence of an appropriate b is guaranteed by the properties of relation $\leq_\delta$, but sometimes we need to look for it outside V.

Note that in this case element a has an R-path in $\mathfrak{A}'$ to every element from V'. Indeed, on a path from a to an element of V there must be an element, say b, which has an R-edge to a point from V. This element b will be made R^+-connected to all elements from V; in particular, if V is non-singular it will have R-edges to infinitely many elements of V.

Case 3: In $\mathfrak{A}$ there is an R-path from an element of V to a. Proceed analogously to Case 2.

Modifying the remaining part of the model. To complete the construction of $\mathfrak{A}'$ consider a pair of elements $a_1, a_2 \in A \setminus V$. If $\mathfrak{A} \models a_1R^+b \wedge b'R^+a_2$ (or symmetrically $\mathfrak{A} \models a_2R^+b \wedge b'R^+a_1$) for some elements $b, b' \in V$ then a_1 becomes R^+-connected to a_2 in $\mathfrak{A}'$, even if they are not connected in $\mathfrak{A}$. Note that in this case $a_1 \in \mathcal{A}$ and $a_2 \in \mathcal{B}$. This means that there is a pair of realizations a'_1, a'_2 of $\text{tp}(a_1)$ and $\text{tp}(a_2)$ in $\mathfrak{A}$ such that $\mathfrak{A} \models a'_1R^+a'_2$. We set in this case $\text{tp}_{\mathfrak{A}'}(a_1, a_2) = \text{tp}_{\mathfrak{A}}(a'_1, a'_2)$ (and proceed analogously in the symmetric case). In the opposite case there is no R^+-path in $\mathfrak{A}'$ between a_1 and a_2 and we can safely set $\text{tp}_{\mathfrak{A}'}(a_1, a_2) = \text{tp}_{\mathfrak{A}}(a_1, a_2)$.

▶ **Proposition 13.** Let $\mathfrak{A}$ be a model of an $\forall^2_{TC}[1]$ formula φ. Then there exists a model $\mathfrak{A}' \models \varphi$ in which all zones are either single cliques or infinite, regular chains of cliques, with regular connections among zones.

Proof. We simply repeat the described procedure successively to all zones, obtaining finally a model of a desired shape. ◀

6.4 Decidability procedure

A structure of a shape as in Proposition 13 can be described in a natural way. Such a description contains for every zone a sequence of clique types guaranteed by Proposition 12, patterns of connections among them, and for every pair of zones a pattern of connection between every clique type from the first zone and every clique type from the second zone.

To check if a given formula φ in $\forall^2_{TC}[1]$ has a model we guess such a description of a regular model. Verifying that a guessed description indeed produces a model of φ is easy and can be done in polynomial time with respect to its size. As the number of zones is bounded doubly exponentially in the size of the signature, and thus also in $|\varphi|$, the whole description of a regular structure is also bounded doubly exponentially. Thus we obtain:

▶ **Theorem 14.** *The satisfiability problem for $\forall^2_{TC}[1]$ is decidable in 2-*NExpTime.

7 NExpTime-upper bound for formulas with three constants

In this section we show that $\forall^2_{TC}[1,3]$, even though it lacks a finite model property, is still decidable in NExpTime. For a given structure $\mathfrak{A}$ we say that a sequence $V_1, \ldots, V_k$ of zones is a *path of zones* if for each i there exist $v_i \in V_i, v_{i+1} \in V_{i+1}$ such that $\mathfrak{A} \models v_i R v_{i+1}$. Note that in this case, in models guaranteed by Proposition 13 a path from each element of V_i to each element of V_j exists for $i < j$.

▶ **Definition 15.** Let $\mathfrak{A}$ be a model with regular zones as in Proposition 13. We say that $\mathfrak{A}$ is *downward fork-like* if it consists of four zones $V_0, \ldots, V_3$, containing all constants, and some number of zones forming a path from V_1 to V_0, a path from V_0 to V_2, and a path from V_0 to V_3. Similarly $\mathfrak{A}$ is *upward fork-like* if it consists of four zones $V_0, \ldots, V_3$, containing all constants, and some number of zones forming a path from V_0 to V_1, a path from V_2 to V_0, and a path from V_3 to V_0. A structure is *fork-like* if it is downward or upward fork-like. Zone V_0 is called a *splitting zone* of the structure. We start from the following observation.

▶ **Lemma 16.** *If an $\forall^2_{TC}[1,3]$ formula φ has a fork-like model $\mathfrak{A}$ then it has a fork-like model in which the number of zones is bounded exponentially in $|\varphi|$.*

Proof. We show a proof for the case in which $\mathfrak{A}$ is downward fork-like. The case of an upward fork-like structure is analogous. Let $\mathfrak{A}'$ be the R-saturation of $\mathfrak{A}$. Note that R-saturation does not change the division into cliques and zones. Let $V_0, \ldots, V_3$ be as in Definition 15. Let π_{10} be some shortest path of zones from V_1 to V_0, π_{02} a shortest path of zones from V_0 to V_2 and π_{03} a shortest paths of zones from V_0 to V_3. Note that $\mathfrak{A}'' = \mathfrak{A}' \restriction \pi_{10} \cup \pi_{02} \cup \pi_{03}$ is still a model of φ. By R-saturation of $\mathfrak{A}'$ and an argument similar to the one used in the proof of Lemma 5 the number of zones in $\mathfrak{A}''$ is bounded exponentially in the size of φ. ◀

Our plan is to show that every satisfiable formula φ in $\forall^2_{TC}[1,3]$ has either a finite, exponentially bounded model, or a fork-like model.

Let $\mathfrak{A} \models \varphi$ be a regular model guaranteed by Proposition 13. Let c_1, c_2, c_3 be the elements satisfying K_1, K_2, K_3, resp.

Simple cases.

- If two of c_1, c_2, c_3 belong to the same zone, then they belong to the same clique. In this case we may construct a finite model as in Section 4.
- If one of the constants, say c_3 is in free position with both the remaining constants, then we construct a model consisting of the cliques of c_1 and c_2, a path between them, if such a path exists, and the clique of c_3. The path between c_1 and c_2 can be then shortened to an exponential length as in Section 4.
- If there exists a path from one of the constants to another, containing the third one then again we may use the construction from Section 4 to obtain a path of exponential size.

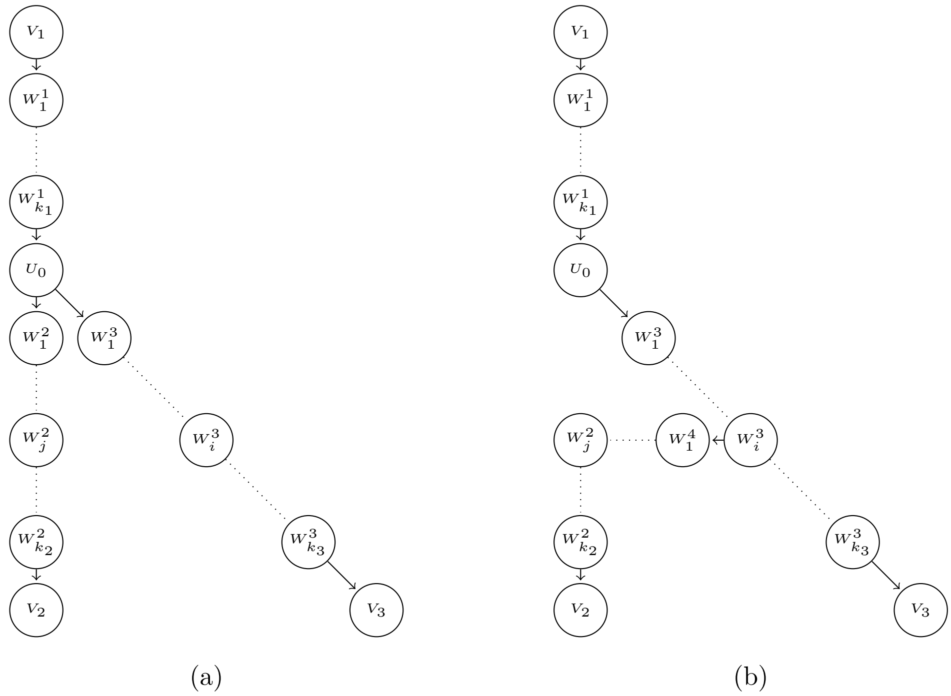

Figure 4 Fork-like structures $\mathfrak{A}_0$ and $\mathfrak{A}_1$ from the proof.

As demonstrated, in each of the above cases there exists a finite, exponentially bounded model of φ.

Fork-like case. A more interesting case is when the constants belong to three distinct zones, two of them, say c_1, c_2 are in free position, and the third one, c_3, can reach both c_1 and c_2 by R-paths, or, symmetrically, can be reached from both c_1 and c_2 by R-paths. Assume, e.g., that $\mathfrak{A} \models c_1R^+c_2 \wedge c_1R^+c_3 \wedge \neg c_2R^+c_3 \wedge \neg c_3R^+c_2$. Let V_1, V_2, V_3 be the zones of c_1, c_2, resp. c_3. Let π_{12} be a path of zones $V_1, W_1^1, \ldots, W_{k_1}^1, U_0, W_1^2, \ldots, W_{k_2}^2, V_2$ from V_1 to V_2, with U_0 being a zone from which a path to V_3 exists. Let π_{03} be a path of zones $U_0, W_1^3, W_2^3, \ldots W_{k_3}^3, V_3$, from U_0 to V_3. See Fig. 4(a).

Note that $\mathfrak{A}_0 \restriction \pi_{12} \cup \pi_{03}$ is a downward fork-like structure, splitting at zone U_0. Observe also that in $\mathfrak{A}_0$ the formula φ cannot be violated by a pair of elements, such that one of them belongs to the fragment $V_1, \ldots, U_0$ of π_{12}. However, it is not necessarily the case that $\mathfrak{A}_0 \models \varphi$, as some elements belonging to zones located below U_0 may be required to be connected by R-paths. Assume e.g. that an element from W_i^3 is connected in $\mathfrak{A}$ to an element in W_j^2. Let $W_i^3, W_1^4, \ldots, W_{k_4}^4, W_j^2$ be a path of zones. Observe now that the structure $\mathfrak{A}'$ consisting of the path of zones $V_1, W_1^1, \ldots, W_{k_1}^1, U_0, W_1^3, \ldots, W_{i-1}^3, W_i^3, W_1^4, \ldots, W_{k_4}^4, W_j^2, W_{j+1}^2, \ldots, W_{k_2}^2, V_2$, and the path $W_{i+1}^3, \ldots, W_{k_3}^3, V_3$ is a downward fork-like structure splitting at zone W_i^3. Denote $U_1 = W_i^3$, and observe that U_1 is located below U_0. See Fig. 4(b). If $\mathfrak{A}_1$ is still not a model of φ we repeat the above step obtaining a fork-like structure $\mathfrak{A}_2$, splitting at U_2, such that U_2 is located below U_1, and so on. Thus a descending sequence of zones $U_0, U_1, \ldots$ is formed, and as the number of zones is finite this process must eventually end in a structure which is a model of φ.

The described construction, together with Lemma 16 allows us to state:

▶ **Theorem 17.** *The satisfiability problem for* $\forall^2_{TC}[1,3]$ *is in* NEXPTIME.

8 Related undecidability results

To complete the picture we observe that the decidable fragment we have identified is very close to the border between decidability and undecidability. Namely, we show that adding a third variable or the transitive closure of a second binary symbol lead to undecidability.

▶ **Theorem 18.** *The satisfiability and the finite satisfiability problems for* $\forall^3_{TC}[1]$ *and* $\forall^2_{TC}[2]$ *are undecidable.*

The proof for $\forall^3_{TC}[1]$ can be obtained by a slight refinement of the proof of Corollary 9 from [7], which states that $\forall^4_{TC}[1]$ is undecidable. In that proof a snake-like path from the upper-left corner to the lower-right corner of the grid is enforced. Additional R-edges, necessary to define vertical adjacency relation, are enforced by a *completing squares* formula with four variables. If we allow for additional diagonal R-edges then this formula can be replaced by a *completing triangles* formula with three variables. We also remark that this proof requires only a single constant: to mark the upper-left corner of the grid. We require this constant to lie on a cycle and accept an incoming edge only from the opposite corner of the grid.

The proof for $\forall^2_{TC}[2]$ can be obtained by an adaptation of the proof of the undecidability of FO^2 with two transitive relations from [10]. This adaptation uses similar ideas to the proof of the 2-ExpTime-lower bound for $\forall^2_{TC}[1,4]$ from Section 5: appropriate neighbors of elements of the grid, which in the proof from [10] are enforced explicitly by formulas with existential quantifiers in our case can be enforced by requiring that some elements have paths to or from some constants, and by appropriate restriction of of 1-types which may be related by R-edges.

9 Conclusions

We have identified an interesting decidable fragment of two-variable logic with transitive closure operator, $\forall^2_{TC}[1]$. This fragment, even though does not allow explicitly for existential quantifiers, is sufficiently strong to admit infinity axioms and encodings of alternating Turing machines with exponentially bounded space.

Regarding the influence of the number of constants k on the finite model property and the complexity of $\forall^2_{TC}[1,k]$ we have drawn the following picture.

complexity:	NExpTime-complete			between 2-ExpTime and 2-NExpTime					
finite model property:	yes			no (infinity axioms)					
# of constants:	0	1	2	3	4	$\cdots$	l	$\cdots$	unbounded

In fact, our construction of a regular model $\mathfrak{A}'$ of φ from its arbitrary model $\mathfrak{A}$ retains more properties than those expressible in $\forall^2_{TC}[1]$. In particular $\mathfrak{A}'$ realizes only clique types realized in $\mathfrak{A}$. Thus we may add for free to our language existential statements of the form $\forall x(\chi_1(x) \rightarrow \exists y(xR^+y \wedge \chi_2(y)))$ or $\forall x(\chi_1(x) \rightarrow \exists y(yR^+x \wedge \chi_2(y)))$, with χ_1, χ_2 quantifier-free.

Without major difficulties it is possible to extend our construction even to a more expressive logic, namely to the fragment of FO^2 with the transitive closure of relation R, with the only restriction that existential subformulas are of the form $\exists y(xR^+y \wedge \psi(x,y))$, $\exists y(yR^+x \wedge \psi(x,y))$ (or formulas obtained by switching the role of x and y). In other words, existential quantifiers are *guarded* by atomic predicates built from R^+.

An important open question arises:

▶ **Open Question 1.** Is the whole two-variable fragment of first-order logic, FO^2, decidable when extended by transitive closure of a fixed binary relation?

In a recent paper [17] it is shown that FO^2 is decidable with one transitive relation. We believe that combining the techniques from that paper with some ideas from our paper may lead to a positive answer to the given open question.

We also leave two open question regarding $\forall^2_{TC}[1]$:

▶ **Open Question 2.** What is the exact complexity of the satisfiability problem for $\forall^2_{TC}[1]$?

▶ **Open Question 3.** Is the finite satisfiability problem for $\forall^2_{TC}[1]$ decidable?

References

1 A. K. Chandra, D. Kozen, and L. J. Stockmeyer. Alternation. *Journal of the ACM*, 28(1):114–133, 1981.

2 W. Charatonik and P. Witkowski. On the complexity of the Bernays-Schönfinkel class with datalog. In *LPAR*, volume 6397 of *LNCS*, pages 187–201, 2010.

3 M. J. Fisher and R. E. Ladner. Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences*, 18(2):194–211, 1979.

4 Martin Fürer. The computational complexity of the unconstrained limited domino problem (with implications for logical decision problems). In *Logic and Machines*, volume 171 of *LNCS*, pages 312–319, 1983.

5 E. Grädel, P. Kolaitis, and M. Vardi. On the decision problem for two-variable first-order logic. *Bulletin of Symbolic Logic*, 3(1):53–69, 1997.

6 E. Grädel, M. Otto, and E. Rosen. Undecidability results on two-variable logic. In *STACS*, volume 1200 of *LNCS*, pages 249–260, 1997.

7 N. Immerman, A. Moshe Rabinovich, T. W. Reps, S. Sagiv, and G. Yorsh. The boundary between decidability and undecidability for transitive-closure logics. In *CSL*, volume 3210 of *LNCS*, pages 160–174, 2004.

8 A.S. Kahr, E.F. Moore, and H. Wang. Entscheidungsproblem reduced to the ∀∃∀ case. 1962.

9 E. Kieroński. The two-variable guarded fragment with transitive guards is 2ExpTime-Hard. In *Proc. Foundations of Software Science and Computational Structures, 6th International Conference, FOSSACS*, volume 2620 of *LNCS*, pages 299–312, 2003.

10 E. Kieroński. Results on the guarded fragment with equivalence or transitive relations. In *CSL*, volume 3634 of *LNCS*, pages 309–324, 2005.

11 E. Kieroński, J. Michaliszyn, I. Pratt-Hartmann, and L. Tendera. Two-variable first-order logic with equivalence closure. *Accepted for LICS*, 2012.

12 E. Kieroński and L. Tendera. On finite satisfiability of the guarded fragment with equivalence or transitive guards. In *LPAR*, volume 4790 of *LNAI*, pages 318–332, 2007.

13 H. R. Lewis. Complexity results for classes of quantificational formulas. *Journal of Computer and System Sciences*, 21:317–353, 1980.

14 J. Michaliszyn. Decidability of the guarded fragment with the transitive closure. In *ICALP (2)*, volume 5556 of *LNCS*, pages 261–272, 2009.

15 M. Mortimer. On languages with two variables. *Zeitschr. f. Logik und Grundlagen d. Math.*, 21:135–140, 1975.

16 W. Szwast and L. Tendera. On the decision problem for the guarded fragment with transitivity. In *LICS*, pages 147–156, 2001.

17 W. Szwast and L. Tendera. FO^2 with one transitive relation is decidable. *Unpublished*, 2012.

A Proof of Claim 3

We argue that the structure $\mathfrak{A}'$ is indeed a model of φ. Let $a, b \in A'$. Clearly, if a, b are members of the same clique they cannot violate φ, as they are either members of the same clique in $\mathfrak{A}$, and in this case they are connected exactly as in $\mathfrak{A}$, or they are members of C' and then $\varphi[a, b]$ holds by Lemma 2.

Assume that $\mathfrak{A}' \models aRb \wedge \neg bRa$. Consider the case when $a, b \in A \setminus C$. Let $\pi = a, a_0, a_1, \ldots, a_l, b$ be an R-path connecting a and b in $\mathfrak{A}$. If π does not visit C then it also exists in $\mathfrak{A}'$. If $a_s, a_{s+1}, \ldots, a_t$ is the maximal fragment of π contained in C then there exists a path π_1 in $\mathfrak{A}'$ from a to an element $a'_s \in C'$ and a path π_2 from an element $a'_t \in \mathfrak{A}$ to b. Concatenating π_1 with a path from a'_s to a'_t and with π_2 we get a path from a to b in $\mathfrak{A}$. Thus a and b are connected in $\mathfrak{A}'$ exactly as they are connected in $\mathfrak{A}$. In case when $a \in A \setminus C$ and $b \in C'$ it can be shown that there exists a path from a to an element $b' \in \mathfrak{C}$ of 1-type equal to the 1-type of b. In this case a and b are connected in $\mathfrak{A}'$ exactly as a and b' are connected in $\mathfrak{A}$. Analogously in the symmetric case when $a \in C'$ and $b \in A \setminus C$.

If $\mathfrak{A}' \models \neg aRb \wedge \neg bRa$, then either $a, b \notin C'$ and in this case they are also not R-connected in $\mathfrak{A}$, or one of them, say a is in C' and the other, $b \notin C'$. In this case there is no R-path from b to C in $\mathfrak{A}$, and thus a and b are connected in $\mathfrak{A}'$ exactly as some $a_0 \in C$ and b in $\mathfrak{A}$.

B Simulating alternating Turing machines

Let M be an alternating Turing machine and w its input. Let $n = |w|$, and assume that M works on w using 2^n tape cells. Without loss of generality we assume that M has exactly two possible transitions in every configuration, enters an accepting or a rejecting state exactly in step 2^{2^n} and that after this step it works further in a trivial way, just staying infinitely in the same configuration.

For each symbol a_i from the alphabet of M we use unary predicates A_i, A_i^L, A_i^R, for each state s_i of M we use unary predicates S_i, S_i^L, S_i^R. A single element a of a structure, satisfying $\bar{P}(a) = k$ is going to contain information about the cell k of the tape in a configuration of M, and about its neighboring cells. The intended meanings of $A_i(a), A_i^L(a), A_i^R(a), S_i(a), S_i^L(a), S_i^R(a)$ are respectively: cell k contains a_i, cell $k+1$ contains a_i, cell $k-1$ contains a_i, cell k is observed by the head and M is in state s_i, cell $k+1$ is observed by the head and M is in state s_i, cell $k-1$ is observed by the head and M is in state s_i.

Let $\kappa_w^M = \lambda \wedge \kappa_I \wedge \kappa_C \wedge \kappa_T \wedge \kappa_R$, where λ is as in the previous subsection, and:

- κ_I says that each node of type H_0^L satisfying H^I represents the initial configuration of M on w,
- κ_C says that the information about the alphabet symbols, states and the head of M kept in an element of the node is consistent with the information kept in the neighboring nodes,
- κ_T takes care of transitions of M; consider a node representing a configuration γ; if γ is universal then we say that every left successor node represents the configuration obtained by the first possible move in γ, and every right successor node represents the configuration obtained by the second possible move in γ; if γ is existential then we say that every left and right successor nodes represent configurations which can be obtained from γ by one of the possible moves (possibly the same move for both successors).
- κ_R says that none of the elements encodes a rejecting state.

Now we sketch an argument for the fact that κ_w^M is satisfiable iff M has an accepting run on w. If κ_w^M has a model then by Claim 8 this model contains an initial node and every node has two successor nodes. κ_I, κ_C, κ_T and κ_R allow now to derive an accepting run of M from $\mathfrak{A}$. If M has an accepting run of w then we construct a model of κ_w^M by expanding the structure $\mathfrak{T}$ in such a way that the initial node of $\mathfrak{T}$ represents the initial configuration of M on w, the left (right) successor of a node representing a universal configuration represents the configuration which can be obtained by applying the first (second) move of M, and both successors of a node representing an existential configuration represent the configuration which is obtained the move of M in an accepting run. An important property of $\mathfrak{T}$ is that only nodes from two consecutive levels are R-connected, so κ_T does not impose restrictions on distant nodes.

C Proof of Lemma 9

Let $\varphi = \forall xy.\varphi_0$ be a $\forall^2_{TC}[1,n]$. Formula φ' will have the same signature as φ, except for unary symbols $K_5, \ldots, K_n$, but extended by fresh unary symbols $S_1, \ldots, S_n$ that simulate constants, and auxiliary predicates $P_1, \ldots, P_n$, $Q_1, \ldots, Q_n$, and t. We define $\varphi' = \forall xy.\psi \wedge (t(x) \wedge t(y) \Rightarrow \varphi_0^*)$, where φ_0^* is the result of replacing all occurrences of K_i in φ_0 by S_i and ψ is the conjunction of the formulas that express the following properties.

(1) Every element satisfies at most one of $P_1, \ldots, P_n, Q_1, \ldots, Q_n, S_1, \ldots, S_n$.
(2) For each i, there is at most one element satisfying P_i, at most one satisfying Q_i, and at most one satisfying S_i.
(3) c_1 satisfies P_1, c_2 satisfies P_n, c_3 satisfies Q_1 and c_4 satisfies Q_n.
(4) There is an R-path from c_1 to c_2 and from c_3 to c_4.
(5) For each i, all R-predecessors of a vertex satisfying P_i satisfy P_{i-1}.
(6) For each i all R-successors of a vertex satisfying Q_i satisfy Q_{i+1}.
(7) All R-successors of a vertex that satisfy P_i satisfy P_{i+1} or S_i.
(8) All R-predecessors of a vertex that satisfy Q_i satisfy Q_{i-1} or S_i.
(9) For each i, j, there is a path from the vertex satisfying P_i to the vertex satisfying Q_j if and only if $j \geq i$.
(10) All constants satisfy $\neg t$.
(11) Each vertex that satisfies S_i satisfies also t.
(12) A vertex v that satisfies t can be connected with a vertex that does not satisfy t only if v satisfies S_i for some i.

A quick check shows that all these properties can be expressed in $\forall^2_{TC}[1,4]$. It can be shown that the model presented in Fig. 3 is minimal (w.r.t. number of vertices), i.e. that each model of $\forall xy\psi$ contains, for all i, at least one point satisfying P_i, at least one satisfying Q_i, and at least one satisfying S_i.

Consider any model of $\forall xy\psi$. Conjunct (4) guarantees that there is a path from c_1 to c_2, by (3), c_1 satisfies P_1 and c_2 satisfies P_n, and by (6) such a path must contains vertices satisfying $P_2, P_3, \ldots, P_{n-1}$. The same holds for c_3, c_4 and $Q_2, \ldots, Q_{n-1}$.

Consider any $i \leq n$. By (9), there has to be a path from (the element satisfying) P_i to Q_i. By (8), the predecessor of Q_i on this path has to satisfy Q_{i-1} or S_i. But if it satisfied Q_{i-1}, then there would be a path from P_i to Q_{i-1}, contradicting (9). Therefore the model contains a vertex that satisfies S_i.

If φ' has a model $\mathfrak{M}$, then the substructure $\mathfrak{M}^t$ of $\mathfrak{M}$ induced by all elements satisfying t satisfies $\forall xy\varphi_0^*$. We define a structure $\mathfrak{M}'$ by replacing in $\mathfrak{M}^t$ all S_i by K_i and erasing auxiliary predicates. Clearly, $\mathfrak{M}'$ is a model of φ and, by (9), in $\mathfrak{M}'$ for all $i < j$ there is no R-path from c_j to c_i.

Assume φ has a model $\mathfrak{M}$ such that for all $i < j$ there is no R-path from c_j to c_i in $\mathfrak{M}$. We build a model $\mathfrak{M}'$ of φ' as follows. We make all vertices of $\mathfrak{M}$ satisfy t, and rename predicates K_i to S_i for all i. Then, we add the structure presented in Fig. 3 by identifying c_i in $\mathfrak{M}$ with S_i. It is not hard to see that $\mathfrak{M}'$ satisfies φ'.