On Logics with Two Variables

Erich Grädel[*] Martin Otto[†]

Abstract

This paper is a survey and systematic presentation of decidability and complexity issues for modal and non-modal two-variable logics.

A classical result due to Mortimer says that the two-variable fragment of first-order logic, denoted FO^2, has the finite model property and is therefore decidable for satisfiability. One of the reasons for the significance of this result is that many propositional modal logics can be embedded into FO^2.

Logics that are of interest for knowledge representation, for the specification and verification of concurrent systems and for other areas of computer science are often defined (or can be viewed) as extensions of modal logics by features like counting constructs, path quantifiers, transitive closure operators, least and greatest fixed points etc. Examples of such logics are computation tree logic CTL, the modal μ-calculus L_μ, or popular description logics used in artificial intelligence. Although the additional features are usually not first-order constructs, the resulting logics can still be seen as two-variable logics that are embedded in suitable extensions of FO^2. Typically, the applications call for an analysis of the satisfiability and model checking problems of the logics employed.

The decidability and complexity issues for modal and non-modal two-variables logics have been studied quite intensively in the last years. It has turned out that the satisfiability problems for two-variable logics with full first-order quantification are usually much harder (and indeed highly undecidable in many cases) than the satisfiability problems for corresponding modal logics. On the other side, the situation is different for model checking problems. The model checking problem of a modal logic has essentially the same complexity as the model checking problem of the corresponding two variable logic with full quantification.

Keywords: Two-variable logics, modal logics, satisfiability, model-checking, decidability

[*] Mathematische Grundlagen der Informatik, RWTH Aachen, D-52056 Aachen, Email: graedel@informatik.rwth-aachen.de

[†] Mathematische Grundlagen der Informatik, RWTH Aachen, D-52056 Aachen, currently UC Santa Cruz and Stanford University, USA. Email: otto@informatik.rwth-aachen.de

1 Introduction

Two-variable logics, more often than not in the disguise of modal logics, are important in many branches of computer science including the specification and verification of concurrent processes, reasoning about knowledge, artificial intelligence etc. Indeed, propositional modal logics and their extensions by mechanisms like path quantifiers, transitive closure operators, least and greatest fixed points or counting constructs have been studied with great success and have lead to interesting, manageable languages, that meet the essential expressive needs for certain applications. Typically, these applications call for an analysis of the satisfiability problem of the logics employed and/or of the complexity of their model checking problems.

The two-variable nature of such languages may be attributed to the fact that they contain propositional modal logic as an essential core – reflecting the eminent role that modal operators (which are viewed here as a restricted form of quantification) tend to play in the above-mentioned areas of applications. For many applications, however, first-order closure properties (i.e. unrestricted quantification over element variables) offer an equally desirable direction for extensions. In this case, the embedding of propositional modal logic into FO^2, i.e. relational first-order logic with two variables, provides the natural starting point.

In this survey we present an overview of recent results pertaining to logics that lift some of the prominent mechanisms of extension from the modal framework to the framework of two-variable first-order logic. The most striking overall result is that, with respect to the satisfiability problem, two-variable first-order logic turns out to be not nearly as robust as modal logic. Several seemingly weak extensions of FO^2 in important directions turn out to be highly undecidable. One notable exception is the extension by counting quantifiers, which does provide a decidable common extension of graded modal logic and two-variable first order logic.

Plan of the paper. In the remainder of this first section, we discuss a number of two-variable properties and introduce the modal and non-modal two-variable logics that we are going to study. Further, we will make precise the notion of a two-variable logic, based on purely semantic game-theoretic criteria.

In section 2 we survey and explain decidability results for satisfiability in two-variable logics. We put the emphasis on the difficult cases, namely two-variable logics with full quantification. We present relevant techniques that are used for proving such results and discuss the decidability proofs for FO^2 and C^2 (the extension of FO^2 by counting quantifiers).

In section 3 however, we will see that, in some sense, C^2 is an exception. For most of the other natural common extensions of FO^2 and a modal language like CTL or L_μ, the satisfiability problem is undecidable in a very strong sense. We will also relate these problems to the study of the FO^2-theories of certain interesting model classes, for instance of the class of structures with built-in equivalence relations or built-in well-orderings. We show that already rather modest built-in predicates lead to undecidable FO^2-theories.

Finally, in section 4 we discuss the complexity of model checking for both modal and non-modal two-variables logics. It turns out that we have here a very different situation than for satisfiability problems. In all cases that we consider the model checking problem of a modal logic has essentially the same complexity as the model checking problem of the corresponding two variable logic with full quantification.

1.1 Examples of two-variable phenomena and two-variable logics

1.1.1 Typical two-variable properties

The following examples are all phrased as monadic queries. In other words, each of them concerns the problem to determine some property of a single element in a given relational structure. To choose a uniform and simple common setting for the structures in the examples, think of directed graphs with two different edge predicates, E_1 and E_2, and with two monadic predicates P_1 and P_2. Let us write $\mathfrak{A} = (A, E_1, E_2, P_1, P_2)$ for a typical structure of this format, where A is the universe of $\mathfrak{A}$. Where we really only need one of the E_i or of the P_i, we drop the indices and consider e.g. structures $\mathfrak{A} = (A, E, P_1, P_2)$ or $\mathfrak{A} = (A, E, P)$. Input $(\mathfrak{A}, a)$, where $a \in A$, as an instance for one of the sample properties $\mathcal{Q}$ is the problem to determine whether a has that property in $\mathfrak{A}$. The corresponding monadic query is the class Q of instances $(\mathfrak{A}, a)$ where a in $\mathfrak{A}$ does have property $\mathcal{Q}$.

Example 1.1 (a) Does a have an outgoing E_1-edge to a vertex in P?

(b) Do all vertices that can be reached from a by traversing one E_2-edge have the property described in (a)?

(c) Is there an E-path of length 17 from a, which ends in a vertex in P?

Example 1.2 (a) Is there an incoming E-edge at a?

(b) Is there a vertex linked to a by both an E_1-edge and an E_2 edge?

(c) Is every vertex in P reachable from a on an E-path of length at most 3?

Example 1.3 (a) Is it possible to reach P from a on an E-path?

(b) Do all E-paths from a eventually hit P_1, and before hitting P_1 only pass through vertices in P_2.

Example 1.4 Think of $\mathfrak{A} = (A, E)$ as the board for a two-person game, in which players move a single pebble on A according to the following rules. Players take alternate moves, player **I** begins. In their moves, players move the pebble from its current position along some E-edge. Who gets stuck first, loses the game (the opponent wins).

(a) Does player **I** have a winning strategy in the game on $\mathfrak{A}$ if the pebble is initially placed on a?

(b) Is the game on $(\mathfrak{A}, a)$ necessarily finite?

The GAME-problem 1.4 (a) is well known for being PTIME-complete. Property (b) is of independent conceptual interest as it concerns the well-foundedness of the converse E^{-1} of E at a.

Example 1.5 (a) Is E deterministic, i.e. are E-successors at all vertices unique?

(b) Is $\mathfrak{A}$ (isomorphic with) the full binary tree with left and right successors E_1 and E_2?

(c) Can you be sure to reach P from a on an E-path of length at most 3, if some adversary may block one E-edge in each step?

(d) In the GAME-problem of Example 1.4, is the strategy for player **I** unique?

However different these problems are, one thing they all have in common is that to determine these properties one would in principle *never have to investigate more than two vertices of the underlying structure at the same time* – provided that, for some of these properties, one can keep on the side certain records of auxiliary properties or of numbers of pairs already inspected, et cetera. This is something that distinguishes these properties crucially from a property like for instance triangle-freeness of a graph, which would intuitively require inspection of triples of vertices. We will show below that the properties in the above examples are definable in natural two-variable logics. We shall also see that there are purely semantic criteria to prove that these are two-variable properties, in a sense that makes the remark about only checking pairs of vertices at any one time precise.

1.1.2 Some typical two-variable logics

We review and introduce a number of prominent two-variable logics, and, by way of indicating their expressive power, apply them to the formalization of those two-variable properties given in the examples. For the most part we may assume we are dealing with a standard relational vocabulary consisting of binary predicates E_1, E_2, ... , and monadic predicates P_1, P_2, ...

Modal logic ML Think of structures $\mathfrak{A} = (A, E_1, \dots, P_1, \dots)$ as *Kripke structures*, regarding their elements as *possible worlds*, the binary E_i as *accessibility relations*, and the monadic P_i as *basic propositions*. The syntax and semantics of (propositional) modal logic ML concerns formulae φ asserting a property of worlds a in some $\mathfrak{A}$: $(\mathfrak{A}, a) \models \varphi$. (If we wanted to speak instead in terms of *transition system*, then elements would be called *states*, the E_i elementary *transitions, actions*, or *programs*, and formulae in one free variable would be *state formulae*). The formulae of ML are inductively generated from atomic formulae of the form P_i, $\perp$ or $\top$ through closure under

(i) Boolean connectives $\neg$, $\wedge$, $\vee$,

(ii) modal quantification: if φ is a formula then so are $\Box_i\varphi$ and $\Diamond_i\varphi$. ([1])

Semantics of ML: For an atomic formula $\varphi = P_i$, $(\mathfrak{A}, a) \models \varphi$ if $a \in P_i$. $\perp$ and $\top$ are universally false, respectively true. Boolean connectives are treated in the natural way. For the semantics of modal quantification let for $a \in A$, $E_i[a]$ be the set of direct E_i-successors of a in $\mathfrak{A}$: $E_i[a] := \{a' \in A : (a, a') \in E_i\}$. Now $(\mathfrak{A}, a) \models \Diamond_i\varphi$ if there is some $a' \in E_i[a]$ such that $(\mathfrak{A}, a') \models \varphi$. Dually, $(\mathfrak{A}, a) \models \Box_i\varphi$ if $(\mathfrak{A}, a') \models \varphi$ for all $a' \in E_i[a]$.

It is easy to check that the properties of Example 1.1 are ML-definable. The formula $\Diamond_1 P$ expresses (a); $\Box_2\Diamond_1 P$ expresses (b); (c) is expressed by $(\Diamond)^{17}P$, where $(\Diamond)^{17}$ is shorthand for 17-fold iteration of the $\Diamond$-operator.

Infinitary modal logic $\mathbf{ML}_\infty$ The infinitary variant of ML, ML_∞, is similar to ML, only that Boolean closure is extended to allow conjunctions and disjunctions over arbitrary sets of formulae. Example 1.3 (a) is expressible in ML_∞ by $\bigvee\{\varphi_n : n \in \mathbb{N}\}$ where $\varphi_n = (\Diamond)^n P$.

The modal μ-calculus L_μ L_μ extends ML with a least fixed point constructor. Starting from atomic modal formulae including atoms for propositional variables $X, Y, \ldots$, the syntax of L_μ is obtained as the closure under

(i) Boolean connectives $\neg$, $\wedge$, $\vee$,

(ii) modal quantification: if φ is a formula then so are $\Box_i\varphi$ and $\Diamond_i\varphi$,

(iii) least fixed points: if $\varphi(X)$ is a formula that is positive in X, then $\mu_X\varphi$ is a formula.

Semantics of L_μ: The semantics naturally extends that of ML, with the following stipulations for the μ-operator. Over $\mathfrak{A}$, the set $[\mu_X\varphi]^{\mathfrak{A}} := \{a : (\mathfrak{A}, a) \models \mu_X\varphi\}$ is the least subset S of the universe satisfying the fixed-point equation $S = \{a : (\mathfrak{A}, a) \models \varphi(S)\}$. Equivalently, this fixed point may be obtained as the limit of an inductively defined monotone sequence of subsets S_α indexed by ordinals α. The S_α are known as the *stages* in the generation of the fixed point: $S_0 = \varnothing$, $S_{\alpha+1} = \{a : (\mathfrak{A}, a) \models \varphi(S_\alpha)\}$, and $S_\lambda = \bigcup_{\alpha<\lambda} S_\alpha$ in limits λ. Then $[\mu_X\varphi]^{\mathfrak{A}} = \bigcup_\alpha S_\alpha = S_{\alpha_0}$, for α_0 the least ordinal such that $S_{\alpha_0+1} = S_{\alpha_0}$ (note that α_0, the *closure ordinal* of the fixed point, depends on the underlying structure $\mathfrak{A}$).

[1] If there is only one accessibility E, one simply writes $\Box$ and $\Diamond$.

One could additionally introduce greatest fixed points without increasing the expressive power, since least and greatest fixed points are related by a straightforward duality.

L_μ is *essentially* contained in ML_∞, in the sense that for every cardinality κ there is a translation of L_μ-formulae into ML_∞ that is sound over all structures whose cardinality is bounded by κ. The reason for this is that the stages of a fixed point can inductively be shown to be ML_∞-definable. Once a uniform bound γ on the closure ordinal is known, the fixed point itself is ML_∞-definable, just as stage S_γ. Generally, and over arbitrarily large structures no such bound is available (and the formal disjunction over the defining formulae for *all* S_α is not admitted even in ML_∞, because they do not form a set).

For some concrete examples of expressibility in L_μ consider the properties in Examples 1.3 and 1.4. The existence of a strategy for player **I** in Example 1.4 (a) is expressed by the formula $\mu_X \Diamond\Box X$. Indeed, **I** playing from b has a strategy to win in one move if $(\mathfrak{A}, b) \models \Diamond\Box\bot$. The set of such b is precisely the first stage w.r.t. the fixed point $\mu_X \Diamond\Box X$. Inductively, **I** has a strategy to win from b in $n+1$ moves if there is a move for **I** such that, no matter which counter-move **II** chooses, **I**'s next move is made from a position in which **I** has a strategy to win in n moves; the $\Box\Diamond$-construct precisely captures this, in the transition from stage n to stage $n+1$. For Example 1.4 (b), note that the game is necessarily finite if and only if there is no infinite E-path from a, which is expressed by $(\mathfrak{A}, a) \models \mu_X \Box X$.

Two-variable first-order $\mathbf{FO^2}$ and its infinitary variant $\mathbf{L^2_{\infty\omega}}$ Two-variable first-order logic FO^2 is just the fragment of ordinary first-order FO (with equality), whose formulae only use variable symbols x and y (free or bound). It is easy to express the properties of Example 1.2 in FO^2.

As for modal logic, we also consider an infinitary variant $L^2_{\infty\omega}$, which is closed under conjunctions and disjunctions over sets of formulae.

Modal versus first-order Note that ML may be embedded into FO^2 (and ML_∞ into $L^2_{\infty\omega}$) through an inductive translation $\varphi \mapsto \varphi^*(x)$ according to

$$\begin{array}{rcl} (P_i)^*(x) & = & P_i x, \\ (\Diamond_i \varphi)^*(x) & = & \exists y \big(E_i xy \wedge \varphi^*(y) \big), \\ (\Box_i \varphi)^*(x) & = & \forall y \big(E_i xy \rightarrow \varphi^*(y) \big), \end{array}$$

where $\varphi^*(y)$ is the result of exchanging variables x and y throughout $\varphi^*(x)$. Indeed, modal quantification may be regarded as a relativized FO^2-quantification, namely relativized to the sets $E_i[x]$.

W.r.t. expressive power, ML is a proper fragment of FO^2: $ML \subsetneq FO^2$, and the strictness of this inclusion can be pinned to a number of different restrictions of ML:

– ML does not have equality.
– ML does not have global quantification (Example 1.2 (c) is not in ML).
– ML does not have mechanisms for defining new binary predicates, not even at the quantifier-free level (Examples 1.2 (a) and (b) are not in ML).

The third point is of particular interest. One might consider logics with modal quantification extended to definable accessibilities. For a formula $\xi(x,y)$ one may consider new modal operators with respect to the accessibility relation defined by ξ, $[\xi]$ and $\langle\xi\rangle$. If ξ is in FO^2 then these generalized modal operators can be defined within FO^2, as for instance in $\big([\xi]\varphi\big)(x) \equiv \forall y\big(\xi(x,y) \rightarrow \varphi(y)\big)$. An interesting feature of plain ML is, however, that it is closed under one particular construction of new accessibilities, which even lies outside FO^2: let $E_1 \circ E_2$ be the composition of the E_i; then $[E_1 \circ E_2]\varphi$ is equivalent with $\Box_1\Box_2\varphi$, while $E_1 \circ E_2$ itself (as a binary relation!) is not definable in FO^{2}from the E_i.

Actually FO^2 (and $L^2_{\infty\omega}$) still have rather limited power for defining new binary predicates. Indeed it is not hard to show inductively that any FO^2-formula $\xi(x,y)$ is equivalent with a Boolean combination of quantifier-free FO^2-formulae $\theta(x,y)$ and FO^2-formulae in one free variable $\theta_1(x)$ and $\theta_2(y)$.

Variants that can count: graded ML, C^2, and their infinitary variants It turns out that none of the properties in Example 1.5 are expressible in $L^2_{\infty\omega}$ even over finite structures. These properties involve some very basic forms of counting.

Two-variable first-order with counting C^2 extends FO^2 by allowing first-order counting quantifiers $\exists^{\leqslant m}$, $\exists^{\geqslant m}$, $\exists^{=m}$ for all $m \geqslant 1$. These quantifiers are first-order definable, but not with a restricted number of variables. Clearly $\mathrm{FO}^2 \subseteq \mathrm{C}^2 \subseteq \mathrm{FO}$, and in fact these inclusions are strict.

Infinitary two-variable logic with counting $C^2_{\infty\omega}$ is obtained through closure of C^2 under conjunctions and disjunctions over arbitrary sets of formulae. Obviously it extends both $L^2_{\infty\omega}$ and C^2.

Graded modal logic extends ML by allowing graded modal operators $\Diamond_i^{\leqslant m}$, $\Diamond_i^{\geqslant m}$, $\Diamond_i^{=m}$ for all $m \geqslant 1$, whose semantics is best indicated through the natural translation into C^2, as in $\big(\Diamond_i^{\leqslant m}\varphi\big)(x) \equiv \exists^{\leqslant m}y\big(E_ixy \wedge \varphi(y)\big)$. ML $\subsetneq$ graded ML $\subsetneq$ C^2.

Again there is the natural infinitary variant *graded* ML_∞ which simultaneously extends ML_∞ and graded ML. $\mathrm{ML}_\infty \subsetneq$ graded $\mathrm{ML}_\infty \subsetneq C^2_{\infty\omega}$.

Turning to the examples, determinism of the edge predicate E, as in Example 1.5 (a), is expressed by the C^2-sentence $\forall x\exists^{\leqslant 1}yExy$ (whose natural translation into FO would require 3 variables, as in $\forall x\forall y\forall z(Exy \wedge Exz \rightarrow y = z)$).

The complete binary tree cannot be characterized in C^2 or in $L^2_{\infty\omega}$, but in $C^2_{\infty\omega}$. For Example 1.5 (b) we may use the conjunction of the following: $\exists^{=1}x\forall y\neg(E_1yx \vee E_2yx)$, $\forall x\forall y\neg(E_1yx \wedge E_2yx)$, $\bigwedge_{i=1,2}\forall x\exists^{=1}yE_ixy$, $\forall x\exists^{\leqslant 1}y(E_1yx \vee E_2yx)$, and $\forall x\bigvee_{n\in N}\varphi_n(x)$, where $\varphi_n(x)$ is an FO^2-formula expressing that there is no descending $(E_1 \cup E_2)$-path of length n at x.

Examples 1.5 (c) and (d) are expressible in graded ML and graded ML_∞, respectively. For (c) consider the formula $P \vee \Diamond^{\geqslant 2}\big(P \vee \Diamond^{\geqslant 2}\big(P \vee \Diamond^{\geqslant 2} P\big)\big)$. (d) is captured by the disjunction over formulae $\big(\Diamond^{=1}\Box\big)^n\bot$, for $n \geqslant 1$.

1.2 A semantic notion of 'requiring only two variables'

Is there any way to specify what it means for a property to be a *two-variable property* – in some more fundamental sense? Since the notion of a property is a semantic one it would be desirable to obtain a semantic specification, one that directly deals with properties rather than with their formalizations in specific logical languages. The basic intuition is that *being two-variable* corresponds to being checkable in a successive analysis of pairs of elements from the structure, rather than requiring some larger simultaneous view of the structure. In other words, a property cannot be two-variable if it distinguishes between structures which are indistinguishable at the level of such successive analysis of pairs. Of course, the right notion of successive analysis of pairs needs to be formalized, through some a set of rules or some protocol. These considerations suggest the use of *logical games*, which are a classical tool to capture notions of similarity or indistinguishability of structures, for the desired kind of protocol.

Consider the two-variable variant of the Ehrenfeucht-Fraïssé game, i.e. the standard *two-pebble game*. This game is played by two players, **I** and **II**, on two relational structures $\mathfrak{A}$ and $\mathfrak{B}$. A current position in the game is given by two pairs of designated elements, one pair in each structure. Think of these pairs as currently pebbled by two pairs of marked pebbles, one pair for each structure. The rules of the game admit the relocation of pebbles within their structure. The basic information in a current position is the isomorphism type of the one- or two-element substructures specified by the pebbled elements. The two structures $\mathfrak{A}$ and $\mathfrak{B}$ can be proved different if **I** has a strategy to conduct the sequence of moves in the game into a position in which the currently marked two-element substructures are different. In other words, the only way to make a structural difference between $\mathfrak{A}$ and $\mathfrak{B}$ manifest, is by pinning it down to some difference at the level of currently inspected pairs of elements – which is in good agreement with the intuition of two-variable properties.

It remains to specify the set of rules by which the players may move pebbles, and this is where some interesting variations come up. These variations give rise to different flavours of the notion of two-variables.

The single round in the standard two-pebble game follows these rules:

I chooses one pebble in one of the structures and moves it to some element of that structure.

II responds by moving the corresponding pebble on the opposite structure.

If the position before the move was $(\mathfrak{A}, a_1, a_2; \mathfrak{B}, b_1, b_2)$, and if for instance **I** chose to move pebble number 2 in $\mathfrak{B}$ to b_2', and if **II** responded with a move onto a_2' in $\mathfrak{A}$, then the new position is $(\mathfrak{A}, a_1, a_2'; \mathfrak{B}, b_1, b_2')$. The game may continue as long as player **II** can maintain the following condition:

(W) the mapping $\pi\colon a_i \mapsto b_i$, is an isomorphism between $\mathfrak{A} \restriction \{a_1, a_2\}$ and $\mathfrak{B} \restriction \{b_1, b_2\}$.

Equivalently, (W) says that the pairs (a_1, a_2) and (b_1, b_2) realize the same quantifier-free types. **I** wins the game as soon as **II** violates this condition. Player **II** is said to have a *winning strategy* in the game if **II** has a strategy to maintain condition (W) indefinitely and in response to any choice of moves for **I**. **II** has a *winning strategy for i rounds* if (W) can be maintained by **II** for at least i rounds. Winning strategies for **I** are defined analogously. In all the games to be considered here it will be the case that exactly one of **I** and **II** has a winning strategy (in any given position), i.e. all these games are determined. The ability of player **II** to respond to challenges of **I** is a measure for the similarity of the underlying positions over $\mathfrak{A}$ and $\mathfrak{B}$. The ability to maintain (W) for more rounds and in response to all manoeuvres of **I** requires a higher degree of similarity of the initial positions. Indeed, the standard two-pebble game just outlined is well known to be a measure for the expressive power of two-variable first-order logic FO^2 and its infinitary variant $L^2_{\infty\omega}$. The following theorem, which we state in its special form for 2 variables, is essentially a straightforward variation on the classical Ehrenfeucht-Fraïssé Theorem for first-order logic. It can be attributed (in its more general k-variable form) and in different formulations to Barwise, Poizat, or Immerman.

Theorem 1.6 *Player* **II** *has a winning strategy for i moves in the standard two-pebble game in position $(\mathfrak{A}, a_1, a_2; \mathfrak{B}, b_1, b_2)$, if and only if (a_1, a_2) in $\mathfrak{A}$ and (b_1, b_2) in $\mathfrak{B}$ satisfy exactly the same* FO^2*-formulae of quantifier-rank at most i.* **II** *has a strategy in the (infinite) game, if and only if (a_1, a_2) in $\mathfrak{A}$ and (b_1, b_2) in $\mathfrak{B}$ satisfy exactly the same $L^2_{\infty\omega}$-formulae.*

A strengthening and a weakening of the game It is not hard to see that the essence of the game would not change if we made the task for **II** seemingly harder by requiring that in each round **II** always tell **I** in advance what the response to any move of **I** would be. Formally, for a move in which **I** would move pebble 1 in $\mathfrak{A}$ say, **II** would have to submit a function $f\colon A \to B$. If **I** now chooses a_1' for the new position of pebble 1 in $\mathfrak{A}$, then **II** is committed to putting pebble 1 in $\mathfrak{B}$ on $f(a_1')$. Equivalence of this variant with the more symmetric standard presentation follows easily from the proof of Theorem 1.6 or from an ad-hoc analysis of the notion of a winning strategy: having a winning strategy is to have an appropriate response to any challenge from **I** and w.l.o.g. one may assume that the players make optimal choices and are omniscient.

A strengthening of the notion of equivalence which is captured by the game does result, however, if **II** is required to choose a bijection f to govern the responses to the choice that **I** makes. (Of course, **II** loses immediately if A and B have different cardinality.) Intuitively, this new game checks not only that **II** has some suitable response to any challenge put forward by **I**, but moreover that the *number* of potential responses matches the numbers of corresponding challenges. In the form just sketched, this strengthening of the game is known as the *bijective pebble game*. Its original and more symmetric presentation as a *pebble game with counting*, involving intermediate set-moves, is due to Immerman and Lander [21]. The corresponding Ehrenfeucht-Fraïssé Theorem, which in the present form only applies to finite and countably infinite structures, is also due to Immerman and Lander.

Theorem 1.7 *Player* **II** *has a winning strategy for i moves in the two-pebble game with counting (or in the bijective two-pebble game) in position $(\mathfrak{A}, a_1, a_2; \mathfrak{B}, b_1, b_2)$ over countable structures $\mathfrak{A}$ and $\mathfrak{B}$, if and only if (a_1, a_2) in $\mathfrak{A}$ and (b_1, b_2) in $\mathfrak{B}$ satisfy exactly the same C^2-formulae of quantifier-rank at most i.* **II** *has a strategy in the corresponding infinite game, if and only if (a_1, a_2) and (b_1, b_2) satisfy exactly the same $C^2_{\infty\omega}$-formulae.*

Instead of restricting **I** we may relax the conditions to obtain a weaker notion of equivalence that corresponds to the logical strength of modal logic rather than first-order. Recall the main differences: modal quantification is restricted to locally accessible vertices, and at the atomic level only monadic predicates (namely the basic propositions) are made available. In view of the latter, the game may actually be thought of as a one-pebble game, and the binary predicates (accessibilities) enter in the rules for the single round which are to reflect the locality of modal quantification.

In a position $(\mathfrak{A}, a; \mathfrak{B}, b)$ **II** selects one of the accessibility relations and one of the structures, say E_i and $\mathfrak{B}$; **I** has to submit a function from E_1-accessible elements in $\mathfrak{B}$ to E_1-accessible elements in $\mathfrak{A}$, $f\colon E_i[b] \to E_i[a]$; **I** chooses $b' \in E_i[b]$, and the resulting position is $(\mathfrak{A}, f(b'); \mathfrak{B}, b')$. The corresponding winning condition (W') only requires the currently pebbled elements in both structures to satisfy exactly the same basic propositions. In its standard, more symmetric, formulation this game is well known as the *bisimulation game*, whose rules for the single round are governed by the following:

- **I** chooses one of the accessibility relations E_i and moves the pebble to an E_i-successor of the currently pebbles vertex in one of the structures.

- **II** responds by moving the pebble in the opposite structure along an E_i-edge.

The less symmetric version of the game given above, and comparison with the corresponding formulation of the standard two-pebble game, immediately show that two-pebble equivalence implies bisimulation equivalence: the functions f, which **II** submits in the two-pebble game, obviously have to respect any accessibility relations E_i if **II** is not to fail the winning condition (W) immediately. But then the restrictions of f to sets of the form $E_i[a]$ are good for the bisimulation game.

In different contexts, the idea of bisimulation equivalence was proposed by van Benthem [3] and by Milner [27]. The following modal Ehrenfeucht-Fraïssé Theorem is due to van Benthem.

Theorem 1.8 *Player* **II** *has a winning strategy for i moves in the bisimulation game in position $(\mathfrak{A}, a; \mathfrak{B}, b)$, if and only if a in $\mathfrak{A}$ and b in $\mathfrak{B}$ satisfy exactly the same* ML-*formulae of quantifier-rank at most i.* **II** *has a strategy in the corresponding infinite game, if and only if a and b satisfy exactly the same* ML_∞*-formulae.*

Of course, there is an immediate strengthening of this game to capture the discerning power of graded modal logic over countable (or indeed over countably branching) structures. Again, we need only require the functions $f\colon E_i[a] \to E_i[b]$, which **II** may submit, to be bijections. Call this game the *bisimulation game with counting*.

Taking the game protocols as natural formalizations of what it means that a property is essentially checkable in recourse to pairs of elements, we obtain a semantic characterization of two-variable properties, with an additional classification according to two main distinctions: modal vs. two-pebble, and counting vs. no counting. The maximal expressive power in the four resulting areas is represented by the infinitary logics ML_∞ graded ML_∞, $L^2_{\infty\omega}$ and $C^2_{\infty\omega}$, at least when we restrict attention to any set-theoretically bounded domain of structures[2]. Note that $C^2_{\infty\omega}$ is the most expressive among these logics, ML_∞ the least expressive, and $L^2_{\infty\omega}$ and graded ML_∞ intermediate and incomparable. We may view these logics as the basic framework for the study of two-variable properties.

$$
\begin{tikzcd}[column sep=small]
 & \text{2-pebble+ctg.} \arrow[dl, dash] \arrow[dr, dash] & & & & C^2_{\infty\omega} \arrow[dl, dash] \arrow[dr, dash] & \\
\text{bisim.+ctg.} \arrow[dr, dash] & & \text{2-pebble} \arrow[dl, dash] & & \text{graded } \mathrm{ML}_\infty \arrow[dr, dash] & & L^2_{\infty\omega} \arrow[dl, dash] \\
 & \text{bisimulation} & & & & \mathrm{ML}_\infty &
\end{tikzcd}
$$

[2]The point of this restriction is only to make sure that we are always dealing with a *set* of isomorphism types of structures rather than a proper class. In the context of actual applications this would always seem to be granted.

A classification of the basic examples according to these four flavours of two-variable properties essentially follows from the formalizations in respective standard two-variable logics given above. An exercise concerning the games would show that, according to the distinctions made here, those formalizations were adequate in the sense that e.g. counting is indeed necessary to deal with the properties in Example 1.5 (they are not definable in $L^2_{\infty\omega}$), or that those in Example 1.2, and (a) and (b) of Example 1.5 are not modal (they are not definable in graded ML_∞).

Proviso As a general proviso for the entire paper, all structures considered are in purely relational vocabularies. Constants could for the most part be admitted without changing the major results; functions, however, have to be excluded if the restriction to two variables is to make sense. In view of the game characterizations we shall further assume from now on that the arity of relation symbols is at most 2. Essentially this is no loss of generality for our purposes: an inspection of the two-pebble games shows that atoms involving more than two elements cannot matter; this observation may actually be turned into a reduction which replaces for instance two-element atoms for higher-arity relations through the introduction of new binary relations.

1.3 More examples of two-variable logics

Two-variable least fixed-point logic $\mathbf{FP^2}$ FP^2 extends FO^2 by a least fixed-point operator, analogous to the extension of ML to L_μ. Syntactically we admit the following: if $\varphi(X, x)$ is positive in a monadic predicate variable X and if x is the only free first-order variable in φ, then $[\mathrm{LFP}_{X,x}\varphi](x)$ is a formula. Semantically $[\mathrm{LFP}_{X,x}\varphi]$ defines the least subset S of the universe satisfying the fixed-point equation $S = \{a : \varphi[S, a]\}$. As with L_μ, the least fixed point can also be obtained from *stages* defined by transfinite induction. Clearly $\mathrm{L}_\mu \subsetneq \mathrm{FP}^2$ just as $\mathrm{ML} \subsetneq \mathrm{FO}^2$.

Similar to the situation discussed for L_μ, definability of the stages in $L^2_{\infty\omega}$ shows that $\mathrm{FP}^2 \subseteq L^2_{\infty\omega}$ over any class of structures of uniformly bounded cardinality. This restricted inclusion is sufficient, however, to guarantee that FP^2 is a two-variable logic in the sense that it is preserved under two-variable equivalence. It is essential for these considerations that the application of the fixed-point operator is restricted to formulae $\varphi(X, x)$ in which y does not occur free [15].

Two-variable transitive closure logic $\mathbf{TC^2}$ The fact that the expressive power of FO^2 (and even of $L^2_{\infty\omega}$) remains very limited for defining binary predicates suggests to extend FO^2 in ways to render specific derived accessibilities definable. One very natural candidate involves the introduction of the transitive closure of already definable binary predicates. Transitive closures of accessibilities are explicitly or implicitly available in several temporal logics and process logics. The natural extension of FO^2 to incorporate transitive closures, TC^2, extends syntax and semantics of FO^2 by a *transitive closure operator*, which applied to a formula $\varphi(x,y)$ yields a new formula $[\mathrm{TC}\varphi](x,y)$, with the following semantics: $\mathfrak{A} \models \big[\mathrm{TC}\varphi\big][a,b]$ if the pair (a,b) is in the transitive closure of the relation $\{(a_1,a_2) : \mathfrak{A} \models \varphi[a_1,a_2]\}$. While TC^2 is formally in two variables, Example 1.9 below reveals that it is *not a two-variable logic* in the sense of the previous section. Observe that of the two graphs involved in that example, one is transitive and the other is not.

Example 1.9 The following two structures with reflexive and symmetric E are equivalent w.r.t. the two-pebble game with counting: a universe of six vertices arranged in one cycle w.r.t. E, and a universe of six vertices which splits into two 3-cycles w.r.t. E. It follows that TC^2 is not a two-variable logic in the strict sense.

Example 1.10 Acyclicity of E over arbitrary structures, and thus also well-foundedness of E over finite structures, are expressible in TC^2 by $\neg\exists x\exists y\big(\mathrm{TC}(Exy)(x,y) \wedge x = y\big)$.

It is shown in [15] that TC^2 cannot express well-foundedness over countably infinite structures. The following well-known process logic has recently been studied in its relationship with TC^2 for model checking applications [22].

Computation tree logic CTL For the typical interpretation of CTL as a branching time logic one usually considers structures with a single binary predicate E (with the intended meaning of temporal successor, or a non-deterministic next-step-operation). The syntax and semantics of CTL is divided between so-called *state formulae* and *path formulae*. State formulae are interpreted as describing properties of single vertices (states), i.e. as defining monadic queries. Path formulae on the other hand describe properties of *maximal E-paths*. An E-path is maximal if it has no proper extension, i.e. if it is infinite or if it terminates in a vertex without E-successors.[3] Let us write state formulae as formulae φ, and path formulae as $\pi(p)$ where p is treated as a formal variable ranging over maximal paths $p = a_0a_1a_2\ldots$ where $(a_i, a_{i+1}) \in E$. The syntax of CTL is defined in a simultaneous induction on state and program formulae:

(i) P_i is a state formula, for each basic predicate P_i, and so are $\bot$ and $\top$.
(ii) Boolean combinations of state formulae are state formulae.
(iii) If φ is a state formula, then $\big(\mathrm{next}\,\varphi\big)(p)$ is a path formula.
(iv) If φ_1 and φ_2 are state formulae, then $\big(\varphi_1 \mathrm{until}\, \varphi_2\big)(p)$ is a path formula.

[3] It is customary with CTL to restrict attention to structures in which every E-path extends infinitely. Our slightly more general convention is more suited to the study of natural extensions of CTL; also we do not lose any of the standard infinite-paths semantics, because terminal vertices are defined by the ML-formula $\Box\bot$.

(v) If $\pi(p)$ is a path formula then $\forall p \pi(p)$ and $\exists p \pi(p)$ are state formulae.

The semantics is clear for (i) and (ii). For (iii) and (iv), a path $p = a_0 a_1 \ldots$ satisfies $\big(\mathrm{next}\,\varphi\big)$ if φ holds at a_1, and p satisfies $\big(\varphi_1 \mathrm{until}\, \varphi_2\big)$ if for some j, φ_2 holds at a_j and φ_1 holds at a_i for all $i < j$. For (v), $\forall p \pi(p)$ is true at a if all paths of the form $p = a \ldots$ satisfy π. Similarly, $\exists p \pi(p)$ requires that there is some path of that form satisfying π.

To give some examples, consider the properties in Example 1.3: (a) is expressed by $\exists p(\top \mathrm{until}\, P)$, and (b) by $\forall p(P_2 \,\mathrm{until}\, P_1)$.

We are here interested in the semantics of CTL only for state formulae and correspondingly regard the path formulae just as auxiliary constructs in the inductive definition of syntax and semantics. It is instructive to review the translation of CTL (meaning its state formulae) into L_μ, which in particular shows that CTL is a two-variable logic of modal type. Inductively the translation is obtained as follows. With (i) and (ii) we trivially remain within ML. (iii) and (iv) do not generate state formulae directly, but do so only if plugged into (v). The next-operator in $\forall p$ or $\exists p$ simply translates into ordinary modal quantification $\Box$ or $\Diamond$, e.g. $\forall p\big(\mathrm{next}\,\varphi\big)$ is equivalent with $\Box\varphi$. For the until-construct consider the formulae

$$\begin{array}{rcl} \psi_\exists & = & \exists p\big(\varphi_1 \,\mathrm{until}\, \varphi_2\big), \\ \psi_\forall & = & \forall p\big(\varphi_1 \,\mathrm{until}\, \varphi_2\big). \end{array}$$

Clearly the following are sound translations into L_μ:

$$\begin{array}{rcl} \psi_\exists & \equiv & \mu_X\big(\varphi_2 \vee (\varphi_1 \wedge \Diamond X)\big), \\ \psi_\forall & \equiv & \mu_X\big(\varphi_2 \vee (\varphi_1 \wedge \neg\Box\bot \wedge \Box X)\big).(^4) \end{array}$$

Interestingly, $\psi_\forall$ turns out to be a well-foundedness statement. Consider the extended modal framework in which modalities for derived accessibilities are available. Then

$$\psi_\forall \equiv \mu_X\big([\xi]X\big)\,, \qquad \text{where } \xi(x,y) = \big(Exy \vee \neg\varphi_1(x) \vee (\Box\bot)(x)\big) \wedge \neg\varphi_2(x).$$

This equivalence is easily verified through comparison of the stages of the respective least fixed points (again, the disjunct $\Box\bot$ is necessary only in our relaxed framework where maximal E-paths need not a priori be infinite). Thus $\psi_\forall$ defines the well-founded part of the converse of the relation defined by ξ.

It follows that over finite structures, CTL translates into TC^2. The obvious TC^2-translation of $\psi_\exists$,

$$\varphi_2(x) \vee \exists y\big(\mathrm{TC}\big(Exy \wedge \varphi_1(x)\big)(x,y) \wedge \varphi_2(y)\big)$$

[4] The conjunct $\neg\Box\bot$ takes care of the possible termination of maximal paths, in our more general setting; it could be dropped under the standard assumption that all E-paths are infinitely extendible.

is actually valid over arbitrary structures. And $\psi_\forall$, being a well-foundedness statement, is equivalent with a TC^2-formula just over finite structures by Example 1.10. Immerman and Vardi [22] extend the inclusion $\mathrm{CTL} \subseteq \mathrm{TC}^2$ over finite structures from CTL to CTL*. CTL* is a variant of CTL with somewhat nicer closure properties since it admits Boolean combinations of path formulae and non-trivial nesting of temporal operators 'until' and 'next'. To capture the semantics of CTL* state formulae they have to use a further extension of TC^2, where transitive closures may involve several Boolean variables along with the usual two element variables. For model checking complexity, this translation proves to be more useful than the standard translation to L_μ.

Actually the specific fragment of L_μ needed to capture CTL and its natural lift to the FO^2-framework, are also rather interesting. Indeed one may propose the following extension, rather than TC^2, as a natural candidate to extend both FO^2 and CTL within the two-variable context. Consider the following two closure operators $\langle\cdot\rangle^\infty$ and $[\cdot]^\infty$, which lead from a pair of formulae $\xi(x, y)$ and $\varphi(x)$ (with free variables as indicated) to new formulae $\psi_1(x) = \big(\langle\xi\rangle^\infty\varphi\big)(x)$ and $\psi_2(x) = \big([\xi]^\infty\varphi\big)(x)$, respectively. Semantically, ψ_1 describes the closure of the set defined by φ under the binary predicate defined by ξ, i.e. we let $\psi_1(x)$ be equivalent with $\bigvee_{n\geqslant 0}\big((\langle\xi\rangle)^n\varphi\big)(x)$. Equivalently, put $\psi_1 \equiv \mu_X\big(\varphi \vee \langle\xi\rangle X\big)$. The semantics of ψ_2 is that of $\mu_X\big(\varphi \vee [\xi]X\big)$. Note, however, that unlike $\mu_X\big(\varphi \vee \langle\xi\rangle X\big)$, the fixed point $\mu_X\big(\varphi \vee [\xi]X\big)$ need not close within ω steps over infinite structures.

Definition 1.11 Let CL^2 be the extension of FO^2, which augments the FO^2-rules for the formation of formulae with the $\langle\cdot\rangle^\infty$- and $[\cdot]^\infty$-constructions.

Clearly $\mathrm{CL}^2 \subseteq \mathrm{FP}^2$, whence CL^2 is indeed a true two-variable logic. Also, by the considerations outlined above, $\mathrm{CTL} \subseteq \mathrm{CL}^2$, in a translation that is sound over arbitrary structures. Again, these inclusions are strict.

We have thus isolated the natural candidates to lift the chain of extensions $\mathrm{ML} \subsetneq \mathrm{CTL} \subsetneq \mathrm{L}_\mu$ to the level of FO^2 within the framework of two-variable logics. It is not hard to see that indeed CL^2 is the natural least common extension of FO^2 and CTL, and FP^2 the natural least common extension of FO^2 and L_μ, under some reasonable closure conditions.

$$\begin{array}{ccccc}
\mathrm{FO}^2 & \subsetneqq & \mathrm{CL}^2 & \subsetneqq & \mathrm{FP}^2 \\
\rotatebox{90}{$\subsetneqq$} & & \rotatebox{90}{$\subsetneqq$} & & \rotatebox{90}{$\subsetneqq$} \\
\mathrm{ML} & \subsetneqq & \mathrm{CTL} & \subsetneqq & \mathrm{L}_\mu
\end{array}$$

We shall see in Section 3, however, that even these minimal lifts of extremely well-behaved extensions of ML to the level of FO^2 turn out to be undecidable for satisfiability.

2 Decidability results

Consider a class of formulae X. Those subclasses of X that come up in the classical decision problem for X are the following.

- $sat(X)$, consisting of those $\varphi \in X$ that have a model;
- $fin\text{-}sat(X)$, consisting of those $\varphi \in X$ that have a finite model;
- $inf\text{-}axioms(X) = sat(X) \setminus fin\text{-}sat(X)$, the *infinity axioms* of X;
- $non\text{-}sat(X) = X \setminus sat(X)$, consisting of the unsatisfiable $\varphi \in X$.

X has the *finite model property* if every satisfiable formula in X even has a finite model: $sat(X) = fin\text{-}sat(X)$. The finite model property is a crucial model theoretic property of many (but not all) classes X for which $sat(X)$ is decidable (see [8]). Note that for every recursive formula class $X \subseteq \mathrm{FO}$, the finite model property of X implies that $sat(X)$ is decidable. Indeed $sat(X)$ is then recursively enumerable (since $fin\text{-}sat(X)$ trivially is), and by the completeness theorem for first-order logic, also $non\text{-}sat(X)$ is recursively enumerable. An easy model theoretic proof for the decidability of propositional modal logic ML, for instance, uses the finite model property of ML and the embedding of ML into FO^2 and hence into FO. Indeed FO^2 itself has the finite model property [26]. The prominent process logics extending ML are decidable and also share the finite model property; e.g. see [24] for L_μ. On the other hand, none of the corresponding extensions of FO^2 retains the finite model property; see [15] for infinity axioms in TC^2, FP^2, C^2 and others. It is important to realize, however, that violation of the finite model property does by no means rule out decidability of either $fin\text{-}sat(X)$, or $sat(X)$, or both.

If nevertheless most interesting extensions of FO^2, with the notable exception of C^2, fail to be decidable, this should not be blamed on the failure of the finite model property, but on the failure of the so-called *tree model property*. The tree model property requires that every satisfiable formula has a tree-like model, a phenomenon that is well known in modal logics. Vardi [38] argues convincingly that the tree model property provides the crucial tools – namely the sophisticated use of tree-automata – to prove decidability (along with good complexity bounds) in the context of modal process logics. The surprising robustness of ML under extensions can thus be attributed to the modal character of these typical extensions. Indeed, the tree model property follows from preservation under bisimulation equivalence, whence all two-variable logics of the modal type share the tree model property. FO^2 on the other hand does not have the tree model property.

In this section we deal with decidability results concerning FO^2 (having the finite model property, but not the tree model property), and C^2 (having neither the finite nor the tree model property).

2.1 Skolemization and Scott's normal form

Explicitly working with FO^2 we review a common technique for reducing quantifier complexity while preserving satisfiability. The basic idea, also known as Skolemization, is to substitute new predicate names for subformulae together with formulae that guarantee the soundness of this substitution. This method has been applied by Scott [34] in 1962 to give an elegant reduction of the satisfiability problem for FO^2 to the Gödel case (the $\forall^2\exists^*$-prefix class) with equality of the classical decision problem. At that time, Gödel's claim that his decidability proof for the $\forall^2\exists^*$-prefix class without equality could be extended 'without difficulty' to formulae containing equality was still believed to be true. Thus, Scott's reduction seemed to imply the decidability of the satisfiability problem for FO^2. Although it turned out later that Gödel's claim was false (see [8]), Scott's reduction became an essential preparatory step in all subsequent proofs of the decidability of FO^2 and with slight modifications carries over to many extensions of FO^2 that have been considered, and in particular to C^2.

Consider the lowest level of quantifier introduction in an FO^2-formula. Up to trivial exchange of variables, we are dealing with one of the following formulae, where displayed variables are actually free and ψ_0 quantifier-free: $\psi(x) = \exists y \psi_0(x,y)$, $\psi(x) = \forall y \psi_0(x,y)$, $\psi = \exists y \psi_0(y)$, or $\psi = \forall y \psi_0(y)$. In the case of the first two, with one remaining free variable, we introduce a new unary predicate P_ψ with intended semantics $P_\psi x \leftrightarrow \psi(x)$. It is readily checked that this stipulation is captured by the following simple FO^2-assertions θ_ψ of quantifier depth 2:

$$\begin{aligned} \theta_\psi &= \forall x \exists y \big(\psi_0(x,y) \leftrightarrow P_\psi x\big) && \text{if } \psi(x) = \exists y \psi_0(x,y), \\ \theta_\psi &= \forall x \forall y \big(\psi_0(x,y) \leftrightarrow P_\psi x\big) && \text{if } \psi(x) = \forall y \psi_0(x,y). \end{aligned}$$

If $\psi(x)$ of either of these forms appears as a subformula of some φ, then satisfiability of φ is equivalent with satisfiability of $\varphi' \wedge \theta_\psi$, where φ' is the result of substituting $P_\psi x$ for every occurrence of the subformula $\psi(x)$ in φ.

In the cases where ψ does not retain any free variables we similarly may simulate the Boolean value of ψ with the use of a unary predicate P_ψ and a dummy constant c (which will be eliminated in the end). Now ψ is to be substituted by $P_\psi c$, and for the semantic adequacy of this substitution we add assertions similar to the above:

$$\begin{aligned} \theta_\psi &= \forall x \exists y \big(\psi_0(y) \leftrightarrow P_\psi x\big) && \text{if } \psi = \exists y \psi_0(y), \\ \theta_\psi &= \forall x \forall y \big(\psi_0(y) \leftrightarrow P_\psi x\big) && \text{if } \psi = \forall y \psi_0(y). \end{aligned}$$

Again, we obtain satisfiability equivalence between φ and $\varphi'_c \wedge \theta_\psi$, where φ' is the result of substituting $P_\psi c$ for subformulae ψ throughout φ.

Starting with an arbitrary sentence $\varphi \in \mathrm{FO}^2$ and applying this procedure $\varphi \mapsto \varphi'$ recursively to the minimal subformulae of quantifier depth 1, and collecting conjuncts

θ_ψ along the way, we finally obtain a quantifier-free FO^2-sentence $\widehat{\varphi}$ and a conjunction θ of prenex FO^2-sentences of type $\forall\forall$ and $\forall\exists$ such that φ is satisfiable (in a finite model) if and only if $\widehat{\varphi} \wedge \theta$ is satisfiable (in a finite model). The dummy constant c, which only occurs in $\widehat{\varphi}$, may be eliminated if we finally replace $\widehat{\varphi}$ by $\varphi^* = \exists x \widehat{\varphi}[x/c]$. Recombining conjuncts in $\varphi^* \wedge \theta$ we obtain a conjunction of at most one $\forall\forall$-sentence with several $\forall\exists$-sentences. It is checked inductively that the length of φ^* is linearly bounded in the length of φ. Note also that we need only introduce new unary predicates, and that their number is bounded by the number of subformulae of the type $Q\psi$ in φ, where Q is $\forall$ or $\exists$.

Theorem 2.1 (Scott) *There is a polynomial-time computable reduction* $\mathrm{NF}\colon \mathrm{FO}^2 \to \mathrm{FO}^2$ *mapping every sentence* $\psi \in \mathrm{FO}^2$ *to a sentence* $\mathrm{NF}(\psi)$ *(with extended vocabulary) of the form*

$$\forall x \forall y \chi_0 \wedge \textstyle\bigwedge_{i=1}^m \forall x \exists y \chi_i$$

with quantifier-free χ_i*, such that* ψ *and* $\mathrm{NF}(\psi)$ *are satisfiable over the same domains. Moreover the length of* $\mathrm{NF}(\psi)$ *is linear in the length of* ψ*.*

As mentioned above, this reduction extends to C^2. It turns out that also here quantifier prefixes can be reduced to the form $\forall\forall$ and $\forall\exists^{\geqslant n}$, $\forall\exists^{\leqslant n}$, respectively $\forall\exists^{=n}$ with a normal form mapping which is linearly bounded w.r.t. formula length. A further and very useful reduction proceeds to eliminate all forms of counting quantifiers apart from the very limited $\exists^{=1}$. This additional reduction step is achieved essentially by paving satisfaction sets with new singleton sets rendered by new predicates. For example, observe that $\exists^{\geqslant n} x \psi(x)$ is faithfully rendered by the conjunction of $\forall x ((\bigvee_{i=1}^n P_i x) \to \psi(x)$, $\forall x \bigwedge_{i \neq j} \neg(P_i x \wedge P_j x)$, and $\bigwedge_{i=1}^n \exists^{=1} x P_i x$, for new P_i. Note, however, that if n is assumed to be encoded in binary in the formula $\exists^{\geqslant n} x \psi(x)$ as usual, then the resulting formula is exponentially longer, as we introduce n new predicates.

Theorem 2.2 (Grädel, Otto, Rosen) *There is a reduction* $\mathrm{NF}\colon \mathrm{C}^2 \to \mathrm{C}^2$ *mapping every sentence* $\psi \in \mathrm{FO}^2$ *to a sentence* $\mathrm{NF}(\psi)$ *(with extended vocabulary) of the form*

$$\forall x \forall y \chi_0 \wedge \textstyle\bigwedge_{i=1}^m \forall x \exists^{=1} y \chi_i$$

with quantifier-free χ_i*, such that* ψ *and* $\mathrm{NF}(\psi)$ *are satisfiable over the same domains. The reduction is computable in exponential time and may increase the formula length exponentially. There also exists a similar but weaker normal form that admits conjuncts of the form* $\forall x \exists^{\geqslant n} y \chi_i$ *and* $\forall x \exists^{\leqslant n} y \chi_i$ *for arbitrary* n *which is computable in polynomial time and remains linearly bounded in terms of formula length.*

2.2 Finding the 'right' models

A key step in the decidability proofs for FO^2, C^2 and related logics consists in passing from arbitrary models to models whose structure is sufficiently regular, so that there is a recursive combinatorial criterion for checking whether such special, regular models exist. The cleanest such approach is exemplified in the treatment of FO^2. FO^2 does possess the *finite model property*, i.e. $sat(\mathrm{FO}^2) = \mathit{fin\text{-}sat}(\mathrm{FO}^2)$. In fact there is a recursive f, $f(\varphi)$ exponentially bounded in the length of φ, such that any FO^2-sentence φ that is satisfiable at all, also has a model of size at most $f(\varphi)$. One usually refers to this as a *small model property.* We may thus use as special models for a sentence φ its small models, namely those whose size is bounded by $f(\varphi)$. Whether φ has such a special model is clearly recursive (here actually in NEXPTIME). The small model property guarantees that the $f(\varphi)$ size bounded structures are fully representative as a class of candidate models for φ.

In the case of C^2 one is not quite as lucky, since C^2 does not have the finite model property: there are satisfiable C^2-sentences without finite models. For instance, the conjunction of $\forall x \exists^{=1} y Exy$, $\forall y \exists^{\leqslant 1} x Exy$, and $\exists y \forall x \neg Exy$ says that E is the graph of a function that is 1–1 but not onto. Hence a representative class of special models for C^2-sentences has to comprise infinite models. It turns out that sufficiently homogeneous models are always available. These admit finite descriptions, from which one can then abstract recursive criteria for satisfiability in infinite models.

2.2.1 Basic types as building blocks

Consider FO^2 and the normal form of Theorem 2.1. Think of normal form FO^2-sentences φ in vocabulary τ_φ consisting of finitely many unary and binary predicates. Prenex quantifier rank 2 suggests to analyze models in terms of quantifier-free 1- and 2-types. A *basic* 1-*type*, respectively *basic* 2-*type*, in a finite and purely relational vocabulary τ, is a maximally consistent finite set of atomic and negated atomic formulae in the single variable x, respectively in variables x and y. For 2-types we also require $\neg x = y$ rather than $x = y$ to be a member. Clearly, for any fixed, finite and relational vocabulary, there are only finitely many different basic 1- and 2-types: their numbers are actually exponentially bounded in the number of predicates in τ. Let $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ stand for the sets of basic 1- and 2-types, respectively.

An element a of a structure $\mathfrak{A}$ realizes the basic 1-type $\alpha = \mathrm{tp}_{\mathfrak{A}}(a)$ consisting of just those (negated) atomic formulae that a satisfies in $\mathfrak{A}$. Similarly for a non-degenerate pair (a, b) and its basic 2-type $\beta = \mathrm{tp}_{\mathfrak{A}}(a, b)$. Let the sets $\boldsymbol{\alpha}_{\mathfrak{A}}$ and $\boldsymbol{\beta}_{\mathfrak{A}}$ be the sets of all basic 1-types, respectively 2-types realized in $\mathfrak{A}$. It is clear that for any quantifier-free FO^2-formula χ the sets $\boldsymbol{\alpha}_\chi = \{\alpha \in \boldsymbol{\alpha} : \alpha \models \chi\}$, respectively $\boldsymbol{\beta}_\chi = \{\beta \in \boldsymbol{\beta} : \beta \models \chi\}$ are easily computable from τ and χ. Furthermore, any 2-type β uniquely determines the 1-types $\beta|_x$ and $\beta|_y$ of its x- and y-component.

Obviously, $\mathfrak{A}$ satisfies a normal form FO^2-sentence $\forall x \forall y \chi_0 \wedge \bigwedge_{i=1}^m \forall x \exists y \chi_i$, if and only if

- $\boldsymbol{\beta}_{\mathfrak{A}} \subseteq \boldsymbol{\beta}_{\chi_0}$.
- for every i, $1 \leqslant i \leqslant m$, and every every a in $\mathfrak{A}$ there is some $b \neq a$ in $\mathfrak{A}$ such that $\mathrm{tp}_{\mathfrak{A}}(a,b) \in \boldsymbol{\beta}_{\chi_i}$.

Any structure $\mathfrak{A}$ can be completely specified by allocating basic 2-types in a consistent manner to all pairs of elements. Here, consistency means, that the 2-types allocated to pairs with a common element assign the same 1-type to that element. Viewing basic types in this way as the building blocks of a model, we find that χ_0 poses a global constraint on the 2-types that may be realized in $\mathfrak{A}$, while the other χ_i require witnesses, in the sense that certain 1-types must always be extendible to appropriate 2-types.

Actually, $\boldsymbol{\beta}_{\mathfrak{A}}$ (respectively $\boldsymbol{\beta}_{\chi_0}$) put some non-trivial conditions about basic 1-types (the trivial ones are those that determine the 1-types as the restrictions of the 2-types): $\alpha \in \boldsymbol{\alpha}_{\mathfrak{A}}$ is realized only once in $\mathfrak{A}$ if and only if there is no $\beta \in \boldsymbol{\beta}_{\mathfrak{A}}$ such that $\beta|_x = \beta|_y = \alpha$. It has become customary to call an element of a structure a *king* if its 1-type is realized by no other element of that structure; accordingly the 1-type of a king is *royal*.

2.2.2 Regular witnessing patterns

Let $\mathfrak{A} \models \varphi$, $\varphi = \forall x \forall y \chi_0 \wedge \bigwedge_{i=1}^m \forall x \exists y \chi_i$. W.l.o.g. suppose that for $i \geqslant 1$, $\chi_i(x,y)$ entails $x \neq y$ (replacing χ_i with $(\chi(x,y) \vee \chi(x,x)) \wedge \neg x = y$ if necessary, which is sound over all structures with at least two elements).

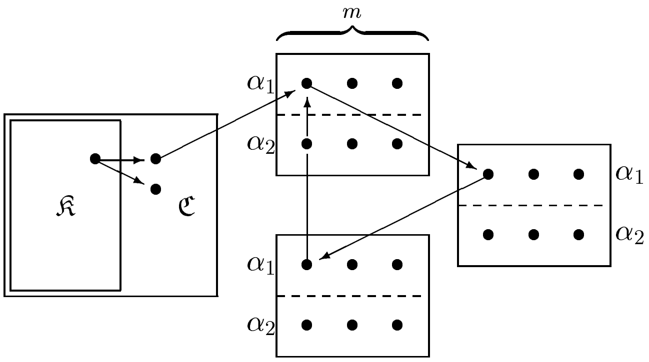
We are looking for some $\mathfrak{B} \models \varphi$ (a small model for φ), where $|B|$ is exponentially bounded in $|\varphi|$. $\mathfrak{B}$ is obtained from $\mathfrak{A}$ in several steps as follows. Recall that $|\boldsymbol{\alpha}|$ is exponentially bounded in the number of predicates, and therefore in $|\varphi|$.

The kings: Let $\mathfrak{K} \subseteq \mathfrak{A}$ be the substructure whose universe K consists of just the set of kings in $\mathfrak{A}$. Clearly $|K| \leqslant |\boldsymbol{\alpha}|$.

The court: For the $\forall\exists$-requirements at kings, select witnesses $a_{k,i}$ in $\mathfrak{A}$ such that $\mathfrak{A} \models \chi_i[k, a_{k,i}]$. Let $\mathfrak{C} \subseteq \mathfrak{A}$ be the substructure whose universe is $C = K \cup \{a_{k,i} : k \in K, 1 \leqslant i \leqslant m\}$. Clearly $|C| \leqslant (m+1)|\boldsymbol{\alpha}|$.

Circular witnessing: Deviating from the given $\mathfrak{A}$, extend $\mathfrak{C}$ to $\mathfrak{B}$ over a universe B consisting of C together with 3 disjoint boxes of new elements, each box containing exactly m new elements of each non-royal 1-type of $\mathfrak{A}$. It follows that $|B| \leqslant (4m+1)|\boldsymbol{\alpha}|$. The new elements are used to take care of the $\forall\exists$-requirements at elements in $C \setminus K$ and among themselves, through a circular witnessing pattern as explained below.

Completion: Settle all remaining 2-types with appropriate $\beta \in \boldsymbol{\beta}_{\mathfrak{A}}$, to complete the interpretation of $\mathfrak{B}$ as a structure. Note that for all tasks that occur there is precedent in $\mathfrak{A}$: if $b \neq b' \in B$, then there are $a \neq a'$ of corresponding 1-types in $\mathfrak{A}$, and one may put $\mathrm{tp}_{\mathfrak{B}}(b, b') = \mathrm{tp}_{\mathfrak{A}}(a, a')$ (it is exactly to make this step go through, that the kings received special treatment).



Concerning the *circular witnessing* scheme, notice first that any element b in $B \setminus K$ is made to realize some 1-type α in $\boldsymbol{\alpha}_{\mathfrak{A}}$. We may thus determine some representative a of its 1-type in $\mathfrak{A}$, and find witnesses $a_1, \ldots, a_m$ in $\mathfrak{A}$, such that $\beta_i = \mathrm{tp}_{\mathfrak{A}}(a, a_i) \in \boldsymbol{\beta}_{\mathfrak{A}}$ are such that β_i entails χ_i and $\beta_i|_x = \alpha$. If the 1-type at the far end of some β_i, $\beta_i|_y$, happens to be royal, then some $k \in K \subseteq B$ realizes it and we may put $\mathrm{tp}_{\mathfrak{B}}(b, k) = \beta_i$ in $\mathfrak{B}$. For the other ones we find distinct b_i of the appropriate 1-type in $B \setminus C$, according to the following circular pattern: if $b \in C \setminus K$, choose b_i in the first box of new elements; if b is form the first/second/third box of new elements, then choose b_i in the second/third/first box.

The finite model property of FO^2, and hence the decidability of $sat(FO^2)$, was established by Mortimer [26]. (Scott's reduction to the Gödel case implies the decidability and finite model property for the equality-free part of FO^2.) A doubly-exponential bound on the size of a minimal model is implicit in Mortimer's proof. The arguments outlined here, due to Grädel, Kolaitis and Vardi [14], provide an exponential upper bound on the model size. This implies that the satisfiability problem for FO^2 can be decided in nondeterministic exponential time. This matches a previously known lower bound [12, 25].

Theorem 2.3 (Grädel, Kolaitis, Vardi) *There exists a constant c such that every satisfiable FO^2-sentence of length n has a model of cardinality at most 2^{cn}. Further, $sat(\mathrm{FO}^2)$ (and hence fin-$sat(\mathrm{FO}^2)$) is* NEXPTIME*-complete.*

2.2.3 Decidability of $sat(\mathrm{C}^2)$

Recall that C^2 does not share the finite model property, so that any decidability proof for C^2 has to take into account infinite models. As $\mathrm{C}^2 \subseteq \mathrm{FO}$, the complement of $sat(\mathrm{C}^2)$ is recursively enumerable by FO-completeness. $\mathit{fin}\text{-}sat(\mathrm{C}^2)$ is trivially recursively enumerable. Hence, for decidability of $sat(\mathrm{C}^2)$ it actually suffices to show that also the class of those C^2-sentences that have infinite models is recursively enumerable.

Again we may view the essential step as an analysis of arbitrary (infinite) models $\mathfrak{A}$ of some normal form C^2-sentence φ which leads to the construction of some special (here: especially regular and homogeneous) model $\mathfrak{B} \models \varphi$. The class of special models can be chosen such that the set of those normal form C^2-sentences that are satisfied in special models becomes recursively enumerable. This immediately also provides an enumeration of $\mathit{inf}\text{-}sat(\mathrm{C}^2)$, thus proving $sat(\mathrm{C}^2)$ decidable.

Recall the normal form of Theorem 2.2 and let $\varphi = \forall x \forall y \chi_0 \wedge \bigwedge_{i=1}^{m} \forall x \exists^{=1} y \chi_i$. The $\forall x \exists^{=1}$-assertions suggest to consider the following *basic counting types* $\gamma = \mathrm{ctp}_{\mathfrak{A}}(a)$ as the fundamental building blocks of models.

$$\begin{array}{rcl} \gamma\colon \boldsymbol{\beta} & \longrightarrow & \{0,1,2^+\} \\ \beta & \longmapsto & \gamma(\beta) = \left|\{b \in A : \mathrm{tp}_{\mathfrak{A}}(a,b) = \beta\}\right|^*, \end{array}$$

where $|S|^*$ is the size of set S counted according to 0, 1, 'many' (coded as 2^+). Let $\boldsymbol{\gamma}$ be the finite set of basic counting types, $\boldsymbol{\gamma}_{\mathfrak{A}} \subseteq \mathfrak{A}$ the set of those realized in $\mathfrak{A}$.

Let us outline the passage from some infinite $\mathfrak{A} \models \varphi$ to the more regular $\mathfrak{B} \models \varphi$. We now regard an element of $\mathfrak{A}$ as a *king* (and its basic counting type as royal) if its basic counting type is realized in $\mathfrak{A}$ only finitely often. $\mathfrak{B}$ is obtained in the following steps.

The kings: Let $\mathfrak{K} \subseteq \mathfrak{A}$ be the substructure whose universe K consists of the set of kings in $\mathfrak{A}$. As $\boldsymbol{\gamma}$ is finite and each royal counting type is realized finitely often, K is finite.

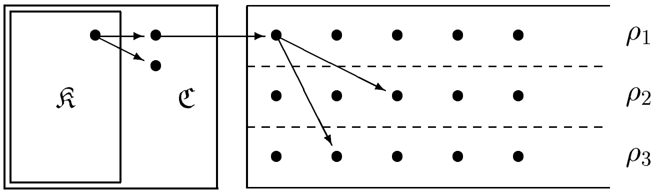
The court: In $A \setminus K$, finitely many basic counting types are each realized infinitely often. Call $a, a' \in A \setminus K$ equivalent, $a \approx a'$, if $\mathrm{ctp}_{\mathfrak{A}}(a) = \mathrm{ctp}_{\mathfrak{A}}(a')$ and if $\mathfrak{A} \restriction (K \cup \{a\})$ and $\mathfrak{A} \restriction (K \cup \{a'\})$ are isomorphic. Let $\mathfrak{C}$ be the substructure of $\mathfrak{A}$ whose universe is K together with all finite $\approx$-classes. As $\approx$ has finite index, C is finite.

Uniform witnessing: $A \setminus C$ consists of finitely many infinite $\approx$-classes $\rho_1, \ldots, \rho_l$, each characterized by its basic counting type and by the basic 2-types its elements realize with kings. We extend $\mathfrak{C}$ to $\mathfrak{B}$ over a universe B consisting of C together with countably many new elements for each of the classes ρ_j. Basic counting types are put right by a systematic prescription of basic 2-types β to pairs (a,b) for suitable $b \in B \setminus C$, whenever a has fewer β-incidences than its basic counting type requires.

Completion: All he remaining 2-types are settled with appropriate $\beta \in \boldsymbol{\beta}_{\mathfrak{A}}$, to complete the interpretation of $\mathfrak{B}$, in such a way that β-counts 0, 1, 2^+ are preserved at both ends whenever a pair is newly attributed atomic 2-type β.

The *uniform witnessing* can actually be organized such that the recipient b, at the other end of a required β-edge at a, is chosen in a uniform way as a minimal as yet unused element in an appropriate class ρ_j, where ρ_j only depends on the $\approx$-class of a.

Even more interestingly, the *completion* step can be handled in an equally uniform manner. A result in the Ramsey theory of complete bipartite graphs gives that, for any two elements $b \neq b' \in B \setminus K$, there is some β depending only on $\gamma = \mathrm{ctp}(b)$ and $\gamma' = \mathrm{ctp}(b')$ such that putting $\mathrm{tp}_{\mathfrak{B}}(b, b') = \beta$ is harmless. Indeed, from the fact that γ and γ' are each realized infinitely often in $\mathfrak{A}$, one infers from Theorem 1 in Chapter 5 of [17] that there is actually some β such that for β and its converse $\beta^{-1}(x, y) = \beta(y, x)$ we have $\gamma(\beta) = \gamma'(\beta^{-1}) = 2^+$. Putting $\mathrm{tp}_{\mathfrak{B}}(b, b') = \beta$ can thus not affect the basic counting types of b and b' (provided they already had the correct β-, respectively β^{-1}-counts before).



It is not surprising that, given the regular pattern of the special structures $\mathfrak{B}$, there is a finite description of these $\mathfrak{B}$ on the basis of which one can check

(a) consistency of that information as the description of an infinite structure, and

(b) whether this structure satisfies a given normal form C^2-sentence φ.

Actually (b) is almost trivial, since $\boldsymbol{\gamma}_{\mathfrak{A}}$ alone (recursively) determines the set of those normal form sentences that are satisfied in $\mathfrak{A}$. Moreover it is clear that $\mathfrak{A}$ and its companion $\mathfrak{B}$ satisfy the same normal form sentences. It is condition (a), which necessitates the rather more involved preparation of $\mathfrak{B}$ as given above. It turns out that for a characteristic description we can use the following: the full specification of the substructures of kings and court, $\mathfrak{K}$ and $\mathfrak{C}$, the specification of the infinite classes ρ_j, and of the finite classes within C. We do not repeat the combinatorially more involved arguments here, but rather refer to the original source [16]. Summing up, we have indicated the main arguments towards the proof of the first part of the following theorem.

Theorem 2.4 (Grädel, Otto, Rosen) *The satisfiability problem for* C^2 *is decidable. Also the finite satisfiability problem for* C^2 *is decidable.*

Currently the best upper bound on the complexity of $sat(\mathrm{C}^2)$ is one of nondeterministic doubly-exponential time, established by Pacholski, Szwast and Tendera [32]. There remains a gap between this and the best known lower bound, which actually is just NEXPTIME (as for FO^2). The exponential gap between these bounds, and the corresponding uncertainty about the actual complexity, is closely linked to the exponential blow up encountered in the normal form for C^2, compare Theorem 2.2. In fact Pacholski, Szwast and Tendera do obtain a NEXPTIME-decision procedure for normal form C^2-sentences.

Theorem 2.5 (Pacholski, Szwast, Tendera) *Satisfiability C^2-sentences in normal form can be decided in* NEXPTIME. *It follows that $sat(\mathrm{C}^2)$ is decidable in nondeterministic doubly-exponential time.*

3 Undecidability results

Modal logics have very robust decidability properties. Extensions of modal logic by temporal operators, least and greatest fixed points, counting constructs provide interesting logical systems that are algorithmically quite manageable and important for applications in a number of areas. It turns out that most of the corresponding extensions of FO^2 are undecidable. In particular this is the case for the logics TC^2 and FP^2 which augment FO^2 by weak forms of recursion, such as transitive closure or (restricted) monadic fixed-point operations. Also the extension of FO^2 by cardinality comparison quantifiers or a choice construct, known as Hilbert's ε-operator are undecidable. In fact all these logics prove to be undecidable both for satisfiability, and for satisfiability in finite models. Moreover most of them are hard for Σ_1^1, the first level of the analytical hierarchy, and thus have a much higher degree of undecidability than first-order logic (see [15] for more details).

A closely related issue is the (un)decidability of the FO^2-theories of certain interesting model classes, defined by constraints on some of the relation symbols. For instance, let $\mathcal{K}$ be the class of structures of the form $\mathfrak{A} = (A, E, R_1, R_2, \dots)$ such that E is an equivalence relation on A (and $R_1, R_2, \dots$ are arbitrary relations). To put it differently, $\mathcal{K}$ is the closure of the class of equivalence relations (A, E) under expansions. We refer to the FO^2-theory of $\mathcal{K}$ as the FO^2-theory of one built-in equivalence relation. Similarly the FO^2-theory of several built-in equivalence relations and the FO^2-theory of several built-in graphs of functions are the FO^2-theories of the classes of structures $\mathfrak{A} = (A, E_1, E_2, \dots, R_1, R_2, \dots)$ where all E_i that are present in $\mathfrak{A}$ are equivalence relations, respectively graphs of unary functions.

Theorem 3.1 *The FO^2-theory of several built-in graphs of unary functions is decidable.*

Proof. This is an immediate consequence of the decidability of C^2. Indeed an FO^2-sentence ψ with relation symbols $E_1, \ldots, E_m, R_1, \ldots, R_k$ belongs to the FO^2-theory of several built-in graphs of unary functions if and only if $\neg\psi \wedge \bigwedge_{i=1}^{m} \forall x \exists^{=1} y E_i xy$ is *not* in $sat(C^2)$. $\square$

Another decidability result of this kind has been proved very recently [30].

Theorem 3.2 (Otto) *The* FO^2*-theory of one built-in equivalence relation is decidable.*

In contrast, we will prove below that the FO^2-theory of several built-in equivalence relations is undecidable in a strong sense. We will further show that this can be viewed as a strengthening of the undecidability of TC^2 and FP^2 in [15].

3.1 Recursive inseparability and strongly undecidable theories

A stronger variant of the unsolvability of the classical decision problem is Trakhtenbrot's *Inseparability Theorem* which uses the concept of recursive inseparability.

Definition 3.3 Two disjoint sets X, Y are called *recursively inseparable* if there is no recursive set R such that $X \subseteq R$ and $R \cap Y = \varnothing$. In particular, neither X nor Y can then be decidable.

Theorem 3.4 (Trakhtenbrot) *The sets fin-sat*(FO)*, inf-axioms*(FO) *and non-sat*(FO) *are pairwise recursively inseparable.*

Definition 3.5 A formula class L is a *conservative reduction class* if there is a recursive function $g : \mathrm{FO} \to L$ that preserves (in the sense of if-and-only-if) satisfiability as well as finite satisfiability.

For a conservative reduction class L it follows from Trakhtenbrot's Theorem that $\mathit{fin\text{-}sat}(L)$, $\mathit{inf\text{-}axioms}(L)$, and $\mathit{non\text{-}sat}(L)$ are pairwise recursively inseparable; in this case $\mathit{fin\text{-}sat}(L)$ and $\mathit{non\text{-}sat}(L)$ are r.e.-hard while $sat(L)$ and $\mathit{inf\text{-}axioms}(L)$ are co-r.e.-hard. For recursive classes $L \subseteq \mathrm{FO}$ it actually suffices to find a *semi-conservative* reduction, i.e. a reduction from FO to L which maps finitely satisfiable formulae to finitely satisfiable ones and unsatisfiable formulae to unsatisfiable ones. A general recursion-theoretic argument then implies that L is a conservative reduction class (see [8, p.37f] for details).

Let $\mathcal{K}$ be a class of structures, and L a class of formula. The L-theory of $\mathcal{K}$, abbreviated $\mathrm{Th}_L(\mathcal{K})$ is the set of all L-sentences that are true in all structures of $\mathcal{K}$. Further, for any theory $T \subseteq L$, we write T_{fin} for the class of L-sentences that hold in all *finite* models of T and $\overline{T_{\mathrm{fin}}}$ for the complement of T_{fin} in L, i.e. for the set of all sentences that are false in some finite model of T.

Definition 3.6 A theory T is *strongly undecidable* if T and $\overline{T_{\text{fin}}}$ are recursively inseparable.

Lemma 3.7 *Let L be a formula class that contains* FO^2 *and is closed under conjunction. If there exists a finite L-axiomatization of the class $\mathcal{K}$ of all models of a strongly undecidable* FO^2*-theory T then* fin-sat(L) *and* non-sat(L) *are recursively inseparable.*

Proof. Suppose that $\alpha \in L$ axiomatizes $\mathcal{K}$. If there exist a recursive set $X \subseteq L$ that separates *fin-sat*(L) from *non-sat*(L), then the set $\{\psi \in \mathrm{FO}^2 : \alpha \wedge \neg\psi \in X\}$ is also recursive and separates $\overline{T_{\text{fin}}}$ from T. $\square$

3.2 Domino problems and grids

Domino or tiling problems provide a simple and powerful method for proving undecidability results. They were introduced in the early 1960s by Wang as a tool to show the unsolvability of the $\forall\exists\forall$-prefix class in the pure predicate calculus. In the last thirty years they have been used to establish many undecidability results and lower complexity bounds for various systems of propositional logic, for subclasses of first-order logic and for decision problems in mathematical theories. The original, 'unconstrained' version of a domino problem is given by a finite set of dominoes or tiles, each of them an oriented unit square with coloured edges. The question is whether it is possible to cover the first quadrant in the Cartesian plane by copies of these tiles, without holes and overlaps, such that adjacent dominoes have matching colours on their common edge. The set of tiles is finite, but there are infinitely many copies of each tile available; rotation of the tiles is not allowed. Variants of this problem require that certain places (e.g. the origin, the bottom row or the diagonal) are tiled by specific tiles. A slightly more convenient definition is the following.

Definition 3.8 A *domino system* $\mathcal{D}$ is a triple (D, H, V) where D is a finite set of dominoes and $H, V \subseteq D \times D$ are two binary relations. Let S be any of the spaces $\mathbb{Z} \times \mathbb{Z}$, $\mathbb{N} \times \mathbb{N}$ or $\mathbb{Z}/s\mathbb{Z} \times \mathbb{Z}/t\mathbb{Z}$.[5] We say that $\mathcal{D}$ *tiles* S if there exists a *tiling* $\tau : S \to D$ such that for all $(x, y) \in S$:

(i) if $\tau(x, y) = d$ and $\tau(x+1, y) = d'$ then $(d, d') \in H$;
(ii) if $\tau(x, y) = d$ and $\tau(x, y+1) = d'$ then $(d, d') \in V$.

We are also interested in periodic solutions of domino problems.

[5] $\mathbb{Z}/s\mathbb{Z}$ stands for $\{0, \dots, s-1\}$ with successor modulo s; this structure is isomorphic with the standard s-cycle.

Definition 3.9 A domino system $\mathcal{D}$ is said to admit a *periodic tiling* if there is a tiling τ of $\mathbb{Z} \times \mathbb{Z}$ by $\mathcal{D}$ that has a horizontal and a vertical period $s, t > 0$ respectively. This means that for all points $(x, y) \in \mathbb{Z} \times \mathbb{Z}$ we have that $\tau(x, y) = \tau(x + s, y) = \tau(x, y + t)$.

A periodic tiling with periods s, t may be pictured as a tiling of a torus $\mathbb{Z}/s\mathbb{Z} \times \mathbb{Z}/t\mathbb{Z}$ obtained from gluing an $s \times t$ rectangle along the edges.

Berger [6] proved that the domino problem is undecidable. Gurevich and Koryakov [18] strengthened this to an inseparability result.

Theorem 3.10 (Berger, Gurevich-Koryakov) *The set of domino systems that admit, respectively, no tiling and a periodic tiling are recursively inseparable.*

For a new proof of this theorem we refer to [8, Appendix A]. The proof shows that one can effectively associate with every first-order sentence ψ a domino system $\mathcal{D}$ which tiles $\mathbb{N} \times \mathbb{N}$ periodically if ψ has a finite model, and which admits no tiling of either $\mathbb{N} \times \mathbb{N}$ or $\mathbb{Z} \times \mathbb{Z}$ if ψ is unsatisfiable. It follows that a formula class X is a conservative reduction class if there exists a recursive function that associates with every domino system $\mathcal{D}$ a formula $\psi_{\mathcal{D}} \in X$ such that:

(i) If $\mathcal{D}$ admits a periodic tiling then $\psi_{\mathcal{D}}$ has a finite model.

(ii) If $\mathcal{D}$ does not tile $\mathbb{N} \times \mathbb{N}$ then $\psi_{\mathcal{D}}$ is unsatisfiable.

Local grids. Two-dimensional grids form the basis of reductions from domino problems. In particular, let $\mathcal{G}_m$ denote the finite standard grid $(\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}, h, v)$ with horizontal and vertical successor functions

$$\begin{aligned} h(x, y) &:= (x + 1, y) \\ v(x, y) &:= (x, y + 1) \end{aligned}$$

In fact it will be sufficient to deal with sufficiently rich classes of structures that locally resemble grids. This is made precise in the following definition.

Definition 3.11 A *rich class of local grids* is a class $\mathcal{C}$ of structures $\mathfrak{A} = (A, h, v)$ with the following properties.

(i) For each structure $\mathfrak{A} = (A, h, v)$ in $\mathcal{C}$, h and v are two unary functions such that $h \circ v = v \circ h$.

(ii) For every $r \in \mathbb{N}$ there exists a multiple $m = kr$ such that $\mathcal{C}$ contains the standard grid $\mathcal{G}_m$.

Let $\mathcal{K}$ be a class of structures that is closed under expansions. That is, $\mathcal{K}$ is defined by imposing semantic conditions on some relations, but is closed under arbitrary variations and additions of other relations. (For instance, consider the case where we have a built-in linear order, or several built-in equivalence relations, but no conditions on other predicates). To prove that the FO^2-theory of $\mathcal{K}$ is strongly undecidable it suffices to present an FO^2-interpretation of some rich class of local grids in $\mathcal{K}$. The original notion of a (first-order) interpretation is due to Tarski [36]. Today, in model theory, interpretations come in many different shapes and sizes (see e.g. [20, Chapter 5]). We use here a specific variant tailored for our particular class of applications. The FO^2-interpretations that we need are given by sequences $I = \langle \delta(x), \varepsilon(x,y), \varphi_h(x,y), \varphi_v(x,y) \rangle$ of FO^2-formulae in the vocabulary of $\mathcal{K}$. Let $\mathfrak{B} \in \mathcal{K}$. I interprets in $\mathfrak{B}$ a structure $I(\mathfrak{B}) = (A, h, v)$ with two unary functions provided that the following conditions are satisfied:

(1) $\mathfrak{B} \models \exists x \delta(x)$

(2) Let $\delta^{\mathfrak{B}} := \{b : \mathfrak{B} \models \delta(b)\}$. The formula $\varepsilon(x,y)$ defines an equivalence relation on $\delta^{\mathfrak{B}}$ that is compatible with φ_h and φ_v. In other words, $\varepsilon^{\mathfrak{B}}$ is a congruence relation on the induced structure $(\delta^{\mathfrak{B}}, \varphi_h^{\mathfrak{B}}, \varphi_v^{\mathfrak{B}})$. We write $[b]$ to denote the congruence class of an element $b \in \delta^{\mathfrak{B}}$. The set of these congruence classes is the universe of $I(\mathfrak{B})$.

(3) The relations defined by φ_h and φ_v on the quotient structure $(\delta^{\mathfrak{B}}, \varphi_h^{\mathfrak{B}}, \varphi_v^{\mathfrak{B}})/\varepsilon^{\mathfrak{B}}$ are the graphs of two unary functions h and v. In other words, for every congruence class $[b]$ there exists precisely one congruence class $[c]$ such that $\mathfrak{B} \models \varphi_h(b,c)$ (and hence $\mathfrak{B} \models \varphi_h(b',c')$ for all $b' \in [b], c' \in [c]$). Similarly for φ_v.

The conditions (1)-(3) are the *admissibility conditions* of I on $\mathfrak{B}$. They are necessary and sufficient for I to define in $\mathfrak{B}$ a structure $I(\mathfrak{B})$ with two unary functions h, v.

Definition 3.12 Let $\mathcal{K}$ a class of structures that is closed under expansions, and $\mathcal{C}$ be a rich class of local grids. We say that FO^2 *interprets* $\mathcal{C}$ *in* $\mathcal{K}$ if there exists an FO^2-interpretation $I = \langle \delta(x), \varepsilon(x,y), \varphi_h(x,y), \varphi_v(x,y) \rangle$ such that

(i) On every structure $\mathfrak{B} \in \mathcal{K}$, the admissibility conditions for I are satisfied and $I(\mathfrak{B})$ is a local grid in $\mathcal{C}$.

(ii) For every finite local grid $\mathfrak{A} \in \mathcal{C}$ there exists a finite $\mathfrak{B} \in \mathcal{K}$ such that $\mathfrak{A} \cong I(\mathfrak{B})$.

Remark. It is sometimes natural to think of a rich class of local grids being interpreted in a finitely FO^2-axiomatizable subclass of $\mathcal{K}$. This means that there exists a sentence $\psi \in \mathrm{FO}^2$ such that only the structures $\mathfrak{B} \in \mathcal{K}$ with $\mathfrak{B} \models \psi$ appear in conditions *(i)* and *(ii)* above. Suppose that $\mathcal{C}$ is a rich class of local grids that contains the trivial grid $\mathcal{G}_1$ with just one node. Note that if FO^2 interprets $\mathcal{C}$ in a finitely FO^2-axiomatizable subclass of $\mathcal{K}$, then FO^2 also interprets $\mathcal{C}$ in $\mathcal{K}$. Indeed suppose that $I = \langle \delta(x), \varepsilon(x,y), \varphi_h(x,y), \varphi_v(x,y) \rangle$ interprets $\mathcal{C}$ in the class axiomatized by ψ. Let I' be the interpretation obtained by replacing each formula η of I by $\psi \to \eta$. Then $I'(\mathfrak{B}) = I(\mathfrak{B})$ for $\mathfrak{B} \models \psi$ and $I'(\mathfrak{B})$ is the trivial grid $\mathcal{G}_1$ for $\mathfrak{B} \models \neg\psi$.

Theorem 3.13 *Suppose that* FO^2 *interprets a rich class of local grids in* $\mathcal{K}$. *Then the* FO^2*-theory of* $\mathcal{K}$ *is strongly undecidable.*

Proof. Let T be the FO^2-theory of $\mathcal{K}$ and suppose that the FO^2-interpretation $I = \langle \delta(x), \varepsilon(x,y), \varphi_h(x,y), \varphi_v(x,y)\rangle$ interprets a rich class of local grids in $\mathcal{K}$. Given a domino system $\mathcal{D} = (D, H, V)$, let $(P_d : d \in D)$ be a collection of new monadic predicates that do not appear in I. Let $\psi_{\mathcal{D}}$ be the conjunction of the following FO^2-sentences:

$$\forall x(\delta(x) \to \bigwedge_{\substack{d,d' \in D\\ d \neq d'}} \neg(P_d x \wedge P_{d'} x))$$

$$\forall x \forall y(\delta(x) \wedge \delta(y) \wedge \varepsilon(x,y) \to \bigwedge_{d \in D} (P_d x \leftrightarrow P_d y))$$

$$\forall x \forall y(\delta(x) \wedge \delta(y) \wedge \varphi_h(x,y) \to \bigvee_{(d,d') \in H} (P_d x \wedge P_{d'} y))$$

$$\forall x \forall y(\delta(x) \wedge \delta(y) \wedge \varphi_v(x,y) \to \bigvee_{(d,d') \in V} (P_d x \wedge P_{d'} y)).$$

Towards a contradiction, assume that some recursive set X separates $\overline{T_{\text{fin}}}$ from T. Then the set $Y := \{\mathcal{D} : \neg\psi_{\mathcal{D}} \in X\}$ is also recursive. We claim that Y separates the domino systems that admit a periodic tiling from those that admit no tiling of $\mathbb{N} \times \mathbb{N}$.

To see this, let first $\mathcal{D}$ be a domino system that tiles $\mathbb{N} \times \mathbb{N}$ periodically. Then there exists an $r \in \mathbb{N}$ such that $\mathcal{D}$ tiles the grid $\mathcal{G}_r$, and hence also the grids $\mathcal{G}_{kr}$, for all k. Every rich class of local grids contains at least one of the grids $\mathcal{G}_{kr}$, so there exists a finite model $\mathfrak{B} \in \mathcal{K}$ such that $I(\mathfrak{B})$ is (isomorphic to) $\mathcal{G}_{kr}$. Recall that the elements of $I(\mathfrak{B})$ are congruence classes $[b]$ of elements b with $\mathfrak{B} \models \delta(b)$ modulo the congruence defined by $\varepsilon(x,y)$. Take a correct tiling τ of $I(\mathfrak{B})$ by the domino system $\mathcal{D}$, and let $\mathfrak{B}'$ be the expansion of $\mathfrak{B}$ by the predicates $P_d := \{\mathfrak{B} \models \delta(x) : \tau([b]) = d\}$, i.e. the set of elements whose equivalence classes are tiled by d. Since $\mathcal{K}$ is closed under expansions, $\mathfrak{B}'$ belongs to $\mathcal{K}$ and $\mathfrak{B}' \models \psi_{\mathcal{D}}$. Hence $\neg\psi_{\mathcal{D}} \in \overline{T_{\text{fin}}} \subseteq X$ and thus $\mathcal{D} \in Y$.

Second, suppose that $\mathcal{D}$ does not tile $\mathbb{N} \times \mathbb{N}$. We claim that in this case $\neg\psi_{\mathcal{D}} \in T \subseteq \overline{X}$ and hence $\mathcal{D} \in \overline{Y}$. Otherwise there exists a model $\mathfrak{B} \in \mathcal{K}$ with $\mathfrak{B} \models \psi_{\mathcal{D}}$. But then $I(\mathfrak{B})$ is a local grid with commuting unary functions h and v. Moreover, the second clause of $\psi_{\mathcal{D}}$ asserts that the predicates P_d are compatible with $\varepsilon(x,y)$. Thus we have a well-defined expansion $\mathfrak{A} = (I(\mathfrak{B}), (P'_d)_{d \in D})$ of $I(\mathfrak{B})$ where P'_d is the quotient of P_d modulo ε. Take any element a of $\mathfrak{A}$ and define a tiling $\tau : \mathbb{N} \times \mathbb{N} \to D$ by

$$\tau(i,j) = d \ \text{ iff } \mathfrak{A} \models P'_d(h^i v^j a).$$

Since $\mathfrak{B} \models \psi_{\mathcal{D}}$ this mapping is well-defined and provides a correct tiling, contradicting the assumption that $\mathcal{D}$ does not tile $\mathbb{N} \times \mathbb{N}$. □

3.3 Several equivalence relations

Theorem 3.14 *The* FO^2*-theory of the class of structures with several built-in equivalence relations is strongly undecidable.*

Proof. We describe a class $\mathcal{K}$ of structures that is finitely FO^2-axiomatizable inside the class of all structures with four built-in equivalence relations, and show $\mathcal{K}$ interprets a rich class of local grids. The result follows by Theorem 3.13. For this purpose we use for each $(i, j) \in \{0,1\} \times \{0,1\}$

- an equivalence relation E_{ij},
- a unary relation A_{ij},
- binary relations $H_{ij}, V_{ij}, D^+_{ij}, D^-_{ij}$.

The idea is that the grids $\mathcal{G}_m$ (for even numbers $m \in \mathbb{N}$) are described by structures of this type as follows:

- A_{ij} contains the points (u, v) such that $u \equiv i \mod 2$ and $v \equiv j \mod 2$.
- Each E_{ij}-equivalence class consists of a square

$$\{(u,v),(u+1,v),(u,v+1),(u+1,v+1)\}$$

  such that the lower left corner (u, v) belongs to A_{ij}.
- H_{ij} and V_{ij} connect the points in A_{ij} with their right and upper neighbours in the grid, respectively.
- D^+_{ij} and D^-_{ij} describe the diagonals in the E_{ij}-equivalence classes. That is, each point $(u, v) \in A_{ij}$ has an outgoing D^+_{ij}-edge to $(u+1, v+1)$ and there is a D^-_{ij}-edge from $(u+1, v)$ to $(u, v+1)$.

We need an FO^2-axiom ψ that enforces the following conditions:

(1) The universe is the disjoint union of A_{00}, A_{01}, A_{10} and A_{11}.
(2) Let $i' := 1 - i$ and $j' := 1 - j$. Every point in A_{ij} has outgoing edges labeled H_{ij}, V_{ij}, D^+_{ij} and $D^-_{i'j}$, incoming edges labeled $H_{i'j}, V_{ij'}, D^+_{i'j'}$ and $D^-_{ij'}$ and no incoming or outgoing H-, V- or D- edges of other kinds.
(3) Each E_{ij} is the reflexive and symmetric closure of the disjoint union of H_{ij}, $H_{ij'}$, V_{ij}, $V_{i'j}$, D^+_{ij} and D^-_{ij}.

It is easy to see that the conditions (1)–(3) can indeed be expressed in FO^2. We claim that a rich class of local grids is interpreted in the class $\mathcal{K}$ axiomatized by ψ (with the additional constraint that the E_{ij} be equivalence relations) via the interpretation $I = \langle \delta(x), \varepsilon(x,y), \varphi_h(x,y), \varphi_v(x,y) \rangle$ with

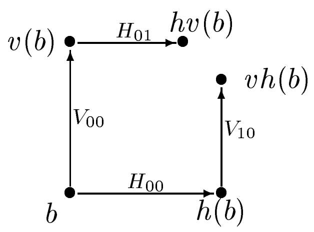
$$\begin{aligned}
\delta(x) &:= (x = x) \\
\varepsilon(x,y) &:= (x = y) \\
\varphi_h(x,y) &:= \bigvee\nolimits_{ij} H_{ij}xy \\
\varphi_v(x,y) &:= \bigvee\nolimits_{ij} V_{ij}xy.
\end{aligned}$$

First, it is obvious, that every finite grid $\mathcal{G}_m$ with even m is interpreted by I in some model of ψ. Just take the expansion of $\mathcal{G}_m$ itself with the intended meaning of the relations $A_{ij}, E_{ij}, \dots$ as described above.

On the other side, let $\mathfrak{B}$ be a model of ψ such that the E_{ij} are equivalence relations on the universe B. We claim that $I(\mathfrak{B})$ is a local grid. We thus have to prove that the formulae $\varphi_h(x,y)$ and $\varphi_v(x,y)$ define commuting functions on B.

Functionality: Let $b \in A_{00}$. By condition (2) there exists an element c such that $\mathfrak{B} \models H_{00}bc$ and hence $\mathfrak{B} \models \varphi_h(b,c)$. Suppose that there exists a $c' \neq c$ such that also $\mathfrak{B} \models \varphi_h(b,c')$. Since outgoing H-edges from nodes in A_{00} are H_{00}-edges, it follows that both $c, c' \in A_{10}$. Further, by (3), c and c' belong to the same E_{00}-equivalence class as b. Hence c and c' have to be connected by an edge labeled by H_{00}, H_{01}, V_{00}, V_{10}, D^+_{00} or D^-_{00}. But this is impossible since both c and c' belong to A_{10}. The arguments for elements of the other A_{ij} and for φ_v are completely analogous. Thus φ_h and φ_v do indeed define functions h and v on B

Commutativity of h and v: Let $b \in A_{00}$. We claim that $hv(b) = vh(b)$. Otherwise we would have the following situation:



Again, since E_{00} is the reflexive and symmetric closure of the union of H_{00}, H_{01}, V_{00}, V_{10}, D^+_{00} and D^-_{00}, it follows that $hv(b)$ and $vh(b)$ belong to the same E_{00}-equivalence class as b, and hence have to be connected by an edge of one of these kinds. But this is impossible since both $hv(b)$ and $vh(b)$ belong to A_{11}. Again the arguments for b in other A_{ij} are analogous. $\square$

3.4 Transitive closure, well-founded relations, and fixed points

The (strong) undecidability of the two-variable theory of several equivalence relations implies a number of further undecidability results. In particular this is the case for the systems TC^2 and FP^2.

For the two-variable transitive closure logic TC^2 undecidability follows immediately because equivalence relations can be axiomatized in TC^2.

Corollary 3.15 (Grädel, Otto, Rosen) *The satisfiability problem and the finite satisfiability problem for* TC^2 *are undecidable.*

In fact $sat(\mathrm{TC}^2)$ is even Σ_1^1-hard [15].

For two-variable fixed-point logic, a little more work is required. We actually prove the undecidability for a weaker system, that just extends FO^2 by *well-foundedness assertions* about compositions of binary relations.

For binary relations R, T we use the notation $R \circ T$ to denote their composition, defined by

$$R \circ T := \{(x, z) : \exists y(Rxy \wedge Tyz)\}.$$

Notice that we can express in FP^2 the well-foundedness of $R \circ T$ by the sentence

$$\forall x[\mathrm{LFP}_{W,x} \forall y(Tyx \to \forall x(Rxy \to Wx))](x)$$

even though the composition $R \circ T$ itself is clearly *not* definable from R and T in FO^2 or even $L^2_{\infty\omega}$. The well-foundedness of any finite composition $R_1 \circ R_2 \circ \cdots \circ R_m$ of binary relations is expressible in FP^2 by a very similar sentence.

Definition 3.16 Let $(\mathrm{FO} + \mathrm{WF}^\circ)$ be the fragment of FP^2 consisting of conjunctions of FO^2-sentences with well-foundedness assertions about compositions (of any finite length) of basic relations.

Recall that a pre-well-ordering is a well-founded, irreflexive and transitive relation $<$, for which the induced relation $x \sim y \Leftrightarrow (\neg x < y \wedge \neg y < x)$ is a congruence. Equivalently, think of $\sim$ as a congruence for $<$, such that the quotient $< / \sim$ is a well-ordering.

Lemma 3.17 *The class of structures* $\mathfrak{A} = (A, <, \sim)$*, such that* $<$ *is a pre-well-ordering and* $\sim$ *is the induced equivalence relation, is axiomatizable in* $(\mathrm{FO}^2 + \mathrm{WF}^\circ)$*, even without equality.*

Proof. $\mathfrak{A} = (A, <, \sim)$ is a pre-well-ordering if and only if the following three conditions are satisfied.

(i) $\mathfrak{A} \models \forall x \forall y \big(x \sim y \leftrightarrow \neg(x < y \lor y < x)\big)$.

(ii) $\mathfrak{A} \models \forall x\, x \sim x$.

(iii) The composition $\sim \circ < \circ \sim$ is well-founded.

For the obvious direction note that, if $(A, <, \sim)$ is a pre-well-ordering, then $\sim \circ < \circ \sim = <$, so that (iii) is clearly satisfied. Conversely, assume that $(A, <, \sim)$ satisfies (i) – (iii). By (i), $\sim$ is symmetric, and $(A, <, \sim)$ satisfies a trichotomy property: $\forall x \forall y \big(x < y \lor y < x \lor x \sim y\big)$. As $\sim$ is reflexive, (iii) clearly implies in particular that $<$ itself (as well as $< \circ \sim$ and $\sim \circ <$) is well-founded, as these compositions are included in $\sim \circ < \circ \sim$. This implies that the trichotomy is strict in the sense that the alternative is exclusive. It suffices to show now that $<$ is transitive and closed under $\sim$ on the left and on the right ($\sim \circ < = < \circ \sim = <$; i.e. $\sim$ is a congruence for $<$).

Transitivity. Suppose that $a_1 < a_2 < a_3$. By trichotomy, it suffices to exclude the possibilities $a_3 < a_1$ (which would violate well-foundedness of $<$) and $a_1 \sim a_3$ (which would violate well-foundedness of $< \circ \sim$ by $a_1 < a_2 \sim a_2 < a_3 \sim a_1$).

Invariance under $\sim$. Suppose, for instance, that $a_1 < a_2 \sim a_3$. We want to show that $a_1 < a_3$, again by excluding the other possibilities: $a_3 < a_1$ would violate well-foundedness of $< \circ \sim$, and $a_1 \sim a_3$ would violate well-foundedness of $\sim \circ < \circ \sim$. $\square$

Since the equivalence relation induced by a pre-well-ordering $<$ is FO^2-definable from $<$ and, conversely, every equivalence relation is induced by a pre-well-ordering (choose a well-ordering of the equivalence classes), we infer the following result.

Proposition 3.18 *The* FO^2*-theory of several built-in pre-well-orderings is strongly undecidable.*

Corollary 3.19 $(\mathrm{FO}^2 + \mathrm{WF}^\circ)$*, and hence also* FP^2*, are conservative reduction classes and hence undecidable for satisfiability and finite satisfiability.*

Again the satisfiability problems for these systems are actually Σ^1_1-hard. A direct application of Lemma 3.17 moreover shows that also the fragment of FP^2 that corresponds to the extension of L_μ by universal FO^2-sentences without equality is undecidable [23].

Another interesting family of extensions of L_μ are the k-dimensional μ-calculi. They have been introduced by Otto [28] who shows that these languages can express precisely those properties of Kripke structures that are invariant under bisimulation and decidable in polynomial time. Unfortunately, these languages do not inherit the nice algorithmic properties of L_μ: already the satisfiability of the two-dimensional μ-calculus is highly undecidable [28].

3.5 Well-orderings and CL^2

The following problems are recursively equivalent (both in their general and in their finitistic versions):

- the FO^2-theory of several built-in well-founded relations;
- the FO^2-theory of several built-in well-orderings;
- the satisfiability problem for CL^2.

That well-orderings and arbitrary well-founded relations carry the same power for the issue of FO^2-satisfiability is a consequence of the fact that

(i) E is well-founded if and only if there is a well-ordering $<$ such that $E \subseteq <$.

(ii) a well-founded relation $<$ is a well-ordering if and only if it satisfies the FO^2-axiom of trichotomy $\forall x \forall y \big(x = y \vee x < y \vee y < x\big)$.

It remains to link CL^2 with FO^2 over well-founded relations. In one direction this connection easily follows from the fact that well-foundedness of a relation E is expressible in CL^2 through $\forall x \big([Eyx]^\infty \bot\big)$, corresponding to the universally quantified, generalized CTL-formula $\forall(E^{-1}\text{-paths}\, p)(\top \,\mathrm{until}\, \bot)$.

In the other direction we have to show that there is a recursive reduction of CL^2-satisfiability to the FO^2-theory of several well-founded relations. Skipping the usual Skolemization procedure which introduces new predicates for subformulae, we need only consider CL^2-formulae $\psi_1(x) = \langle \xi \rangle^\infty \varphi$ and $\psi_2(x) = [\xi]^\infty \varphi$, where $\xi(x,y) = Exy$ and $\varphi(x) = Px$ are atomic, i.e.

$$\psi_1(x) \equiv \Diamond^\infty P \equiv \mu_X \big(P \vee \Diamond X\big) \qquad \text{and} \qquad \psi_2(x) \equiv \Box^\infty P \equiv \mu_X \big(P \vee \Box X\big).$$

It suffices to find FO^2-sentences θ_i in an extended vocabulary with new monadic predicates S, T, and binary E_1, E_2, such that:

(a) every structure $\mathfrak{B} = (B, E, P)$ can be expanded to a model of $\theta_1 \wedge \theta_2$.

(b) all structures in the extended vocabulary that interpret the E_i as well-founded relations satisfy $\theta_1 \longrightarrow \forall x \big(\psi_1(x) \leftrightarrow Tx\big)$ and $\theta_2 \longrightarrow \forall x \big(\psi_2(x) \leftrightarrow Sx\big)$.

Consider first ψ_1 over some structure $\mathfrak{B}$. Let $T \subseteq B$ be the set of vertices b such that $\mathfrak{B} \models \psi_1[b]$. Define a function $\mathrm{rk}\colon T \to \mathbb{N}$, where $\mathrm{rk}(b) = i$ if i is the minimal length of an E-path from b to a vertex in $P^{\mathfrak{B}}$. Let $E_1 \subseteq E^{-1}$ be the set of all pairs $(b, b') \in T^2$ for which $(b', b) \in E$ and $\mathrm{rk}(b') = \mathrm{rk}(b) + 1$. Note that E_1 is well-founded. Clearly $(\mathfrak{B}, T, E_1)$ satisfies

$$\begin{array}{rcl} \theta_1 & = & \forall x \forall y (E_1 xy \to Eyx) \\ & \wedge & \forall x \big[Tx \to \big(Px \vee \exists y (E_1 yx \wedge Ty)\big)\big] \\ & \wedge & \forall x \big[\neg Tx \to \big(\neg Px \wedge \forall y (Exy \to \neg Ty)\big)\big]. \end{array}$$

It is not hard to check that θ_1 is also as desired for (b) over all structures that interpret E_1 as a well-founded relation.

Turning to ψ_2, let S, in some $\mathfrak{B}$, be the set of vertices that satisfy ψ_2. Let $S = \bigcup_\alpha S_\alpha$, where the S_α are the stages of the least fixed point associated with ψ_2, i.e. $S_0 = P$, $S_{\alpha+1} = \{b : (\mathfrak{B}, S_\alpha) \models (\Box S_\alpha)[b]\}$, and $S_\lambda = \bigcup_{\alpha<\lambda} S_\alpha$ in limits λ. Let rk be the ordinal-valued rank function on S defined as $\mathrm{rk}(b) = \min\{\alpha : b \in S_\alpha\}$. Let E_2 be the set of all pairs $(b, b') \in S^2$ for which $\mathrm{rk}(b) < \mathrm{rk}(b')$. Clearly E_2 is well-founded, and also $E \cap S \times B \subseteq (E_2)^{-1}$. It follows that $(\mathfrak{B}, S, E_2)$ satisfies

$$\begin{array}{rcl} \theta_2 & = & \forall x \forall y (Exy \wedge Sx \to E_2 yx) \\ & \wedge & \forall x \big[Sx \to \big(Px \vee \forall y (E_2 yx \to Sy)\big)\big] \\ & \wedge & \forall x \big[\neg Sx \to \big(\neg Px \wedge \exists y (Exy \wedge \neg Sy)\big)\big]. \end{array}$$

Moreover, θ_2 is indeed as desired in (b), in forcing S to capture the semantics of ψ_2 over all structures in which E_2 is well-founded.

The following results about the behaviour of FO^2 over well-founded relations are quite recent and will be fully treated elsewhere [31].

Theorem 3.20 (Otto) *The* FO^2*-theory of one built-in well-ordering and the* FO^2*-theory of one built-in finite linear order are decidable in* Co-Nexptime. *On the other hand, the* FO^2*-theory of several built-in well-orderings is strongly undecidable.*

As indicated above, well-orderings may be replaced by arbitrary well-founded relations without affecting the statements of the last theorem.

Corollary 3.21 *The least common extension* CL^2 *of* FO^2 *and* CTL *is undecidable for satisfiability as well as for finite satisfiability.*

Remark. The observation that the embedding of propositional modal logic into FO^2 does not really explain the robustness and the nice algorithmic and model-theoretic properties of modal logics (see also [38] in this context) has lead to the study of another fragment of first-order logic, the so-called *guarded fragment*. Here the number of variables and the arities of the relation symbols are not restricted, but only a restricted form of quantification (relativized by atoms) is allowed. It seems that the guarded fragment indeed shares many of the nice properties of modal logics (see [1, 4, 5, 13]). In particular it has both the finite model property and (a generalized variant of) the tree model property.

4 Model checking in two variables

We have seen that the satisfiability problems for two-variable logics *with full first-order quantification* such as FO^2, TC^2, CL^2 or FP^2 are much harder (and indeed undecidable in most cases) than the satisfiability problems for corresponding *modal* logics such as ML, CTL or L_μ. Our point in this section is that for *model checking problems* the situation is different: in all cases that we consider, the model checking problem of a modal logic has essentially the same complexity as the model checking problem of the corresponding two variable logic with full quantification. In fact, we can even drop the restriction to two variables, and admit instead any bounded number of variables without a significant increase of complexity.

Complexity issues for model checking. The model checking problem for a logic L, denoted $\mathrm{MC}(L)$, is the following: Given a formula $\psi \in L$ and an appropriate finite structure $\mathfrak{A}$ (including, if necessary, constants interpreting the free variables of ψ), determine whether $\mathfrak{A} \models \psi$.

There are different possibilities to study the complexity of model checking problems. The general measure is the *combined complexity*. Here both the structure and the formula are considered as variable inputs for $\mathrm{MC}(L)$ and the complexity is measured in terms of the combined length of both inputs. But in many cases it makes sense to fix either the formula or the structure and measure the complexity just in terms of the other input. If the formula is fixed, and complexity is measured in terms of the length (essentially: the cardinality) of the structure, then we speak of the *structure complexity*[6] of $\mathrm{MC}(L)$. The structure complexity is meaningful and important because in many situations, the formula (i.e. the query or the specification) is rather short but the structure (the database to be queried or the program to be verified) may be huge. On the other hand, if the structure is fixed and only the formula varies, we speak of the *expression complexity*[7]. In more classical terms, the expression complexity of a logic L on a fixed structure $\mathfrak{A}$ is just the complexity of the L-theory of $\mathfrak{A}$.

Our default here will be the combined complexity. Whenever we consider the structure complexity or the expression complexity, this will be mentioned explicitly.

Aside Consider once more the semantically distinguished fields of two-variable logics, which were characterized in section 1.2 by invariance under bisimulation and two-pebble equivalence (with or without counting). It is a curious phenomenon that, within each of the corresponding maximal logics ML_∞, graded ML_∞, $L^2_{\infty\omega}$, and $C^2_{\infty\omega}$, recursive syntax can be given to fragments that are *semantically complete* for PTIME structure complexity within that field. These *capturing results* point at a very special nature of the two-variable scenarios also with respect to descriptive complexity [29]. PTIME within the bisimulation invariant, modal world, for instance, is captured by a higher-dimensional variant of L_μ [28].

[6] In the context of database application or automatic verification one also uses the terms *data complexity* or *program complexity*.

[7] or *specification complexity* in automatic verification.

4.1 Model checking for modal logic and FO^k.

The model checking problem for first-order logic is well known to be PSPACE-complete. We explain this result and draw some conclusions for modal logic and the bounded-variable fragments of first-order logic.

We assume that the reader is familiar with the notion of an alternating algorithm and we will use the facts that ALOGSPACE = P and APTIME = PSPACE [8].

Without loss of generality, we can restrict attention to formulae in positive normal form. This means that all negations are driven inwards and occur only in front of atomic formulae. Recall that a literal is an atomic formula or its negation. To check whether $\mathfrak{A} \models \psi(a_1, \ldots, a_m)$ for a given first-order formula $\psi(x_1, \ldots, x_m)$, a finite structure $\mathfrak{A}$, and elements $a_1, \ldots, a_m$ of $\mathfrak{A}$, we use the obvious recursive algorithm (or equivalently, the first-order model checking game, see below). It is instructive to describe it as an alternating procedure:

ModelCheck$(\psi, \mathfrak{A}, a_1, \ldots, a_m)$

Input: a first-order formula $\psi(x_1, \ldots, x_m)$,
 a finite structure $\mathfrak{A}$,
 a tuple $\bar{a} = (a_1, \ldots, a_m)$ of elements of $\mathfrak{A}$
if ψ is a literal **then**
 if $\mathfrak{A} \models \psi(\bar{a})$ **accept else reject**
if $\psi = \eta \vee \vartheta$ **then do**
 existentially guess $\varphi \in \{\eta, \vartheta\}$
 ModelCheck$(\varphi, \mathfrak{A}, \bar{a})$
if $\psi = \eta \wedge \vartheta$ **then do**
 universally choose $\varphi \in \{\eta, \vartheta\}$
 ModelCheck$(\varphi, \mathfrak{A}, \bar{a})$
if $\psi = \exists x_j \varphi$ **then do**
 existentially guess an element a of $\mathfrak{A}$
 ModelCheck$(\varphi, \mathfrak{A}, \bar{a}\frac{a}{j})$
if $\psi = \forall x_j \varphi$ **then do**
 universally choose an element a of $\mathfrak{A}$
 ModelCheck$(\varphi, \mathfrak{A}, \bar{a}\frac{a}{j})$

Here $\bar{a}\frac{a}{j}$ is the tuple obtained from $\bar{a}$ by changing the j-th component to a or adding a as j-th component.

As every alternating procedure, this algorithm can be described more intuitively as a game between the existential and the universal player. In this case the game is the obvious model checking game that can also be used to define the semantics of first-order logic. On input $(\psi, \mathfrak{A}, \bar{a})$ the positions of the games are pairs $(\varphi, \bar{b})$ where φ is a

[8] For background on alternating complexity classes, see for instance [2, Chapter 3] or [33].

subformula of ψ and $\bar{b}$ is a tuple of elements from $\mathfrak{A}$. The initial position is $(\psi, \bar{a})$. At position $(\varphi, \bar{b})$, the type of φ determines the kind of move to be played next: if φ is a disjunction, the existential player selects one of the disjuncts; if φ is a conjunction, the universal player selects one of the conjuncts; if the formula starts with an existential or universal quantifier, the corresponding player selects an element of the structure and modifies the tuple $\bar{b}$ accordingly. The final positions are the pairs $(\varphi, \bar{b})$ where φ is a literal. At such a position, the existential player wins if $\varphi(\bar{b})$ is true in the given structure $\mathfrak{A}$, otherwise the universal player wins. The procedure accepts if and only if the existential player has a winning strategy (which is true if and only if $\mathfrak{A} \models \psi(\bar{a})$).

The alternating model checking procedure runs in time $O(|\psi| \log n)$ and uses work space at most $r \log n + \log |\psi|$, where n is the cardinality of $\mathfrak{A}$ and r is the maximal number of free variables in any subformula of ψ. Indeed, the structure $\mathfrak{A}$ is never modified, so in any situation the procedure needs at most $r \log n$ bits to describe the current tuple and a pointer of length $\log |\psi|$ to specify the current subformula of ψ. Together with the facts that alternating polynomial time coincides with deterministic polynomial space and that alternating logspace coincides with deterministic polynomial time, the following results follow immediately.

Proposition 4.1 *The model checking problem for* FO *is in* PSPACE. *For every fixed k,* MC(FO^k) *is in* P.

It is a trivial consequence of the PSPACE-completeness of quantified propositional logic that the first-order theory of any structure with at least two elements is PSPACE-hard. Thus the expression complexity and hence also the combined complexity of first-order model checking is PSPACE-complete. Vardi [37] has proved that the model checking problems for FO^k are P-complete, for all $k \geqslant 3$. We have a simple proof that the same holds for FO^2, and in fact also for propositional modal logic ML.

Proposition 4.2 *The model-checking problem for* ML *is* P*-complete. As a consequence, the same holds for* FO^k *for all $k \geqslant 2$.*

Proof. ML is a sublogic of FO^2 and we have already seen that the model checking problem of FO^k is in P for every fixed k.

To prove hardness, we present a reduction from the GAME problem (see Example 1.4) to the model-checking problem for ML. Recall that an instance of the GAME problem is a Kripke structure $\mathfrak{A} = (A, E)$ with a element a, and it is asked whether player **I** has a winning strategy for the two-player game on board $\mathfrak{A}$ with one pebble and the following rules: Player **I** begins with the pebble at position a. The players alternate; in each move they bring the pebble from its current position along some E-edge to a next position. Who gets stuck first, loses the game (the opponent wins).

We define a sequence of propositional modal formulae by

$$\varphi_1 := \Diamond\Box\bot, \qquad \varphi_{i+1} := \Diamond\Box\varphi_i.$$

Clearly $\mathfrak{A}, a \models \varphi_i$ if and only if player **I** has a strategy to win the game from a in at most i moves. Further, if player **I** has a winning strategy, then she also has a winning strategy in at most n moves, where n is the total number of positions in the game.

Thus, the function taking $G = (\mathfrak{A}, a)$ to the pair (G, φ_n) is a logspace reduction from GAME to MC(ML). $\Box$

Structure and expression complexity. The alternating procedure for first-order model checking runs in alternating logarithmic time for fixed formulae ψ. A more sophisticated argument shows that also the expression complexity for FO^k-model checking is in ALOGTIME (see [37]).

4.2 Polynomial-time model checking for extensions of FO^k

We next consider the logics C^k, the extensions of FO^k by counting quantifiers $\exists^{\geqslant m}$. Recall that C^k is a fragment of first-order logic but there exists no function $f : \mathbb{N} \to \mathbb{N}$ such that $C^k \subseteq \mathrm{FO}^{f(k)}$. With an appropriate modification of the model checking game for FO^k we can show that also the C^k admit efficient model checking algorithms.

Theorem 4.3 *For every $k \geqslant 2$, the model checking problem for C^k is* P*-complete.*

Proof. We first observe that we can easily reduce the model checking problem for C^k to the case where the parameters m in the counting quantifiers $\exists^{\geqslant m}$ and $\exists^{\leqslant m}$ are bounded by the cardinality of the given structure. Indeed, given a formula $\psi(\bar{x}) \in C^k$, and a structure $\mathfrak{A}$ of cardinality n, replace for all $m > n$, the subformulae $\exists^{\geqslant m} x\varphi$ of ψ by $\bot$ and the subformulae $\exists^{\leqslant m} x\varphi$ by $\top$. Further we assume, without loss of generality, that the universe of the given structure $\mathfrak{A}$ is of the form $A = \{1, \ldots, n\}$, and we slightly extend C^k to allow expressions of the form $(\exists^{\geqslant m} x_j > a)$ and $(\exists^{\leqslant m} x_j > a)$ where $a \in A \cup \{0\}$. The value $a = 0$ is admitted so that we can rewrite statements $\exists^{\geqslant m} x_j \varphi$ with unrelativized counting quantifiers by relativized quantifiers $(\exists^{\geqslant m} x_j > 0)\varphi$. We can thus assume that all counting quantifiers are relativized.

We extend the alternating procedure $\mathbf{ModCheck}(\varphi, \mathfrak{A}, \bar{a})$ given in the previous section by rules applying to subformulae of the form $(\exists^{\geqslant m} x_j > a)\varphi$ or $(\exists^{\leqslant m} x_j > a)\varphi$. We use the game description of the model checking procedure: at a position given by a formula $(\exists^{\geqslant m} x_j > a)\varphi$ and a tuple $\bar{b}$ the existential player selects a value $c > a$ for x_j. The universal player now has two options (if $m \geqslant 2$): she can either challenge c or accept c. To challenge c means that she moves to the position $(\varphi, \bar{b}\frac{c}{j})$. Thus, in the rest of the game, the existential player has to prove that c was a valid choice for x_j. But (unless

$m = 1$) the universal player also has the option to accept c and to force her opponent to produce a next value for x_j. In that case c becomes the new value of a (i.e. perform the update $a := c$) and the game proceeds at the position given by $((\exists^{\geqslant m-1} x_j > a)\varphi, \bar{b})$. The rules for formulae $(\exists^{\leqslant m} x_j > a)\varphi$ are analogous.

It is clear that the existential player has a winning strategy for the modified game **ModCheck**$(\psi, \mathfrak{A}, \bar{a})$ if and only if $\mathfrak{A} \models \psi(\bar{a})$. Further, the game only requires logarithmic space. Indeed only one relativized quantifier $(\exists^{\geqslant m} x_j > a)$ is treated at a time, so, compared to the model checking game for FO^k, only two additional variables storing the current values of m and a are needed. Hence the total space required by the game is bounded by $(2 + k)\log n + \log|\psi|$. This proves that the model checking problem for C^k is solvable in alternating logspace, hence in deterministic polynomial time. □

It is known that the model checking problem for CTL is in P [9]. Also the model checking problems for the bounded-variable transitive closure logics TC^k are solvable in polynomial time.

Theorem 4.4 *For all $k \geqslant 2$, the model checking problem for TC^k (and hence in particular for CL^2) is* P*-complete.*

Proof. It only remains to be shown that the problems are in P. For simplicity, we just consider TC^2. On a fixed structure $\mathfrak{A}$ we need to look only for paths of length bounded by the cardinality of $\mathfrak{A}$, so we can rewrite a formula $(\mathrm{TC}\ \varphi)(x, y)$ as $\varphi^{(n)}(x, y)$, saying that there exists a φ-path of length at most n from x to y. Due to the problem of nested TC-operators, we avoid giving a direct reduction to the model checking problem for some FO^k (which would be possible for $k = 3$ if a graph representation of the formulae is used). Instead we describe the necessary modifications of the model checking game.

At a position given by a formula $\varphi^{(m)}$ and a pair (a, b) the existential player selects an element c, with the claim that $\varphi(a, c)$ and $\varphi^{(m-1)}(c, b)$. The universal player can challenge either of these claims, and then the game proceeds either at the position $(\varphi, (a, c))$ or at the position $(\varphi^{(m-1)}, (b, c))$. Clearly the game describes the semantics of the formula in the correct way and requires only logarithmic space. □

4.3 L_μ and bounded-variable fixed point logics

The complexity of the model checking problem for the μ-calculus is probably the major open problem in this area. It has been extensively studied due to its importance for application in automatic verification. Nevertheless the problem could not yet be solved in a satisfactory way.

Theorem 4.5 *The model checking problem for L_μ is* P*-hard and is contained in* NP $\cap$ Co-NP*.*

This result has first been established explicitly in [11] and [7]. A very nice proof based on a model checking game for L_μ has been given by Stirling [35]. This model checking game was in fact discovered earlier by Herwig [19], and Theorem 4.5 is implicit in [19]. Although the model checking game for L_μ is (in some sense) a logspace game it is not clear whether the associated strategy problem is solvable in deterministic polynomial-time. The reason is that the game does not always reach a final position; instead it may get into an infinite loop. To deal with this case the definition of the winning conditions has to be extended, depending on whether the outermost fixed-point formula on the loop is a least or a greatest fixed point (for details see [19, 35]). As a result the question whether the existential player has a winning strategy is not know to be in P. However, it can be shown to be in NP. Therefore the model checking problem for L_μ is in NP, and since the μ-calculus is closed under negation it also is in Co-NP.

Remark. Unlike the case for propositional modal logic ML, it is not hard to see that also the data complexity and the expression complexity of L_μ are P-hard.

The natural question arises, whether FP^2 is more complicated to check than the μ-calculus. The next result shows that this is not the case. We can even reduce the model checking problems for more general bounded-variable fixed-point logics to the μ-calculus.

Definition 4.6 For any $k \in \mathbb{N}$, let LFP^k denote the k-variable fixed-point logic that extends FO^k by least-fixed point formulae of the form

$$[\mathrm{LFP}_{X,x_1,\dots,x_s}\ \varphi(X, x_1, \dots, x_s)](z_1, \dots, z_s)$$

where X is a relation symbol of arity $s \leqslant k$ occurring only positively in φ; φ is an LFP^k-formula having no free first-order variables other than $x_1, \dots, x_s$; and $z_1, \dots, z_s$ are arbitrary variables from $x_1, \dots, x_k$.

Note that LFP^2 extends FP^2 since in LFP^2 one can apply fixed-point operators to build binary relations.

Proposition 4.7 *For every $k \in \mathbb{N}$, the model checking problem for* LFP^k *is* LOGSPACE-*reducible to the model checking problem for L_μ.*

Proof. We present a reduction which, given a finite structure $\mathfrak{A} = (A, R_1, \dots, R_m)$ and a formula $\psi(x_1, \dots, x_k) \in \mathrm{LFP}^k$, produces a Kripke structure $\mathfrak{K}$ with universe A^k and a formula $\psi^* \in L_\mu$ such that for all $\bar{a} \in A^k$

$$\mathfrak{A} \models \psi(\bar{a}) \iff (\mathfrak{K}, \bar{a}) \models \psi^*.$$

Every relation R of $\mathfrak{A}$ is represented by a unary relation R^* of $\mathfrak{K}$ such that $R^* = \{\bar{a} \in A^k : (a_1, \dots, a_s) \in R\}$ where s is the arity of R (we can assume that $s \leqslant k$). Further $\mathfrak{K}$ has for all $i, j \in \{1, \dots, k\}$ unary relations $I_{ij} = \{\bar{a} \in A^k : a_i = a_j\}$. The binary accessibility relations (actions) of $\mathfrak{K}$ are $E_1, \dots, E_k$ with

$$(\bar{a}, \bar{b}) \in E_j \ \text{ iff } a_i = b_i \text{ for all } i \neq j$$

and for each function $\sigma : \{1 \dots k\} \to \{1, \dots, k\}$ an accessibility relation E_σ with

$$(\bar{a}, \bar{b}) \in E_\sigma \ \text{ iff } a_{\sigma(i)} = b_i \text{ for } i = 1, \dots k.$$

The translation $\psi \mapsto \psi^*$ is defined by induction:

$$\begin{aligned} (x_i = x_j)^* &:= I_{ij} \\ (Px_{\sigma(1)} \cdots x_{\sigma(s)})^* &:= \Box_\sigma P \\ (\neg\varphi)^* &:= \neg\varphi^* \\ (\varphi \vee \eta)^* &:= \varphi^* \vee \eta^* \\ (\exists x_j \varphi)^* &:= \Diamond_j \varphi^* \\ [\mathrm{LFP}_{X, x_1, \dots x_s} \varphi](x_{\sigma(1)}, \dots, x_{\sigma(s)})^* &:= \Box_\sigma(\mu_X \varphi^*) \end{aligned}$$

A straightforward induction shows that this reduction has the desired properties. □

Corollary 4.8 (Vardi) *The model-checking problem for* LFP^k *is in* NP ∩ Co-NP.

This result was first established in [37] by a different method. It should however be noted that there also exists a more powerful (and perhaps more natural) variant of fixed-point logic with k variables, permitting first-order parameters inside fixed points. This means that given a formula $\varphi(X, \bar{x}, \bar{y})$ we can build a fixed-point formula of the form

$$[\mathrm{LFP}_{X, \bar{x}} \ \varphi(X, \bar{x}, \bar{y})](\bar{z}, \bar{y}).$$

This more powerful variant can apparently not be reduced to the μ-calculus. In fact, Dziembowski [10] showed, that even in the two-variable case the free parameters inside nested fixed points can be used to simulate arbitrary sequences of Boolean variables. Thus the model checking of quantified Boolean formulae is reducible to this liberalized variant of FP^2.

Theorem 4.9 (Dziembowski) *For every $k \geqslant 2$, the model checking problem for k-variable fixed-point logic with parameters is* PSPACE*-complete. Indeed, there exists a very simple structure $\mathfrak{B}$ (just a set with three elements) such that the expression complexity of the liberalized* FP^2 *on $\mathfrak{B}$ is* PSPACE*-complete.*

References

1. H. ANDRÉKA, J. VAN BENTHEM, AND I. NÉMETI, *Modal Languages and Bounded Fragments of Predicate Logic*, ILLC Research Report ML-96-03 (1996), 59 pages.

2. J. BALCÁZAR, J. DÍAZ, AND J. GABARRÓ, *Structural Complexity II*, Springer 1990.

3. J. VAN BENTHEM, *Modal Logic an Classical Logic*, Bibliopolis, Napoli 1983.

4. J. VAN BENTHEM, *Exploring Logical Dynamics*, CSLI-Publications, Stanford 1996.

5. J. VAN BENTHEM, *Dynamic bits and pieces*, ILLC Research Report (1997).

6. R. BERGER, *The undecidability of the domino problem*, Mem. AMS, 66 (1966).

7. O. BERNHOLTZ, M. Y. VARDI, AND P. WOLPER, An automata-theoretic approach to branching-time model checking. *Poc. of 6th Inter. Conf. on Computer Aided Verification, CAV '94*, Lecture Notes in Computer Science Nr. 818 (1994), 142–155.

8. E. BÖRGER, E. GRÄDEL, AND Y. GUREVICH, *The Classical Decision Problem*, Springer, 1997.

9. E. M. CLARKE, E. A. EMERSON, AND A. P. SISTLA. *Automatic verification of finite-state concurrent systems using temporal logic specifications*, ACM Transactions on Programming Languages and Systems, 8 (1986), 244–263.

10. S. DZIEMBOWSKI, *Bounded-Varioable Fixpoints Queries are PSPACE-complete*, Computer Science Logic CSL '96, Selected Papers, Lecture Notes in computer Science Nr. 1258, Springer 1997, 89–105.

11. E. EMERSON, C. JUTLA AND A. SISTLA, On model-checking for fragments of μ-calculus, Proc. of 5. Int. Workshop on Computer-Aided Verification CAV 93, Lecture Notes in Computer Science Nr. 697, Springer 1993, 385–396.

12. 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, Lecture Notes in Computer Science No. 171, Springer, 1984, pp. 312–319.

13. E. GRÄDEL, *On the restraining power of guards*, submitted for publication (1997).

14. E. GRÄDEL, P. KOLAITIS, AND M. VARDI, *On the complexity of the decision problem for two-variable first-order logic.* Bulletin of Symbolic Logic, 3, 1997, pp. 53–69.

15. E. GRÄDEL, M. OTTO, AND E. ROSEN, *Undecidability results on two-variable logics*, to appear in Archive of Mathematical Logic. A preliminary version has appeared in Proceedings of 14th Symposium on Theoretical Aspects of Computer Science STACS'97, Lecture Notes in Computer Science Nr. 1200, Springer (1997), 249–260.

16. ——, *Two-variable logic with counting is decidable.* Proc. 12th IEEE Symp. on Logic in Computer Science LICS '97, Warsaw 1997, pp. 306–317.

17. R. GRAHAM, B. ROTHSCHILD, AND J. SPENCER, *Ramsey Theory.* John Wiley and Sons, 1980.

18. Y. Gurevich and I. Koryakov, *Remarks on Berger's paper on the domino problem*, Siberian Math. Journal, 13 (1972), pp. 319–321.

19. B. Herwig, *Zur Modelltheorie von L_μ*, Dissertation, Universität Freiburg (1989).

20. W. Hodges, *Model Theory*, Cambridge University Press 1993.

21. N. Immerman and E. Lander, *Describing graphs: a first-order approach to graph canonization*, in A. Selman, ed., Complexity Theory Retrospective, Springer, 1990, pp. 59–81.

22. N. Immerman and M. Vardi, *Model checking and transitive closure logic*, Proc. 9th Conf. Computer-Aided Verification, (1997), pp. 291–302.

23. P. Kolaitis and M. Otto, *The boundedness problem for two-variable first-order logic*, to appear in Proc. 13th IEEE Symp. on Logic in Computer Science LICS '98, Indianapolis, 1998.

24. D. Kozen, *A finite model theorem for the propositional μ-calculus*, Studia Logica, 47 (1983), pp. 233–241.

25. H. Lewis, *Complexity Results for Classes of Quantificational Formulas*, Journal of Computer and System Sciences21 (1980), pp. 317–353.

26. M. Mortimer, *On languages with two variables*, Zeitschr. f. math. Logik u. Grundlagen d. Math., 21 (1975), pp. 135–140.

27. R. Milner, *A Calculus of Communicating Systems*, Lecture Notes in Computer Science, vol. 92, Springer, 1980.

28. M. Otto, *Capturing bisimulation-invariant Ptime*, in Proc. 4th Symposium on Logical Foundations of Computer Science, 1997, LNCS 1234, 1997, pp. 294–305.

29. ——, *Bounded-variable logics: two, three, and more*, to appear in Archive for Mathematical Logic (special issue), 1998.

30. ——, *Two-variable first-order over equivalence relations*, unpublished manuskript,1997.

31. ——, *Two-variable first-order logic over ordered domains*, unpublished, 1998.

32. L. Pacholski, W. Szwast, and L. Tendera, *Complexity of Two-Variable Logic with Counting*, Proc. 12th IEEE Symp. on Logic in Computer Science LICS '97, Warsaw 1997, pp. 318–327.

33. C. Papadimitriou, *Computational Complexity*, Addison-Wesley (1994).

34. D. Scott, *A decision method for validity of sentences in two variables*, Journal of Symbolic Logic, 27 (1962), p. 377.

35. C. Stirling, *Local Model Checking Games (Extended Abstract)*, in: CONCUR 95. Concurrency Theory, Lecture Notes in Computer Science Nr. 962, Springer (1995), 1–11.

36. A. Tarski, A. Mostowski, and R. Robinson, *Undecidable Theories*, North-Holland, Amsterdam 1953.

37. M. Vardi, *On the complexity of bounded-variable queries*, Procceedings of 14th Ann. ACM Symposium on Principles of Database Systems PODS 1995, pp 266–276.

38. ——, *Why is modal logic so robustly decidable*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science 31, American Math. Society, 1997, 149–184.