Bisimulation-Invariant Ptime and Higher-Dimensional $\boldsymbol{\mu}$-Calculus

Martin Otto

Mathematische Grundlagen der Informatik RWTH Aachen

currently visiting UC Santa Cruz and Stanford University

otto@informatik.rwth-aachen.de

Revised, May 98

Abstract. *Consider the class of all those properties of worlds in finite Kripke structures (or of states in fiinite transition systems), that are*

- *recognizable in polynomial time, and*
- *closed under bisimulation equivalence.*

It is shown that the class of these bisimulation-invariant PTIME queries *has a natural logical characterization. It is captured by the straightforward extension of propositional μ-calculus to arbitrary finite dimension. Bisimulation-invariant* PTIME*, or the* modal fragment of PTIME*, thus proves to be one of the very rare cases in which a logical characterization is known in a setting of unordered structures.*

It is also shown that higher-dimensional μ-calculus is undecidable for satisfiability in finite structures, and even Σ^1_1-hard over general structures.

keywords: logic in computer science, finite model theory, modal logic, descriptive complexity, process logics, model checking

0 Introduction

An outstanding issue in the study of the relation between computational complexity and logical definability concerns the search for exact matches. Paradigmatic results in this area are, for instance, Fagin's Theorem (the NP-recognizable properties of finite structures are exactly those that can be formalized in existential second-order logic), the Büchi-Elgot-Trakhtenbrot Theorem (the automaton-recognizable properties of finite words are those that are definable in monadic second-order logic), or the Immerman-Vardi Theorem (the PTIME properties of finite linearly ordered structures are exactly those that are definable in least fixed-point logic).

It is a characteristic feature in these examples that they either concern complexity classes beyond PTIME or else concern classes of linearly ordered structures. Indeed, no logical characterization has been found for any of the standard complexity classes below NP, that would cover arbitrary rather than linearly ordered structures. In particular, the question whether PTIME itself – regarded as the class of all those properties of finite structures that can be recognized by PTIME algorithms – admits a logical characterization, is a central open problem in finite model theory. This fundamental issue was raised by Chandra and Harel [10] and more rigorously formalized by Gurevich [16], cf. also [12].

The present investigation deals with, and offers a positive solution for, a semantically defined fragment of PTIME concerning finite Kripke structures. Kripke structures not only form the natural models for *modal logics* but also play an important role as formalizations of *transition systems*. Under both aspects, *bisimulation equivalence* is the adequate notion of indistinguishability. It is therefore natural in this framework to consider the class of those PTIME properties of finite Kripke structures (transition systems), that are preserved under bisimulation. It turns out that this class possesses an exact logical match in higher-dimensional μ-calculus, which is here introduced as the obvious extension of ordinary propositional μ-calculus L_μ to arbitrary arities. Apart from its theoretical appeal, this result is of potential interest for model-checking applications. Just as a natural logic for PTIME would, if it exists, be the theoretically ideal database language in the world of relational databases, the higher-dimensional μ-calculus is, in a precise sense, an optimal logic for all efficient model checking tasks in a bisimulation-invariant framework. Moreover, as a natural extension of the standard propositional μ-calculus L_μ, this language makes close connections with theoretically well developed areas in model checking, and also highlights the fundamental role of the μ-calculus in a new way.

A similar claim can of course be made with a view to the logically more fundamental framework of modal logic. As bisimulation-invariance may reasonably be regarded as the defining characteristic of modal properties, the higher-dimensional μ-calculus provides a modal fixed-point logic which precisely captures PTIME in the modal world – and does so without having to resort to a given ordering. The fact that the question about a logic for PTIME is thus answered affirmatively for the modal world, may be seen as yet another indication of the more general phenomenon that the model theory of modal logic shows a much neater behaviour in restriction to finite structures than does classical first-order logic.

In a further study of its expressive power over finite and infinite structures, the higher-dimensional μ-calculus is shown to be undecidable for satisfiability in finite models as well as in general models. In sharp contrast with L_μ itself, even two-dimensional L_μ does no longer have the finite model property, and its satisfiability problem is hard for the first level of the analytical hierarchy (Σ^1_1-hard).

1 Preliminaries, basic definitions, and the main theorem

We deal with *Kripke structures* that form the appropriate models for propositional modal logic ML, its infinitary variant ML_∞, and the propositional μ-calculus L_μ. Of course, Kripke structures may be identified with transition systems, with only some minor changes in terminology (and beyond these introductory remarks, we choose to stick with the terminology of Kripke structures). We fix a finite set of basic propositions or propositional constants $\overline{P} = P_1, \dots, P_l$. A Kripke structure (transition system) for $\overline{P}$

then is a structure

$$\mathfrak{A} = \big(A, E^{\mathfrak{A}}, P_1^{\mathfrak{A}}, \dots, P_l^{\mathfrak{A}}\big),$$

where

- A is the universe or *set of worlds* (states) of $\mathfrak{A}$,
- $E^{\mathfrak{A}} \subseteq A^2$ is the binary relation of *accessibility between worlds* (atomic transitions between states),
- each $P_i^{\mathfrak{A}} \subseteq A$ interprets the set of worlds (states) in which P_i *holds true*.

For the standard semantics of ML, ML_∞ and L_μ one usually deals with Kripke structures $(\mathfrak{A}, a)$ in which one element is designated, one also speaks of *model-world pairs*:

- $a \in A$ the *distinguished world* of $(\mathfrak{A}, a)$.

In more first-order minded terms, a Kripke structure for $\overline{P}$ is just a τ-structure of vocabulary $\tau = \{E, P_1, \dots, P_l\}$ with binary E and unary P_i. A distinguished world a may be regarded as a fixed parameter or as a constant in $\mathfrak{A}$.

It is customary to consider Kripke structures, and in particular transition systems, which have more than one accessibility relation (atomic transition). All results presented here extend to that more general, multi-modal framework with only minor modifications, which are summarily indicated in Section 2.4. It therefore seems justified to simplify the formal presentation of the main development through restriction to the basic case of just one binary relation.

1.1 Bisimulation equivalence

A fundamental notion of equivalence between Kripke structures with distinguished worlds is *bisimulation equivalence*, which we denote by $\sim$. This equivalence has a natural motivation as a notion of behavioural indistinguishability, if Kripke structures are taken as descriptions of transition systems [26, 31]. There is also a very elegant (in fact also earlier) Ehrenfeucht-Fraïssé style characterization of bisimulation equivalence , due to van Benthem [6, 7], by means of a game in which the two players can move a single pebble in each structure along forward E-edges.

Definition 1.1 Let $\mathfrak{A} = (A, E, P_1, \dots, P_l)$ and $\mathfrak{A}' = (A', E', P_1', \dots, P_l')$ be Kripke structures. A relation $R \subseteq A \times A'$ is a *bisimulation* between $\mathfrak{A}$ and $\mathfrak{A}'$ if the following conditions hold for all $(a, a') \in R$:

(i) $a \in P_i \Leftrightarrow a' \in P_i'$, for $1 \leqslant i \leqslant l$.

(ii) for all b such that $(a, b) \in E$, there is some b' such that $(a', b') \in E'$ and $(a', b') \in R$.

(iii) for all b' such that $(a', b') \in E'$, there is some b such that $(a, b) \in E$ and $(a', b') \in R$.

There is a natural *largest bisimulation* $\sim$ between $\mathfrak{A}$ and $\mathfrak{A}'$, which is obtained as the union over all bisimulations between $\mathfrak{A}$ and $\mathfrak{A}'$.

$(\mathfrak{A}, a)$ and $(\mathfrak{A}', a')$ are *bisimulation equivalent*, or just bisimilar, $(\mathfrak{A}, a) \sim (\mathfrak{A}', a')$, if $(a, a') \in R$ for some (and hence for the largest) bisimulation between $\mathfrak{A}$ and $\mathfrak{A}'$.

We shall mostly work with the following equivalent, inductive characterization of *bisimulation inequivalence* $\not\sim$ as a least fixed point – which clearly corresponds to the natural co-inductive definition of the largest bisimulation $\sim$ itself as a greatest fixed point. Let $\mathfrak{A} = (A, E, P_1, \dots, P_l)$ and $\mathfrak{A}' = (A', E', P_1', \dots, P_l')$ be two Kripke structures. Put

- $(\mathfrak{A}, a) \not\sim_0 (\mathfrak{A}', a')$ if $\bigvee_{i=1}^{l} \neg\big(a \in P_i \leftrightarrow a' \in P_i'\big)$
- $(\mathfrak{A}, a) \not\sim_{\alpha+1} (\mathfrak{A}', a')$ if

$$\begin{array}{ll} & (\mathfrak{A}, a) \not\sim_\alpha (\mathfrak{A}', a') \\ \text{or} & \exists b \in A \, \Big[(a,b) \in E \wedge \forall b' \in A' \Big((a',b') \in E' \to (\mathfrak{A}, b) \not\sim_\alpha (\mathfrak{A}', b')\Big)\Big] \\ \text{or} & \exists b' \in A' \, \Big[(a',b') \in E' \wedge \forall b \in A \Big((a,b) \in E \to (\mathfrak{A}, b) \not\sim_\alpha (\mathfrak{A}', b')\Big)\Big] \end{array} \tag{1}$$

- $(\mathfrak{A}, a) \not\sim_\lambda (\mathfrak{A}', a')$ if $\exists \alpha < \lambda \ (\mathfrak{A}, a) \not\sim_\alpha (\mathfrak{A}', a')$ for limits λ.

- $(\mathfrak{A}, a) \not\sim (\mathfrak{A}', a')$ if $(\mathfrak{A}, a) \not\sim_\alpha (\mathfrak{A}', a')$ for some α.

For cardinality reasons, the sequence of the $\not\sim_\alpha$ becomes stationary at some stage in restriction to any given pair of structures $\mathfrak{A}$ and $\mathfrak{A}'$. Thus $(\mathfrak{A}, a) \not\sim (\mathfrak{A}', a')$ if $(\mathfrak{A}, a) \not\sim_\alpha (\mathfrak{A}', a')$ for the least α such that

$$\forall a \in A \, \forall a' \in A' \, \Big((\mathfrak{A}, a) \not\sim_{\alpha+1} (\mathfrak{A}', a') \Longrightarrow (\mathfrak{A}, a) \not\sim_\alpha (\mathfrak{A}', a') \Big).$$

This least α is the closure ordinal of the inductive definition. Over finite structures $\mathfrak{A}$ and $\mathfrak{A}'$ the limit is reached within polynomially many steps: indeed, the closure ordinal is generally bounded by $(|A| \cdot |A'|)^+$, which is $|A| \cdot |A'| + 1$ if these cardinalities are finite. This implies in particular that bisimulation equivalence over finite Kripke structures is PTIME computable. In fact it is PTIME-complete [2].

There is an Ehrenfeucht-Fraïssé type theorem associated with bisimulation equivalence which involves the infinitary variant ML_∞ of ordinary propositional modal logic ML, compare e.g. [5]. Recall that ML for $\overline{P} = P_1, \dots, P_l$ has atomic formulae $P_i x$, is closed under boolean operations, and under the modal constructors $\Diamond$ and $\Box$, where the semantics of the latter is given by

$$\begin{array}{lll} (\mathfrak{A}, a) \models \Diamond\varphi & \text{iff} & \exists b \in A \big((a,b) \in E^{\mathfrak{A}} \wedge (\mathfrak{A}, b) \models \varphi\big) \\ (\mathfrak{A}, a) \models \Box\varphi & \text{iff} & \forall b \in A \big((a,b) \in E^{\mathfrak{A}} \to (\mathfrak{A}, b) \models \varphi\big). \end{array}$$

ML_∞ further enriches the syntax and semantics of ML through closure under conjunctions and disjunctions over arbitrary sets of formulae.

Clearly $\Diamond$ and $\Box$ may be pictured as existential and universal first-order quantifications along accessibility edges. It is therefore straightforward that $\mathrm{ML} \subsetneq L^2_{\omega\omega}$ and $\mathrm{ML}_\infty \subsetneq L^2_{\infty\omega}$, where $L^2_{\omega\omega}$ is first-order logic with only two variable symbols, $L^2_{\infty\omega}$ its infinitary variant. This translation of modal logics into first-order or infinitary logic (with just two variables) is explored in model-theoretic terms in the work of van Benthem. Note that $\Diamond$ is the dual of $\Box$ so that only one of these operators need be retained in the presence of negation.

We shall write all modal formulae φ as formulae $\varphi(x)$ in a single formal element variable x, which is ultimately interpreted by the distinguished world in a Kripke structure. The semantics of φ is here associated with the monadic predicate defined by $\varphi(x)$ over the $\mathfrak{A}$:

$$\varphi[\mathfrak{A}] = \{a \in A \mid \mathfrak{A} \models \varphi[a]\} = \{a \in A \mid (\mathfrak{A}, a) \models \varphi\},$$

whereby formulae of ML or ML_∞ define *monadic global relations* over Kripke structures.

Theorem 1.2 (Barwise, van Benthem, Moss [3, 5]) $(\mathfrak{A}, a) \sim (\mathfrak{A}', a')$ *if and only if* $(\mathfrak{A}, a)$ *and* $(\mathfrak{A}', a')$ *satisfy exactly the same formulae of* ML_∞. *For finite (in fact even for finitely branching) Kripke structures:* $(\mathfrak{A}, a) \sim (\mathfrak{A}', a')$ *if and only if* $(\mathfrak{A}, a)$ *and* $(\mathfrak{A}', a')$ *satisfy exactly the same formulae of* ML.

1.2 Propositional μ-calculus

Propositional μ-calculus L_μ, as introduced in [23], augments the syntax and semantics of ML by constructors for least and greatest fixed points. To this end one firstly admits propositional variables $X, Y, Z, \ldots$ in formulae, with corresponding new atomic assertions Xx etc. Free propositional variables get interpreted by subsets of the universe (sets of worlds) like the propositional constants.

If φ is a formula of L_μ that is positive in X (meaning that X does not occur free in the scope of an odd number of negations), then $\psi = \mu_X \varphi$ and $\psi' = \nu_X \varphi$ are also formulae of L_μ (in which X no longer occurs free).

The semantics can without loss of generality be explained in the case that no propositional variable apart from X is free in φ. Then φ induces a monotone operator on subsets of A according to

$$F^{\mathfrak{A}}_\varphi \colon P \subseteq A \longmapsto \{a \in A \mid (\mathfrak{A}, a) \models \varphi[P/X]\},$$

where $\varphi[P/X]$ denotes φ under the interpretation that assigns P to X. Being a monotone operator, $F^{\mathfrak{A}}_\varphi$ possesses least and greatest fixed points $\mathrm{LFP}(F^{\mathfrak{A}}_\varphi)$ and $\mathrm{GFP}(F^{\mathfrak{A}}_\varphi)$. Now

$$(\mathfrak{A}, a) \models \mu_X \varphi \quad \text{iff} \quad a \in \mathrm{LFP}(F^{\mathfrak{A}}_\varphi) \quad \text{and} \quad (\mathfrak{A}, a) \models \nu_X \varphi \quad \text{iff} \quad a \in \mathrm{GFP}(F^{\mathfrak{A}}_\varphi).$$

As usual, least and greatest fixed points may be generated inductively. For instance, putting inductively

$$\begin{array}{lcl} X^{\mathfrak{A}}_0 & = & \emptyset \\ X^{\mathfrak{A}}_{\alpha+1} & = & F^{\mathfrak{A}}_\varphi(X^{\mathfrak{A}}_\alpha) \\ X^{\mathfrak{A}}_\lambda & = & \bigcup_{\alpha<\lambda} X^{\mathfrak{A}}_\alpha \text{ for limit } \lambda \end{array} \qquad \text{and} \qquad \begin{array}{lcl} X'^{\mathfrak{A}}_0 & = & A \\ X'^{\mathfrak{A}}_{\alpha+1} & = & F^{\mathfrak{A}}_\varphi(X'^{\mathfrak{A}}_\alpha) \\ X'^{\mathfrak{A}}_\lambda & = & \bigcap_{\alpha<\lambda} X'^{\mathfrak{A}}_\alpha \text{ for limit } \lambda, \end{array}$$

it is easy to see that the X_α are increasing, the X'_α decreasing, and that

$$\mathrm{LFP}(F^{\mathfrak{A}}_\varphi) = \bigcup_\alpha X^{\mathfrak{A}}_\alpha \quad \text{and} \quad \mathrm{GFP}(F^{\mathfrak{A}}_\varphi) = \bigcap_\alpha X'^{\mathfrak{A}}_\alpha.$$

As with $\Diamond$ and $\Box$, the μ- and ν-operators are related by a straightforward duality so that it suffices to retain one of them in the presence of negation.

It is well known that L_μ is preserved under bisimulation: if $(\mathfrak{A}, a) \sim (\mathfrak{A}', a')$ and $\varphi \in \mathrm{L}_\mu$, then $(\mathfrak{A}, a) \models \varphi$ iff $(\mathfrak{A}', a') \models \varphi$. Indeed, this is a consequence of the following statement, which is proved by syntactic induction (and using the inductive generation of least and greatest fixed-points in the μ- and ν-steps).

Fact 1.3 *Let κ be any infinite cardinal. Then there is for each $\varphi \in \mathrm{L}_\mu$ a formula $\varphi^{(\kappa)} \in \mathrm{ML}_\infty$ such that φ and $\varphi^{(\kappa)}$ are equivalent over all Kripke structures of cardinality less than κ.*

It follows that $\mathrm{L}_\mu \subseteq \mathrm{ML}_\infty \subseteq L^2_{\infty\omega}$ in restriction to any class of Kripke structures with a uniform bound on the cardinality. Even over the class of finite Kripke structures the inclusion $\mathrm{L}_\mu \subseteq \mathrm{ML}_\infty$ is strict as $\mathrm{L}_\mu \subseteq \textsc{Ptime}$. More background on L_μ, its variants, and its role as a process logic can be found in [13]. It is customary to introduce L_μ in a multi-modal framework, i.e. with several accessibility relations and corresponding modalities rather than just one. As pointed out above, all results presented here have straightforward extensions to that scenario, see also Section 2.4.

1.3 A k-dimensional μ-calculus

We introduce extensions of L_μ which roughly correspond to the expressive power of L_μ over the k-th Cartesian power of the given Kripke structures. The elements of this power are k-tuples of worlds, $\overline{a} \in A^k$, and there are k different accessibility relations E_j corresponding to E-accessibility in the j-th component, for $1 \leqslant j \leqslant k$:

$$(\overline{a}, \overline{a}') \in E_j^{\mathfrak{A}} \quad \text{iff} \quad a_i = a_i' \text{ for } i \neq j, \text{ and } (a_j, a_j') \in E^{\mathfrak{A}}.$$

We write formulae of L^k_μ in a k-tuple of element variables $\overline{x} = (x_1, \dots, x_k)$ and semantically associate them with k-ary global relations over Kripke structures.

The syntax is governed by the following clauses:

atomic formulae: for $1 \leqslant i \leqslant l$ and $1 \leqslant j \leqslant k$, $P_i x_j$ is an atomic formula of L^k_μ. For k-ary second-order variables $X, Y, Z, \dots$ L^k_μ has atomic formulae $X\overline{x}$, $Y\overline{x}$, $Z\overline{x}$, …

booleans: L^k_μ is closed under the boolean operations $\neg$, $\wedge$, $\vee$.

modalities: L^k_μ is closed under modalities $\Diamond_j$ and $\Box_j$ for $1 \leqslant j \leqslant k$.

variable substitutions: L^k_μ is closed under substitutions $\sigma\colon \{1, \dots, k\} \to \{1, \dots, k\}$, which operate on variables according to $\sigma\colon (x_1, \dots, x_k) \mapsto (x_{\sigma(1)}, \dots, x_{\sigma(k)})$.

least and greatest fixed points: L^k_μ is closed under applications of μ_X- and ν_X-operators to φ whenever X occurs positively. Note that X plays the role of a k-ary relation variable, which we stress by speaking of μ^k- and ν^k-operations.

Semantically we associate with each formula $\varphi \in \mathrm{L}^k_\mu$ a k-ary predicate

$$\varphi[\mathfrak{A}] = \{\overline{a} = (a_1, \dots, a_k) \in A^k \mid \mathfrak{A} \models \varphi[\overline{a}]\}$$

over each $\mathfrak{A}$ (with an interpretation for all free second-order variables in φ):

- The atomic and boolean cases are obvious.
- $\Diamond_j$ and $\Box_j$ are modal operators for the accessibility relation $E_j^{\mathfrak{A}}$ over A^k.
- Substitutions σ operate as $\varphi^\sigma[\mathfrak{A}] = \{(a_1, \dots, a_k) \in A^k \mid \mathfrak{A} \models \varphi[a_{\sigma(1)}, \dots, a_{\sigma(k)}]\}$.
- Operators μ_X and ν_X correspond to least and greatest fixed points of the induced monotone operation that sends $P \subseteq A^k$ to $\{\overline{a} \in A^k \mid \mathfrak{A} \models \varphi[P/X, \overline{a}]\}$.

For ease of notation we shall also just write $\varphi(x_{\sigma(1)}, \dots, x_{\sigma(k)})$ for φ^σ, or indicate simple substitutions as in $\varphi(x_k/x_i)$ for φ^σ if σ maps i to k and fixes all other indices.

For L_μ^k, too, it would clearly suffice to retain only one of the μ^k- or ν^k-operators and only one of $\Diamond_j$ and $\Box_j$ for each j, since the semantics is in accordance with the usual duality relations between these.

Towards a *standard monadic semantics* of formulae of L_μ^k over Kripke structures $(\mathfrak{A}, a)$ with the single distinguished world a, we ultimately pass to substitution instances $\varphi(x_1, x_1, \dots, x_1)$ (i.e. σ identically 1) of formulae without free second-order variables and write these as $\varphi(x)$ to stress their role as defining monadic global relations.

Remark For a substitution instance which reduces the number of variables that occur free in a formula φ, we may either think of the result as still defining a k-ary global relation (with trivial factors in those components that are not actually free any more), or as defining a global relation of correspondingly reduced arity. For instance a formula $\varphi(x_1, x_1) \in \mathrm{L}_\mu^2$ could alternatively be associated with the monadic global relation given by $Q_1^{\mathfrak{A}} = \{a \in A \mid \mathfrak{A} \models \varphi[a, a]\}$, or with the binary global relation according to $Q_2^{\mathfrak{A}} = \{(a_1, a_2) \in A^2 \mid \mathfrak{A} \models \varphi[a_1, a_1]\} = Q_1^{\mathfrak{A}} \times A$. The translation between these interpretations is straightforward and we shall often increase the arity of relations to some uniform common value by this type of *padding* that occurs in the translation from Q_1 to Q_2. [1]

Definition 1.4 L_μ^k is k-dimensional μ-calculus as introduced above, L_μ^ω denotes the union $\bigcup_k \mathrm{L}_\mu^k$. ML^k is k-dimensional modal logic, the fragment of L_μ^k without μ^k- or ν^k-operators. ML_∞^k stands for the corresponding infinitary k-dimensional modal logic, which is closed under disjunctions and conjunctions over arbitrary sets of formulae.

With each of these logics we associate a *general semantics*, in which for instance L_μ^k defines global relations of arity (up to) k, and a *standard monadic semantics* as outlined for L_μ^k above. The latter is the main concern in the present investigation, and will always be tacitly implied in considerations about monadic global relations.

Clearly ML^k, L_μ^k, ML_∞^k for $k = 1$ are the familiar ML, L_μ, and ML_∞. For each of these logics the $k+r$-dimensional variant is at least as expressive as the k-dimensional one, in standard monadic semantics as well as for the general semantics (up to the necessary padding). We shall see that the extension to higher dimension adds crucial expressive power to L_μ even with respect to the standard monadic semantics. With the

[1] A semantics we shall *not* associate with $\varphi(x_1, x_1)$ is that of $\{(a, a) \in A^2 \mid \mathfrak{A} \models \varphi[a, a]\}$. This would introduce a hidden equality which is completely alien to the bisimulation-oriented framework.

ML^k_∞ versus ML_∞, and ML^k versus ML, on the other hand, there is actually no gain in expressiveness for the standard monadic semantics, and only a trivial one even with respect to the general semantics (see Remark 1.6 below).

For the monadic semantics over any class of bounded cardinality, the inclusion structure is in fact the following, for $k \geqslant 2$ (proofs are indicated below):

$$\begin{array}{ccccc}
\mathrm{ML}^k & \subsetneq & \mathrm{L}^k_\mu & \subsetneq & \mathrm{ML}^k_\infty \\[1em]
\| & & \rotatebox[origin=c]{90}{$\subsetneq$} & & \| \\[1em]
\mathrm{ML} & \subsetneq & \mathrm{L}_\mu & \subsetneq & \mathrm{ML}_\infty
\end{array}$$

The following L^k_μ-analogue of Fact 1.3 has a direct inductive proof, based on the inductive generation of fixed points.

Fact 1.5 *For any infinite cardinal κ and any $\varphi \in \mathrm{L}^k_\mu$ without free second-order variables there is a formula $\varphi^{(\kappa)} \in \mathrm{ML}^k_\infty$ that is equivalent with φ over all Kripke structures of cardinality less than κ.*

Let $L^k_{\infty\omega}$ stand for infinitary logic with k variables. An argument that is strictly analogous to that for $\mathrm{ML}_\infty \subseteq L^2_{\infty\omega}$ shows that $\mathrm{ML}^k_\infty \subseteq L^k_{\infty\omega}$ for $k \geqslant 2$. Therefore also $\mathrm{L}^k_\mu \subseteq L^k_{\infty\omega}$ in restriction to any class of structures of bounded cardinality, and for the general semantics. But in fact much more can be said. The following is a theorem in [7] for ML, which, together with its inductive proof immediately generalizes to ML_∞.

Remark 1.6 *Any formula $\varphi(x_1, \dots, x_k) \in \mathrm{ML}^k_\infty$ is equivalent with an infinitary boolean combination of formulae $\chi(x_j/x)$, where $\chi \in \mathrm{ML}_\infty$ and $1 \leqslant j \leqslant k$.*

This implies in particular that $\mathrm{ML}^k_\infty \equiv \mathrm{ML}_\infty$ just as $\mathrm{ML}^k \equiv \mathrm{ML}$ with respect to the standard monadic semantics.

Corollary 1.7 *Over any class of structures of bounded cardinality, and in particular over the class of all finite Kripke structures: $\mathrm{L}^k_\mu \subseteq \mathrm{ML}_\infty \subseteq L^2_{\infty\omega}$ for the standard monadic semantics.*

It is also easy to show inductively that the ML^k_∞ respects bisimulation equivalence, in the sense that for $\varphi \in \mathrm{ML}^k_\infty$, and Kripke structures $\mathfrak{A}$ and $\mathfrak{A}'$ and $\overline{a} \in A^k$, $\overline{a}' \in A'^k$ are such that $(\mathfrak{A}, a_j) \sim (\mathfrak{A}', a'_j)$ for $j = 1, \dots, k$: $\overline{a} \in \varphi[\mathfrak{A}]$ if and only if $\overline{a}' \in \varphi[\mathfrak{A}']$. Of course this also follows from Remark 1.6 and the bisimulation-invariance of ML_∞. As a corollary we have that L^k_μ respects bisimulation equivalence. As this fact is so important in the present investigation, we also sketch a direct proof.

Lemma 1.8 *Let $\varphi(\overline{x}) \in \mathrm{L}^k_\mu$, $\mathfrak{A}$ and $\mathfrak{A}'$ Kripke structures. If $\overline{a} \in A^k$ and $\overline{a}' \in A'^k$ are such that for $j = 1, \dots, k$ $(\mathfrak{A}, a_j) \sim (\mathfrak{A}', a'_j)$, then $\overline{a} \in \varphi[\mathfrak{A}]$ if and only if $\overline{a}' \in \varphi[\mathfrak{A}']$. In particular any monadic global relation that is definable in L^k_μ is bisimulation-invariant.*

Sketch of proof. Let us write $\overline{a} \sim \overline{a}'$ for tuples $\overline{a} \in A^k$ and $\overline{a}' \in A'^k$ if $(\mathfrak{A}, a_j) \sim (\mathfrak{A}', a'_j)$ for $j = 1, \dots, k$. An inductive proof of the lemma has to allow for free second-order variables in (subformulae of) φ. The version of the statement that lends itself to this treatment is the following. Let $\varphi(\overline{X}, \overline{x})$ have free variables as displayed. Let $(\mathfrak{A}, \overline{P})$ and $(\mathfrak{A}', \overline{P}')$ with interpretations $\overline{P}$ and $\overline{P}'$ for $\overline{X}$ be such that for all $\overline{a} \in A^k$ and $\overline{a}' \in A'^k$, $\overline{a} \sim \overline{a}'$ implies $\overline{a} \in P \Leftrightarrow \overline{a}' \in P'$. Then $\overline{a} \sim \overline{a}'$ also implies $\overline{a} \in \varphi[\mathfrak{A}, \overline{P}] \Leftrightarrow \overline{a}' \in \varphi[\mathfrak{A}', \overline{P}']$.

The claim is clear for atomic φ, obviously carries over to boolean combinations and to substitution instances. For $\varphi = \Diamond_j \psi$, an appeal to $(\mathfrak{A}, a_j) \sim (\mathfrak{A}', a'_j)$ and to the inductive hypothesis for ψ yields the desired result. For $\varphi = \mu_X \psi(X, \overline{X}, \overline{x})$, it is first shown inductively that the individual stages $X_\alpha^{\mathfrak{A}}$ and $X_\alpha^{\mathfrak{A}'}$ in the inductive generation of the least fixed points in $\mathfrak{A}$ and $\mathfrak{A}'$, respectively, conform to the requirement that $\overline{a} \sim \overline{a}'$ implies $\overline{a} \in X_\alpha^{\mathfrak{A}} \Leftrightarrow \overline{a}' \in X_\alpha^{\mathfrak{A}'}$. This then immediately implies the desired result for $\varphi[\mathfrak{A}, \overline{P}] = \bigcup X_\alpha^{\mathfrak{A}}$ and $\varphi[\mathfrak{A}', \overline{P}'] = \bigcup X_\alpha^{\mathfrak{A}'}$ as well. □

A crucial example of the expressive power of the L_μ^k is the following:

Lemma 1.9 *Consider bisimulation equivalence $\sim$ as a binary global relation on individual Kripke structures, according to $(a, a') \in \sim^{\mathfrak{A}}$ if $(\mathfrak{A}, a) \sim (\mathfrak{A}, a')$. Then $\sim$ is definable in L_μ^k for all $k \geqslant 2$.*

In fact, a comparison with the inductive generation of the $\not\sim_\alpha$ in (1) in section 1.1 shows that $\not\sim$ is defined as a global relation in variables x_1 and x_2 in this sense by the formula

$$\varphi_{\not\sim} = \mu_X \Big(\bigvee_{i=1}^{l} \neg(P_i x_1 \leftrightarrow P_i x_2) \vee \Diamond_1 \Box_2 X \overline{x} \vee \Diamond_2 \Box_1 X \overline{x} \Big).$$

1.4 The capturing result

We now consider monadic queries or global relations on finite Kripke structures. Formally such a query Q is a mapping that associates with each finite Kripke structure $\mathfrak{A}$ (of the fixed type $\overline{P}$) a subset

$$Q^{\mathfrak{A}} \subseteq A$$

such that any isomorphism $f : \mathfrak{A} \simeq \mathfrak{A}'$ is also an isomorphism of the expansions $(\mathfrak{A}, Q^{\mathfrak{A}})$ and $(\mathfrak{A}', Q^{\mathfrak{A}'})$. We are here interested in the stronger invariance condition imposed by bisimulation equivalence.

Definition 1.10 A monadic query Q on finite Kripke structures is

(i) *bisimulation-invariant* ($\sim$-invariant) if $(\mathfrak{A}, a) \sim (\mathfrak{A}', a')$ implies that $a \in Q^{\mathfrak{A}}$ iff $a' \in Q^{\mathfrak{A}'}$.

(ii) *polynomial time computable* (in PTIME) if there is a PTIME algorithm for deciding on input $(\mathfrak{A}, a)$ whether $a \in Q^{\mathfrak{A}}$.

If a monadic query Q on Kripke structures is regarded as a class of finite Kripke structures with designated worlds, namely as the class of those $(\mathfrak{A}, a)$ for which $a \in Q^{\mathfrak{A}}$, then Q is $\sim$-invariant if it is closed with respect to $\sim$ within the class of all finite Kripke structures with designated worlds. A standard argument (using the boundedness of the class of finite Kripke structures with designated worlds and ML_∞-Scott sentences for these) shows that a monadic query on finite Kripke structures is $\sim$-invariant if and only if it is definable in ML_∞. The class of those monadic PTIME queries on finite Kripke structures that also are $\sim$-invariant may therefore suggestively be described as the intersection of PTIME with ML_∞:

$$\text{PTIME} \cap \mathrm{ML}_\infty.$$

Definition 1.11 A monadic query Q on finite Kripke structures is definable in L^k_μ if there is a formula $\varphi(x) \in \mathrm{L}^k_\mu$ such that for all finite $\mathfrak{A}$: $\varphi[\mathfrak{A}] = Q^{\mathfrak{A}}$.

As stated above, any L^k_μ-definable query is bisimulation invariant, and also in PTIME. To see the latter, recall that any k-ary monotone fixed-point operation over finite $\mathfrak{A}$ reaches saturation within $|A|^k + 1$ steps. We now state the main theorem as follows.

Theorem 1.12 *A monadic query Q on finite Kripke structures is in* PTIME and *bisimulation invariant if and only if it is definable in L^k_μ for some k. In formulae*

$$\text{PTIME} \cap \mathrm{ML}_\infty \;\equiv\; \mathrm{L}^\omega_\mu.$$

Corollary 1.13 *In particular* $\text{PTIME} \cap \mathrm{ML}_\infty$ *admits a recursive presentation in the sense of descriptive complexity: there is a language with recursive syntax and with* PTIME *semantics which is semantically complete for the class of all* PTIME *bisimulation-invariant queries over finite Kripke structures.*

For some background on the underlying notion of *capturing complexity classes* compare [16, 12, 29]. The following section is devoted to the proof of the non-trivial inclusion $\text{PTIME}\cap\mathrm{ML}_\infty \subseteq \mathrm{L}^\omega_\mu$, i.e. to showing that any given PTIME computable and bisimulation-invariant monadic query is definable in L^k_μ for sufficiently large k.

2 Proof of the main theorem

2.1 Canonical structures and a normal form

Let for a Kripke structure $\mathfrak{A} = (A, E, P_1, \dots, P_l)$ and an element $a \in A$, $\langle a\rangle^E$ denote the *forward E-closure of a*:

$$\langle a\rangle^E = \big\{ b \in A \;\big|\; \exists n\, \exists a_1 \dots a_n \text{ such that } a_1 = a, a_n = b \text{ and } (a_i, a_{i+1}) \in E^{\mathfrak{A}} \big\}.$$

Clearly $(\mathfrak{A}, a) \sim (\mathfrak{A}, a) \restriction \langle a\rangle^E$. Bisimulation equivalence over $\mathfrak{A}$ may be factored out to obtain a quotient Kripke structure which is a minimal representative for the $\sim$-class of the given $(\mathfrak{A}, a)$. We want to denote the result of applying this process to $(\mathfrak{A}, a) \restriction \langle a\rangle^E$ by $\mathrm{can}(\mathfrak{A}, a)$ and call it the *canonical structure for* $(\mathfrak{A}, a)$. Compare the *strongly extensional quotients* in [4, 5]. Explicitly,

$$\mathrm{can}(\mathfrak{A}, a) = \big(\langle a\rangle^E/\!\sim, E^\sim, P_1^\sim, \dots, P_l^\sim, [a]\big),$$

where $\sim$ is bisimulation equivalence (viewed as a binary relation over A), $[b]$ denotes the $\sim$-equivalence class of b, and the predicates $E^{\sim}$ and $P_i^{\sim}$ are defined as follows. $P_i^{\sim} = \{[b] \mid b \in P_i^{\mathfrak{A}} \cap \langle a \rangle^E\}$, which is sound, since $\sim$ is a congruence for the P_i. $E^{\sim}$ is defined according to

$$([b], [b']) \in E^{\sim} \quad \text{if} \quad \exists b'' \, (b, b'') \in E^{\mathfrak{A}} \wedge b' \sim b'',$$

where again, it is checked that the defining clause is independent of the choice of b within its equivalence class. Note, however, that $\sim$ is not a congruence with respect to E.

It is obvious that $\mathrm{can}(\mathfrak{A}, a)$ is computable from $(\mathfrak{A}, a)$ in PTIME. It is also straightforward to verify that for all $(\mathfrak{A}, a)$: $(\mathfrak{A}, a) \sim \mathrm{can}(\mathfrak{A}, a)$ and that for all $(\mathfrak{A}, a)$, $(\mathfrak{A}', a')$

$$(\mathfrak{A}, a) \sim (\mathfrak{A}', a') \quad \text{iff} \quad \mathrm{can}(\mathfrak{A}, a) \simeq \mathrm{can}(\mathfrak{A}', a').$$

In the special case that $\mathfrak{A} = \mathfrak{A}'$, i.e. that we consider $\sim$ as a binary global relation over individual Kripke structures, this becomes: $a \sim^{\mathfrak{A}} a'$ iff $\mathrm{can}(\mathfrak{A}, a) = \mathrm{can}(\mathfrak{A}, a')$, since $a \sim a'$ in particular also implies that $\langle a \rangle^E / \sim \, = \langle a' \rangle^E / \sim$ (even though not necessarily $\langle a \rangle^E = \langle a' \rangle^E$).

Definition 2.1 Let $\mathcal{CAN}$ stand for the class of all *canonical Kripke structures* with distinguished worlds, $\mathcal{CAN} = \{(\mathfrak{A}, a) \mid \mathrm{can}(\mathfrak{A}, a) \simeq (\mathfrak{A}, a)\}$. Let $\mathcal{CAN}_{\mathrm{fin}}$ be the class of finite structures in $\mathcal{CAN}$.

It is immediate that $(\mathfrak{A}, a) \in \mathcal{CAN}$ if and only if the following are satisfied:

(i) A consists of the forward E-closure of a: $\langle a \rangle^E = A$.

(ii) $\sim$ on $\mathfrak{A}$ coincides with the identity relation: for all $b, b' \in A$, $b \sim b' \Leftrightarrow b = b'$.

Returning to can: $(\mathfrak{A}, a) \mapsto \mathrm{can}(\mathfrak{A}, a)$, we note that the passage to $\mathrm{can}(\mathfrak{A}, a)$ may be seen as a canonization procedure that (almost) picks unique representatives from bisimulation equivalence classes.[2] The interested reader is referred to the account of canonization and invariants that characterize structures up to corresponding notions of equivalence in a different though related context of k-variable equivalence [29].

Passage to $\mathrm{can}(\mathfrak{A}, a)$ may also be used as a filter to enforce bisimulation invariance of queries. Let $\mathcal{A}$ be an algorithm that recognizes some monadic query Q_0 on finite Kripke structures and identify Q_0 with the class of those $(\mathfrak{A}, a)$ for which $a \in Q^{\mathfrak{A}}$. Let $\mathcal{A} \circ \mathrm{can}$ denote the algorithm which on input $(\mathfrak{A}, a)$ first computes $\mathrm{can}(\mathfrak{A}, a)$ and then applies $\mathcal{A}$ to the outcome. Then $\mathcal{A} \circ \mathrm{can}$ recognizes the $\sim$-invariant query $Q = \{(\mathfrak{A}, a) \mid \mathrm{can}(\mathfrak{A}, a) \in Q_0\}$. Furthermore, if $\mathcal{A}$ is in PTIME then so is $\mathcal{A} \circ \mathrm{can}$, and Q_0 is itself $\sim$-invariant if and only if $Q_0 = Q$. In other words we have the following.

Proposition 2.2 *For any monadic query Q on finite Kripke structures the following are equivalent:*

(i) Q is $\sim$-invariant.

[2] '*almost*' because in general the result is really only determined up to isomorphism. For finite Kripke structures we could do better: we shall see that the $\mathrm{can}(\mathfrak{A}, a)$ carry a definable and PTIME computable linear ordering. This can be used to obtain standard representations of the $\mathrm{can}(\mathfrak{A}, a)$ with initial segments of the natural numbers for their universes. See Proposition 2.12

(ii) $Q = \{(\mathfrak{A}, a) \mid \mathrm{can}(\mathfrak{A}, a) \in Q\}$.

Note that $\{(\mathfrak{A}, a) \mid \mathrm{can}(\mathfrak{A}, a) \in Q\}$ is in PTIME if Q itself is.

We obtain two corollaries from this, the first one gives an *identity property* for bisimulation-invariant queries, the second one a normal form for PTIME $\cap$ ML_∞.

Corollary 2.3 *For any two bisimulation-invariant queries on finite Kripke structures, Q and Q': $Q = Q'$ if $Q \cap \mathcal{CAN}_{fin} = Q' \cap \mathcal{CAN}_{fin}$.*

Corollary 2.4 *The queries in* PTIME $\cap$ ML_∞ *are exactly those that are obtained as compositions of arbitrary* PTIME *queries with* can*:*

$$\mathrm{PTIME} \cap \mathrm{ML}_\infty \equiv \mathrm{PTIME} \circ \mathrm{can}.$$

The notorious difficulties of capturing complexity classes – e.g. of finding a recursive syntax with PTIME semantics for the class of all PTIME queries on finite graphs – has to do with the implicit isomorphism invariance that is required of queries. For Kripke structures we have so far established a smooth passage from arbitrary queries (subject to isomorphism-invariance) to bisimulation-invariant queries. This passage lends itself to a translation of the capturing issue. But why should it help? The answer is, that the domain to which we have reduced the original problem through Proposition 2.2 has *the* crucial advantage on which all known capturing results ultimately rely: over $\mathcal{CAN}$, there is a definable and PTIME computable, in fact LFP-definable, global linear ordering. This allows us to reduce the present capturing issue to the well known Immerman-Vardi Theorem, that least fixed-point logic LFP captures PTIME over linearly ordered finite structures. This reduction is dealt with in the following section.

Before that, we briefly give some indication that $\mathcal{CAN}_{\mathrm{fin}}$ is sufficiently rich for an interesting complexity theory. The descriptive complexity of bisimulation-invariant queries actually reflects all the richness of structural complexity.

Richness of the bisimulation-invariant scenario *Word models* are structural encodings of words. Let some alphabet $\{p_1, \ldots, p_l\}$ be fixed. A word $w \in \{p_1, \ldots, p_l\}^n$ of length $n > 0$ is encoded by the structure

$$\mathfrak{A}_w = \big(\{0, \ldots, n-1\}, \mathrm{succ}^n, P_1^w, \ldots, P_l^w\big),$$

where succ^n is the usual successor relation in restriction to $n = \{0, \ldots, n-1\}$, and each $P_i^w \subseteq n$ contains those positions j in w that carry the letter p_i. If we regard $(\mathfrak{A}_w, 0)$ as a Kripke structure with designated world 0 (marking the first position in w) then clearly $(\mathfrak{A}_w, 0) \in \mathcal{CAN}_{\mathrm{fin}}$. Let $\mathcal{W} \subseteq \mathcal{CAN}_{\mathrm{fin}}$ be the class of these Kripke word models.

The encoding of words through canonical Kripke structures immediately allows to embed all string languages faithfully into corresponding bisimulation-invariant classes of finite Kripke structures: to $L \subseteq \{p_1, \ldots, p_l\}^+$ associate $W_L \subseteq \mathcal{W}$, the class of $(\mathfrak{A}_w, 0)$ for $w \in L$, and $Q_L \subseteq \mathcal{CAN}_{\mathrm{fin}}$, the class of all $(\mathfrak{B}, b)$ that are bisimulation-equivalent with some $(\mathfrak{A}_w, 0) \in W_L$. At the level of PTIME and above, the complexity of L and W_L is the same as that of Q_L.

Since the considerations of Proposition 2.2 apply to complexity classes above PTIME just as for PTIME, it is clear that for instance PSPACE $\cap$ $\mathrm{ML}_\infty \equiv$ PSPACE $\circ$ can just as

$\mathrm{P}\textsc{time} \cap \mathrm{ML}_\infty \equiv \mathrm{P}\textsc{time} \circ \mathrm{can}$. It therefore follows from the embeddability of string languages that for instance $\mathrm{P}\textsc{space} \cap \mathrm{ML}_\infty = \mathrm{P}\textsc{time} \cap \mathrm{ML}_\infty$ if and only if PSPACE collapses to PTIME over $\mathcal{CAN}_{\mathrm{fin}}$ if and only if PTIME = PSPACE.

The embeddability of string languages also shows that $\mathrm{L}_\mu \subsetneq \mathrm{P}\textsc{time} \cap \mathrm{ML}_\infty$, and indeed that $\mathrm{L}_\mu \subsetneq \mathrm{L}^2_\mu$. Recall that by the theorem of Büchi, Elgot and Trakhtenbrot, $W_L \subseteq \mathcal{W}$ is MSO-definable (over $\mathcal{W}$) exactly if L is regular, see for instance [12]. It is well known on the other hand that $\mathrm{L}_\mu \subseteq \mathrm{MSO}$ – the essential observation is that monadic least fixed points have obvious MSO-definitions as $\subseteq$-minimal fixed points.

Example 2.5 Consider any PTIME recognizable language L which is not regular, for instance $L = \{p_1^n p_2^n \mid n \geqslant 1\}$. Then Q_L is a bisimulation-invariant PTIME query but not definable in MSO. Otherwise $Q_L \cap \mathcal{W} = W_L$ would be MSO-definable over $\mathcal{W}$ and this would imply that L is regular. Now $W_{L'}$ for $L' = \{p_1^n p_2^m \mid n, m \geqslant 1\}$ is MSO-definable. The following L^2_μ-formula defines W_L within $W_{L'}$:

$$\varphi(x) = \mu_Y \Big[\Big(\mu_X \big[(\varphi_1(x_1) \wedge \varphi_2(x_2)) \vee \Diamond_1 \Diamond_2 X x_1 x_2 \big] \wedge \varphi_1(x_2) \Big) \vee \Diamond_2 Y x_1 x_2 \Big] (x/x_1, x/x_2),$$

where $\varphi_1(x_i) = P_1 x_i \wedge \Box_i P_2 x_i$ and $\varphi_2(x_i) = \Diamond_i (P_2 x_i \wedge \Box_i \bot)$. Here $\bot$ stands for any universally false statement, like $P_1 x_1 \wedge \neg P_1 x_1$. Over any structure in $W_{L'}$, $\varphi_1(x_i)$ defines the last position carrying the letter p_1, and $\varphi_2(x_i)$ the position before the last p_2. Observe that the $\Diamond_1\Diamond_2$-process for X corresponds to a synchronized back-stepping in x_1 and x_2. So $W_L = \big\{ (\mathfrak{A}_w, 0) \in W_{L'} \;\big|\; \mathfrak{A}_w \models \varphi[0] \big\}$. It follows that φ is not equivalent with any MSO-formula, and we have a separation of L_μ from L^2_μ.

2.2 Order and the Immerman-Vardi Theorem

Among the prominent logics in finite model theory are various fixed-point extensions of first-order logic, most notably *least fixed-point logic* LFP. We here only sketch the definitions of LFP and the related *inductive fixed-point logic* IFP to make them available for technical applications. For more background the reader should see for instance [12].

LFP is the extension of first-order logic that is obtained through closure under the formation of least fixed points of positively defined operations on predicates. If φ is a formula in which the second-order variable X occurs positively (no free occurrence in the scope of an odd number of negations), X of arity r and $\overline{x}$ a tuple of r distinct first-order variables, $\overline{y}$ any tuple of r first-order variables, then

$$\psi = \big[\mathrm{LFP}_{X,\bar{x}} \varphi \big] (\overline{y})$$

is also a formula (in which the $\overline{y}$ are free and X is not). ψ asserts of $\overline{y}$ that it is contained in the least fixed point of the monotone operation

$$P \mapsto \big\{ \overline{x} \;\big|\; \varphi[P/X] \big\}.$$

Least-fixed point logic LFP is the smallest extension of first-order logic that is closed under first-order operations and the LFP-constructor.

Inductive fixed-point logic IFP similarly extends first-order by the formation of *inductive fixed points* of operations on predicates. For φ and X, $\overline{x}$ and $\overline{y}$ as above, but φ not necessarily positive in X,

$$\psi = \big[\mathrm{IFP}_{X,\bar{x}} \varphi \big] (\overline{y})$$

is also a formula which asserts that $\overline{y}$ is in the limit of the inductive sequence of r-ary relations X_α generated as

$$\begin{array}{lcl} X_0 & = & \emptyset \\ X_{\alpha+1} & = & X_\alpha \cup \{\overline{x} \mid \varphi[X_\alpha/X]\} \\ X_\lambda & = & \bigcup_{\alpha<\lambda} X_\alpha \quad \text{for limits } \lambda. \end{array}$$

Theorem 2.6 (Gurevich, Shelah, [18]) *Over finite structures* LFP *and* IFP *have the same expressive power: a query is* LFP*-definable if and only if it is* IFP*-definable.*

A class of finite structures admits an LFP-definable global ordering if there is an LFP-definable binary query which on all members of that class evaluates to a total linear ordering.

Theorem 2.7 (Immerman [21], Vardi [34]) *A query on linearly ordered structures is in* PTIME *if and only if it is definable in* LFP*. The same holds of* PTIME *queries in restriction to classes of structures that admit an* LFP*-definable global linear ordering.*

Recall that $\mathcal{CAN}_{\rm fin}$ is the class of all $\mathrm{can}(\mathfrak{A}, a)$ where $(\mathfrak{A}, a)$ is any finite Kripke structure (of our fixed type $\overline{P}$).

Proposition 2.8 $\mathcal{CAN}_{\mathit{fin}}$ *admits an* LFP*-definable global ordering.*

Proof. In view of the Gurevich-Shelah Theorem we need only show that a linear order can uniformly be obtained through an IFP-like inductive process over all $\mathrm{can}(\mathfrak{A}, a)$. The idea is the same as in the colour refinement for finite graphs (cf. the detailed treatment in [29]): we know that $\sim$ (or rather its complement $\not\sim$) is definable in a fixed-point process, and the elements of $\mathrm{can}(\mathfrak{A}, a)$ just are the $\sim$-classes. The refinement steps in the generation of $\not\sim$ as a limit of successive stages $\not\sim_i$ according to (1) in section 1.1 may be adapted to generate an ordered representation of the $\sim_i$-classes in each stage. The resulting limit will be a linear ordering of the $\sim$-classes.

Recall that $\sim_0$ is atomic equivalence with respect to $\overline{P} = (P_1, \ldots, P_l)$. Enumerate the 2^l atomic $\overline{P}$-types in some fixed order, and let $\prec_0$ be the strict pre-ordering induced by this enumeration: $[b] \prec_0 [b']$ if the atomic type of b precedes that of b' in this fixed enumeration. Clearly $\prec_0$ is a first-order definable global relation.

The elements of $\mathrm{can}(\mathfrak{A}, a)$ are $\sim$-classes. Since $\sim$ is the common refinement of the $\sim_i$, the $\sim_i$-classes are represented over $A/{\sim}$ by sets of $\sim$-classes. In this picture, the limit $\sim$ is reached with that $\sim_i$ whose classes are singleton subsets of $A/{\sim}$. We shall inductively define the $\prec_i$ so that $\prec_i$ induces a linear ordering of the $\sim_i$-classes, i.e. such that each

$$\big(A/{\sim}, \prec_i\big) \;/\; {\sim_i}$$

is a well defined strict linear ordering. The limit $\prec \;:=\; \bigcup_i \prec_i$ will provide a linear ordering on $A/{\sim}$ and similarly on the universe $\langle a\rangle^E/{\sim}$ of $\mathrm{can}(\mathfrak{A}, a)$.

Define $\prec_{i+1}$ in terms of $\prec_i$ through

$$[b] \prec_{i+1} [b'] \quad \text{iff} \quad \begin{array}{l} [b] \prec_i [b'] \\ \quad\quad \text{or} \\ [b] \sim_i [b'] \text{ but } \Gamma\big([b], \prec_i\big) <_{\text{lex}} \Gamma\big([b'], \prec_i\big), \end{array}$$

where $<_{\text{lex}}$ is lexicographic comparison, applied to boolean tuples $\Gamma\big([b], \prec_i\big)$ which are defined as follows. Let $\gamma_1, \ldots, \gamma_t$ be the enumeration of $\sim_i$-classes in $\prec_i$-increasing order. Then $\Gamma\big([b], \prec_i\big) = (\imath_1, \ldots, \imath_t) \in \{0,1\}^t$ where

$$\imath_s = \begin{cases} 1 & \text{if } \exists\, [c] \Big([c] \in \gamma_s \wedge ([b],[c]) \in E^{\sim} \Big) \\ 0 & \text{else.} \end{cases}$$

In comparison with the inductive generation of the $\not\sim_\alpha$ in (1) in section 1.1 it can be shown inductively that $\prec_i$ is a linear ordering of the $\sim_i$-classes. In other words

$$[b] \not\sim_i [b'] \quad \text{iff} \quad [b] \prec_i [b'] \ \text{ or } \ [b'] \prec_i [b].$$

It follows that the limit $\prec := \bigcup_i \prec_i$ does indeed define a linear ordering of the $\sim$-classes, i.e. a linear ordering on the universe of $\operatorname{can}(\mathfrak{A}, a)$.

It is also not hard to see that the crucial lexicographic comparison in the refinement step is first-order definable in the sense that there is a first-order formula that defines $\prec_{i+1}$ in terms of $\prec_i$ and $E^{\sim}$. Altogether this shows that the global ordering $\prec$ is IFP-definable, hence LFP-definable, over $\mathcal{CAN}_{\text{fin}}$. $\Box$

Corollary 2.9 *A bisimulation-invariant monadic query Q on finite Kripke structures is in* PTIME *if and only if* $\operatorname{can}(Q) = \big\{ \operatorname{can}(\mathfrak{A}, a) \bigm| a \in Q^{\mathfrak{A}} \big\}$ *is definable in* LFP *over* $\mathcal{CAN}_{\mathit{fin}}$*. In formulae this may be summed up more suggestively as*

$$\text{PTIME} \cap \mathrm{ML}_\infty \equiv \mathrm{LFP} \circ \operatorname{can}.$$

The above method of constructing inductively an ordering of the types was conceived by Abiteboul and Vianu [1] in the setting of relational computation, where it became instrumental in their fundamental investigation of least versus partial fixed point. The logical and game-theoretic formulations, on which the present treatment is modeled, were abstracted and applied to bounded-variable logics in [11, 14, 27]. [28] actually contains an account of the bounded-variable case of this technique in terms of the bisimulation analysis of Kripke structures that encode the bounded-variable games.

Recently Sazonov and Lisitsa [32] have employed the same idea of a lexicographic pre-ordering of bisimulation equivalence classes to obtain a definable order on the hereditarily finite sets in non-wellfounded set theory with the anti-foundation axiom. The connection with the above development is apparent, since in that framework bisimulation equivalence is set identity.

2.3 Getting it all into L_μ^k

Consider L_μ^k-definability. From Lemma 1.8 we know that any L_μ^k-definable query is bisimulation-invariant. We wish to show that conversely any query in $\mathrm{PTIME} \cap \mathrm{ML}_\infty$ is L_μ^k-definable for some k. From Corollary 2.9 we already have over $\mathcal{CAN}_{\mathrm{fin}}$ that $\mathrm{PTIME} \cap \mathrm{ML}_\infty \equiv \mathrm{LFP}$. Using the identity property for bisimulation-invariant queries expressed in Corollary 2.3, it will suffice to show that $\mathrm{LFP} \equiv \mathrm{L}_\mu^\omega$ over $\mathcal{CAN}_{\mathrm{fin}}$, or, for the non-trivial inclusion, $\mathrm{LFP} \subseteq \mathrm{L}_\mu^\omega$ over $\mathcal{CAN}_{\mathrm{fin}}$.

Proposition 2.10 *Let $\varphi(\overline{x})$ be an* LFP*-formula in the language of Kripke structures with distinguished worlds $(\mathfrak{A}, a)$ (where a is regarded as a constant). Then there is some k and $\varphi^* \in \mathrm{L}_\mu^k$ such that the free variables $\overline{x}$ of φ are among $x_1, \dots, x_{k-1}$, and such that for all $(\mathfrak{A}, a) \in \mathcal{CAN}$ and all $\overline{b} \in A^{k-1}$:*

$$(\mathfrak{A}, a) \models \varphi[\overline{b}] \quad \textit{iff} \quad \mathfrak{A} \models \varphi^*[\overline{b}, a].$$

For LFP*-sentences φ in particular, $(\mathfrak{A}, a) \models \varphi$ if and only if $(\mathfrak{A}, a) \models \varphi^*$ (in the standard monadic semantics):* $\mathrm{LFP} \subseteq \mathrm{L}_\mu^\omega$ *for monadic queries over $\mathcal{CAN}$.*

This proves Theorem 1.12 as follows. Let $Q \in \mathrm{PTIME} \cap \mathrm{ML}_\infty$ be a monadic query. By Corollary 2.9, Q is LFP-definable in restriction to $\mathcal{CAN}_{\mathrm{fin}}$. The proposition shows that Q is definable in restriction to $\mathcal{CAN}_{\mathrm{fin}}$ by some $\varphi(x) \in \mathrm{L}_\mu^k$ for suitable k. As the query defined by $\varphi(x)$ over arbitrary finite Kripke structures is $\sim$-invariant, and coincides with Q over $\mathcal{CAN}_{\mathrm{fin}}$, it coincides with Q over all finite Kripke structures by Corollary 2.3. So Q is L_μ^k-definable.

In preparation for Proposition 2.10 we prove a weak normal form for LFP.

Lemma 2.11 *Any formula of* LFP *is logically equivalent with one that – for some suitable value of r – satisfies the following conditions:*

(i) all first-order variables that occur (free or bound) are among $x_1, \dots, x_r$.

(ii) all applications of the fixed-point operator are of the form $[\mathrm{LFP}_{X,\bar{x}}\psi]\overline{x}'$ where the arity of X is r, $\overline{x} = (x_1, \dots, x_r)$, $\overline{x}'$ some r-tuple with entries from $x_1, \dots, x_r$.

Sketch of proof. Towards this claim consider a formula of type $[\mathrm{LFP}_{Y,\bar{x}_1}\psi(Y, \overline{x}_1, \overline{x}_2)]\overline{y}$, where we may assume without loss of generality that for some sufficiently large r and suitable tuple of variables $\overline{x}_3$:

– no variables besides $x_1, \dots, x_r$ occur in ψ or in $\overline{y}$,

– the concatenation $\overline{x} = \overline{x}_1\,\overline{x}_2\,\overline{x}_3$ is a permutation of $(x_1, \dots, x_r)$,

– no variable from $\overline{x}_2$ occurs bound, and no variable from $\overline{x}_3$ free, in ψ.

Let then X be a new second-order variable of arity r and let $\psi'(X, \overline{x})$ be obtained from ψ through replacing all atoms $Y\overline{z}$ by atoms $Z\overline{z}\,\overline{x}_2\,\overline{x}_3$. Then $[\mathrm{LFP}_{Y,\bar{x}_1}\psi(Y, \overline{x}_1, \overline{x}_2)]\overline{y}$ is universally equivalent with

$$[\mathrm{LFP}_{X,\bar{x}}\psi'(X, \overline{x})]\overline{y}\,\overline{x}_2\,\overline{x}_3.$$

This is verified via comparison of the individual stages in the respective fixed-point generations. For all $\mathfrak{A}$ and tuples $\overline{a}_1$ and $\overline{a}_2$ from A, if $Y_\alpha^{\mathfrak{A},\bar{a}_2}$ denotes the α-th iteration for ψ with parameters $\overline{a}_2$ for $\overline{x}_2$, and $X^{\mathfrak{A}}_\alpha$ the α-th iteration for ψ', then for all α these satisfy

$$\begin{aligned} \overline{a}_1 \in Y^{\mathfrak{A},\bar{a}_2}_\alpha \quad &\Leftrightarrow \quad \overline{a}_1\,\overline{a}_2\,\overline{a}_3 \in X^{\mathfrak{A}}_\alpha \text{ for some } \overline{a}_3 \\ &\Leftrightarrow \quad \overline{a}_1\,\overline{a}_2\,\overline{a}_3 \in X^{\mathfrak{A}}_\alpha \text{ for all } \overline{a}_3. \end{aligned}$$

Finally we apply a renaming of first-order variables throughout $[\mathrm{LFP}_{X,\bar{x}}\psi'(X,\overline{x})]$ to put $\overline{x} = (x_1,\dots,x_r)$ as required. □

Proof of Proposition 2.10. The proof is by induction on the syntax of LFP-formulae in normal form for fixed r. Put $k = r+1$. Let c be the constant symbol available in LFP for the designated worlds of Kripke structures. Note that, for the sake of the induction, free second-order variables X of arity r have to be admitted in LFP-formulae. These translate into free second-order variables X^* of arity $k = r+1$. Inductively we show that for $\varphi(\overline{X}, x_1,\dots,x_r)$ (in normal form), there is $\varphi^*(\overline{X}^*, x_1,\dots,x_k)$ such that for all $(\mathfrak{A}, a) \in \mathcal{CAN}$, and all interpretations $\overline{R}$ of the $\overline{X}$ over A, and all $\overline{b} \in A^r$:

$$(\mathfrak{A}, a) \models \varphi[\overline{R}, \overline{b}] \quad \Leftrightarrow \quad \mathfrak{A} \models \varphi^*[\overline{R}^*, \overline{b}, a],$$

where, for R in $\overline{R}$, R^* is any extension of R to arity $k = r+1$ such that $R = \{\overline{b} \,|\, \overline{b}a \in R^*\}$. For the induction it is also checked that positive occurrences of X in φ translate into positive occurrences of X^* in φ^*.

Recall from Lemma 1.9 that $\sim$ is L^k_μ-definable for each $k \geqslant 2$. Let $\varphi_\sim(\overline{x})$ be an L^k_μ-formula which defines (padded) bisimulation equivalence in the first and second component as an r-ary global relation $\sim^{\mathfrak{A}} = \{\overline{b} \,|\, (\mathfrak{A}, b_1) \sim (\mathfrak{A}, b_2)\}$. We shall simply write $x_i \sim x_j$ as shorthand for the corresponding relation defining $\sim$ in the i-th and j-th component as obtained from $\varphi_\sim$ through substitutions. Recall that over $(\mathfrak{A}, a) \in \mathcal{CAN}$, $b \sim b'$ if and only if $b = b'$, whence over $\mathcal{CAN}$ the L^k_μ-formula for $x_i \sim x_j$ defines equality $x_i = x_j$. The other characteristic property of $(\mathfrak{A}, a) \in \mathcal{CAN}$, that will become important for first-order quantification, is that $A = \langle a \rangle^E$. We turn to the syntactic induction.

(i) atomic φ. The following translations $\varphi \longmapsto \varphi^*$ satisfy the claim, for $1 \leqslant i, j, i_1, \dots, i_r < k$ and for all basic propositions P:

$$\begin{aligned} \big(x_i = x_j\big)^* &= x_i \sim x_j \\ \big(x_i = c\big)^* &= x_i \sim x_k \\ \big(Px_i\big)^* &= Px_i \\ \big(Xx_{i_1}\dots x_{i_r}\big)^* &= X^* x_{i_1}\dots x_{i_r} x_k \\ \big(Ex_i x_j\big)^* &= \Diamond_i \big(x_i \sim x_j\big). \end{aligned}$$

(ii) boolean combinations translate trivially.

(iii) existential quantification. Let $\varphi = \exists x_j \psi$ for some $1 \leqslant j < k$, and assume that ψ^* is as required for ψ. Then

$$\varphi^* = \big[\mu_X \big(\psi^*(\overline{x}) \vee \Diamond_j X\overline{x}\big)\big](x_k/x_j)$$

is as desired (where X is new). The least fixed point $\mu_X\big(\psi^*(\overline{x}) \vee \Diamond_j X\overline{x}\big)$ exactly comprises those tuples of worlds from which a positive instance for ψ^* can be reached on a forward E-path in the j-th component. The substitution of x_k (which in the semantic requirements for φ^* is set to a) in the j-th component thus extends the existential quantification appropriately to $\langle a\rangle^E = A$, for $(\mathfrak{A}, a) \in \mathcal{CAN}$.

(iv) fixed-point applications. Let $\varphi = \left[\mathrm{LFP}_{X,\bar{x}}\psi\right]\overline{x}'$, where $\overline{x} = (x_1, \ldots, x_r)$ according to normal form. As L^k_μ is closed under variable substitutions we may assume $\overline{x}' = \overline{x}$. Suppose that $\psi^*(X^*, x_1, \ldots, x_k)$ is as required for ψ, where second-order parameters may be suppressed in the notation since they play no active role. Then

$$\varphi^* = \mu_{X^*}\big(\psi^*(X^*, \overline{x})\big)$$

is as desired for φ. For this claim we consider inductively the appropriate claim for the individual stages in the fixed-point generation: if the stages of the inductive evaluation of $\mathrm{LFP}_{X,\bar{x}}\psi$ are denoted R_α, and if those of $\mu_{X^*}\psi^*$ are denoted R^*_α, then $\overline{b} \in R_\alpha \Leftrightarrow \overline{b}a \in R^*_\alpha$. This is obviously true for $R_0 = \emptyset = R^*_0$. Inductively the claim carries over to successor stages, since by the assumption on ψ^*, always

$$\begin{array}{rcl} \overline{b} \in R_{\alpha+1} & \Leftrightarrow & (\mathfrak{A}, a) \models \psi[R_\alpha, \overline{b}] \\ & \Leftrightarrow & \mathfrak{A} \models \psi^*[R^*_\alpha, \overline{b}, a] \\ & \Leftrightarrow & \overline{b}a \in R^*_{\alpha+1}. \end{array}$$

$\square$

2.4 Some further remarks

The capturing result Theorem 1.12 as proved in this section, generalizes in an obvious manner to a multi-modal framework of Kripke structures with several modalities and accessibilities $(E^{(\jmath)})_{\jmath=1\ldots m}$ rather than a single E. Bisimulation and bisimulation equivalence are defined analogously, only that the typical back-and-forth clauses (cf. (ii) and (iii) in Definition 1.1) have to be formulated for each of the individual $E^{(\jmath)}$. In the L^2_μ-definition of bisimulation inequivalence (cf. Lemma 1.9) this modification leads to the replacement of the $\Diamond_1\Box_2 X \vee \Diamond_2\Box_1 X$-part by a corresponding expression

$$\bigvee_{\jmath=1}^{m} \Diamond_1^{(\jmath)}\Box_2^{(\jmath)} X \vee \bigvee_{\jmath=1}^{m} \Diamond_2^{(\jmath)}\Box_1^{(\jmath)} X.$$

Canonical structures are obtained as quotients with respect to bisimulation equivalence just as before (cf. Section 2.1), only that the restriction of the domain to the forward E-closure of the distinguished world has to be replaced by the corresponding closure for $E := \bigcup_\jmath E^{(\jmath)}$. All the essential steps in the above treatment go through as before, in particular Corollary 2.4, Proposition 2.8, and Proposition 2.10 continue to hold unchanged. One of the few minor technical modifications is, that in the proof of Proposition 2.8, the lexicographic comparison obviously has to be carried out with respect to the tuple describing $E^{(\jmath)}$-incidence with classes of the previous level of refinement, for each $\jmath$. It follows that Theorem 1.12 holds also in this extended setting.

It should be noted that the *existence* of a logic for bisimulation-invariant PTIME, in the sense of descriptive complexity, follows directly from Corollary 2.9 which may be looked at as a normal form theorem or as a logical characterization. Putting the reasoning that leads to Corollary 2.9 in different perspective, the abstract capturing result we have here is due to the following.

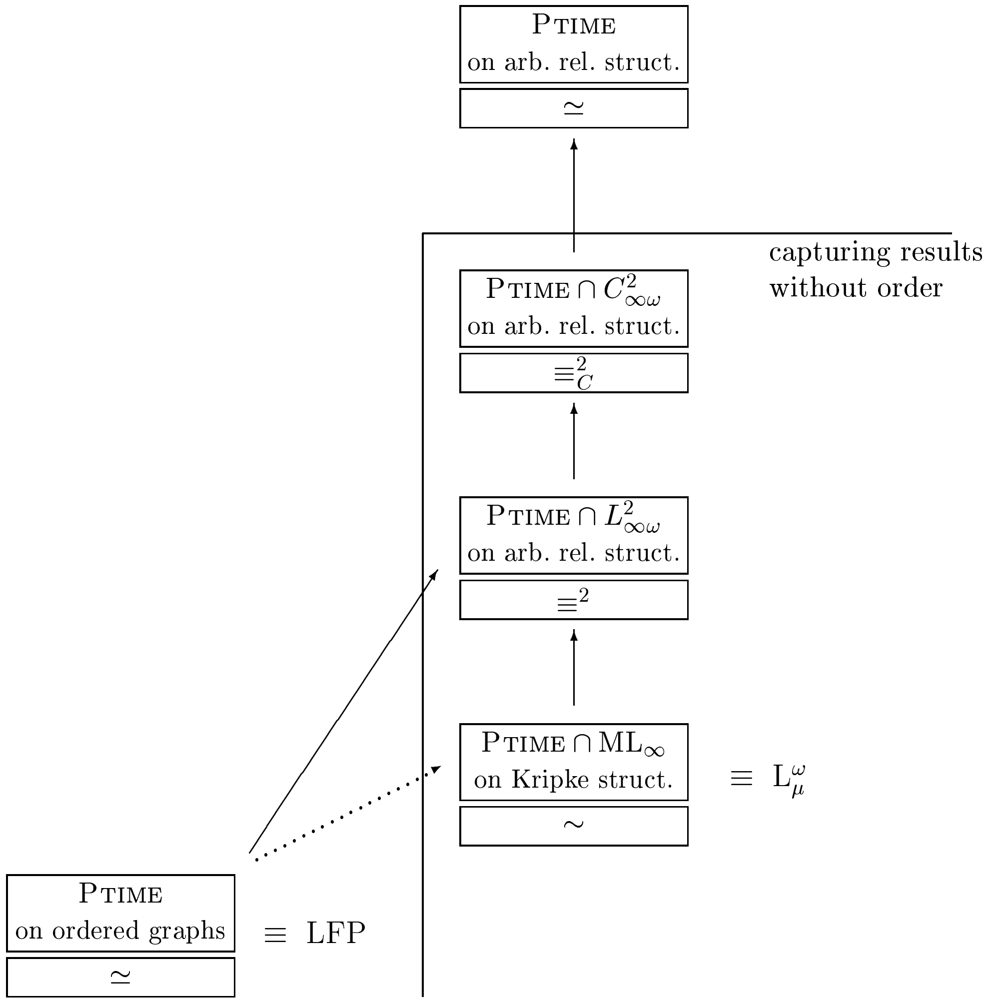
Proposition 2.12 *There is a* PTIME *functor* H *defined on finite Kripke structures with distinguished worlds, for canonization up to bisimulation equivalence:*

$$\begin{aligned}&\forall(\mathfrak{A},a) \quad (\mathfrak{A},a) \sim H(\mathfrak{A},a),\\ &\forall(\mathfrak{A},a)\forall(\mathfrak{A}',a') \quad (\mathfrak{A},a) \sim (\mathfrak{A}',a') \Rightarrow H(\mathfrak{A},a) = H(\mathfrak{A}',a').\end{aligned}$$

Such H is obtained through $(\mathfrak{A},a) \overset{\mathrm{can}}{\longmapsto} \mathrm{can}(\mathfrak{A},a) \overset{\mathrm{stan}}{\longmapsto} H(\mathfrak{A},a)$, where the functor 'can' is composed with a functor 'stan' that maps $\mathrm{can}(\mathfrak{A},a)$ to its isomorphic standard representation over an initial segment of the natural numbers, naturally ordered by the PTIME computable global ordering $\prec$ (according to Proposition 2.8).

PTIME canonization functors generally induce capturing results for corresponding fragments of PTIME, as outlined in [29]. Indeed, two other interesting capturing results for fragments of PTIME could be obtained along these lines [28]. These concern the full two-variable fragments of infinitary first-order logic and its extension by counting quantifiers. In those cases, canonization requires far more elaboration, since the passage to quotients (as in the formation of $\mathrm{can}(\mathfrak{A},a)$) does not lead to structures of the original kind. The reconstruction of a standard representative from its concise quotient description becomes an essential, and technically involved, step in those arguments. It just so happens that for bisimulation invariance the quotient structures themselves are canonical representatives of their equivalence class. Note, however, that the present capturing result is not a consequence of the capturing result for the two-variable fragments PTIME $\cap\, L^2_{\infty\omega}$ and PTIME $\cap\, C^2_{\infty\omega}$ in [28], although the latter results are stronger in the sense that PTIME $\cap\, \mathrm{ML}_\infty \subsetneq$ PTIME $\cap\, L^2_{\infty\omega} \subsetneq$ PTIME $\cap\, C^2_{\infty\omega}$. Compared with the harder capturing results for those two-variable fragments, however, we have here obtained a syntactically much nicer and more natural explicit capturing result.

Let PTIME *on ordered graphs* stand for the class of all PTIME boolean queries on ordered graphs, which by the Immerman-Vardi Theorem is captured by LFP. Any class of ordered graphs that is closed under isomorphism is also $L^2_{\infty\omega}$-definable, so PTIME on ordered graphs is a subclass of PTIME $\cap\, L^2_{\infty\omega}$. Interestingly, it may even be embedded into PTIME $\cap\, \mathrm{ML}_\infty$, though. To this end we need merely regard an ordered graph as a Kripke structure, in the multi-modal framework. We may take the edge relation as one accessibility relation and the successor relation, which comes with the ordering, as a second one. The initial vertex with respect to the ordering is regarded as the distinguished world, just as for the Kripke versions of word models considered above. For these canonical representations of ordered graphs as Kripke structures, bisimulation equivalence coincides with isomorphism. In terms of this representation, therefore, PTIME on ordered graphs translates into a subclass of PTIME $\cap\, \mathrm{ML}_\infty$. The containments between these fragments of PTIME, together with a specification of the associated invariance conditions, are indicated in the following diagram, where arrows stand for strict inclusion and the dotted arrow indicates strict inclusion under the translation just outlined.



For the analysis of bisimulation-invariant queries in higher arity, the right notion of equivalence is *component-wise bisimulation equivalence*: $(\mathfrak{A},\overline{a}) \sim (\mathfrak{A}',\overline{a}')$ for k-tuples of worlds $\overline{a}$ and $\overline{a}'$, if $(\mathfrak{A},a_j) \sim (\mathfrak{A}',a'_j)$ for $j = 1,\dots,k$. Compare Lemma 1.8. Over finite Kripke structures k-ary bisimulation-invariant queries are exactly the ML^k_∞-definable queries, so the corresponding fragment of PTIME is PTIME $\cap\, \mathrm{ML}^k_\infty$. The analogue of the functor can: $(\mathfrak{A},a) \mapsto \mathrm{can}(\mathfrak{A},a)$ then is the extension to arguments $(\mathfrak{A},\overline{a})$ for $\overline{a} \in A^k$ through $\mathrm{can}(\mathfrak{A},\overline{a}) := \big(\mathrm{can}(\mathfrak{A},a_j)\big)_{1\leqslant j\leqslant k}$. Note the high degree of independence between the components a_j, which is appropriate according to the definition of $(\mathfrak{A},\overline{a}) \sim (\mathfrak{A}',\overline{a}')$. Passage to $\mathrm{can}(\mathfrak{A},\overline{a})$ serves to translate $\sim$-invariant k-ary queries into queries over $(\mathcal{CAN}_{\mathrm{fin}})^k$, or over the class of disjoint unions of k canonical Kripke structures with one distinguished world in each component. In analogy with the monadic case one could similarly show that PTIME $\cap\, \mathrm{ML}^k_\infty$ coincides with the class of k-ary queries definable L^ω_μ.

A more technical remark concerns the relation between L_μ^k and plain multi-modal L_μ in application to the k-th Cartesian power of the $\mathfrak{A}$ with accessibility relations E_j and propositional constants P_{ij} for atomic L_μ^k-formulae $P_i x_j$. This translation is actually reminiscent of a similar reduction of finite-variable equivalence to bisimulation equivalence of suitably defined powers that is sketched in [30]. The only construct in L_μ^k that introduces a difference is variable substitution. It should be noted that variable substitutions introduce some degree of non-locality (in terms of E and the E_j). This raises the question to which extent variable substitutions are actually necessary in order to guarantee the expressive power needed for $\mathrm{ML}_\infty \cap \mathrm{P}\textsc{time}$. There seems to be one essential application of variable substitutions that is not easily avoidable. This occurs in the translation of existential quantification as carried out in the proof of Proposition 2.10. In this context substitutions are applied to set one component to the fixed parameter designating the distinguished world a of $(\mathfrak{A}, a)$. Allowing for just these special substitutions (of a constant c, say, for variables) one could obtain a system that otherwise resembles even more closely plain L_μ in application to k-th Cartesian powers.

3 Undecidability of L_μ^2

For the considerations of this section, again, the intended semantics of L_μ^ω is the standard monadic one. Satisfiability considerations thus concern the existence of Kripke structures with designated worlds satisfying L_μ^k-formulae in a single free variable.

Recall from Corollary 1.7 that L_μ^ω is a two-variable logic in the sense that $\mathrm{L}_\mu^k \subseteq L_{\infty\omega}^2$ for the standard monadic semantics over finite structures. Of course $\mathrm{L}_\mu^2 \subseteq \mathrm{ML}_\infty \subseteq L_{\infty\omega}^2$ over finite structures even for the general semantics. The undecidability result presented here should thus be seen in connection with the programme to investigate the borderline of decidability in the vicinity of two-variable logics and of L_μ. Several undecidable extensions of two-variable first-order logic $L_{\omega\omega}^2$ are exhibited in [15]. Among these undecidable extensions is a weak two-variable fragment $\mathrm{LFP}_{\mathrm{mon}}^2$ of monadic least fixed-point logic. $\mathrm{LFP}_{\mathrm{mon}}^2$ is the least common extension of $L_{\omega\omega}^2$ and L_μ, which are both decidable. L_μ^2 goes beyond L_μ in a different direction, however. $L_{\omega\omega}^2 \not\subseteq \mathrm{L}_\mu^2$, and, unlike $\mathrm{LFP}_{\mathrm{mon}}^2$, $\mathrm{L}_\mu^2 \not\subseteq \mathrm{MSO}$. Thus $\mathrm{LFP}_{\mathrm{mon}}^2$ and L_μ^2 are incomparable (even over finite Kripke structures, and with respect to monadic semantics).

It is not hard to see that the k-dimensional modal logics ML^k (i.e. L_μ^k without fixed points) are decidable even for the general semantics; for the monadic semantics ML^k collapses to plain ML anyway, [7].

We now consider the satisfiability problem for the L_μ^k and show that satisfiability in finite Kripke structures with designated worlds (*finite satisfiability*) is undecidable (r.e.-complete) for L_μ^2 while *satisfiability* for L_μ^2 – in arbitrary Kripke structures with designated worlds – is even Σ_1^1-hard (Σ_1^1 is the first level of the analytical hierarchy, above all arithmetical levels of undecidability). This situation is in sharp contrast with that for the classical μ-calculus: for L_μ the finite satisfiability problem and the satisfiability problem coincide (L_μ has the finite model property) [24], and are decidable [25, 33].

For notational convenience, let $sat(\mathcal{L})$, respectively $\mathit{fin}\text{-}sat(\mathcal{L})$, stand for the sets of those $\mathcal{L}$-formulae that have a model, respectively have a finite model. $\mathcal{L}$ has the finite model property if $sat(\mathcal{L}) = \mathit{fin}\text{-}sat(\mathcal{L})$.

Theorem 3.1 L^2_μ *does not have the finite model property, fin-sat*(L^2_μ) *is r.e.-complete, and sat*(L^2_μ) *is* Σ^1_1*-hard.*

As a logic for bisimulation-invariant PTIME, $\mathrm{L}^\omega_\mu = \bigcup \mathrm{L}^k_\mu$ is of course semantically determined only with respect to finite structures. The general satisfiability problem $sat(\mathrm{L}^\omega_\mu)$ thus concerns L^ω_μ as a natural extension of the classical μ-calculus, rather than as a logic for PTIME$\cap\mathrm{ML}_\infty$. But also for finite satisfiability, decidability is a property of the syntax of a logic. Could it be then that there is a different syntax for PTIME$\cap\mathrm{ML}_\infty$ that would be decidable for finite satisfiability? The answer is no, if the question is formalized in such a way as to preclude obvious trivializations of the issue. The undecidability argument for L^2_μ will show that no logic which *effectively captures* PTIME $\cap\, \mathrm{ML}_\infty$ over finite Kripke structures can be decidable for finite satisfiability. The crucial point about *effective* capturing is, that we require a recursive passage from some standard encoding of algorithms (like for instance natural finite encodings of Turing machines) to formulae. This condition is indeed satisfied for L^ω_μ: for LFP on $\mathcal{CAN}_{\mathrm{fin}}$ this is part of the capturing result of Immerman and Vardi, and the passage from LFP to L^ω_μ over $\mathcal{CAN}$ in the proof of Proposition 2.10 is recursive, too.

Remark 3.2 *If $\mathcal{L}$ is a logic that effectively captures bisimulation-invariant* PTIME *over finite Kripke structures, then fin-sat*$(\mathcal{L})$ *is undecidable.*

Indeed, this abstract undecidability claim may directly be inferred from the observation that there are sets of PTIME-recognizable subclasses of $\mathcal{CAN}_{\mathrm{fin}}$ for which the emptiness problem is undecidable. But rather than pursue the issue on this level, we take a more concrete and specific look at *domino problems* and their formalization in L^2_μ. These domino problems can serve as particular and intuitive examples for the claim just made. But the technical results of Theorem 3.1 and Lemma 3.7 below, for L^2_μ yield more:

- undecidability is located at the lowest possible level in terms of the L^k_μ, given that $\mathrm{L}^1_\mu = \mathrm{L}_\mu$ is decidable.
- since the natural semantics of L^2_μ extends to infinite Kripke structures, we get the extra result about Σ^1_1-hardness of the satisfiability problem (while in the abstract setting of capturing PTIME $\cap\, \mathrm{ML}_\infty$ this question would be meaningless).

A simple example showing that L^2_μ does not have the finite model property is the following. Recall from Lemma 1.9 that there is a formula of L^2_μ that defines bisimulation equivalence as a binary relation $x_1 \sim x_2$.

Example 3.3 The formula $\chi(x_1, x_2) = \neg\mu_X\big(\Diamond_1 (x_1 \sim x_2 \vee Xx_1x_2)\big)$ asserts that there is no E-path of length greater than 0 leading from the $\sim$-class $[x_1]$ into $[x_2]$. Let $\varphi(x)$ (in standard monadic semantics) be the conjunction of the universal closure in $\langle x\rangle^E$ of $x_1 \sim x_2 \rightarrow \chi(x_1, x_2)$, and the L_μ-formula $\neg\mu_X\Box Xx$ which asserts that E^{-1} is not well-founded at x (i.e. there is an infinite E-path starting at x). We claim that $\varphi \in sat(\mathrm{L}^2_\mu) \setminus \mathit{fin}\text{-}sat(\mathrm{L}^2_\mu)$. If $(\mathfrak{A}, a) \models \varphi$, then in $\mathrm{can}(\mathfrak{A}, a)$ there cannot be loops with respect to $E^\sim$, by the first conjunct. The second conjunct then forces an infinite $E^\sim$-path from a which cannot loop back, so A has to be infinite. To see that φ is satisfiable, consider $(\mathbb{N}, \mathrm{succ}, P, 0)$, with the standard successor relation for accessibility, distinguished world 0, and basic proposition P chosen such that no two vertices are bisimulation equivalent: e.g. let $P = \{2^n \mid n \in \mathbb{N}\}$. Then $(\mathbb{N}, \mathrm{succ}, P, 0) \models \varphi$.

3.1 Classical domino problems and reductions to L^2_μ

Definition 3.4 (i) A *domino* $\mathcal{D}$ is a triple (D, R_H, R_V) where D is a finite set of (types of) domino pieces, and $R_H, R_V \subseteq D \times D$ are binary relations for admissible horizontal or vertical adjacency of pieces.

(ii) A *tiling* of $\mathbb{N} \times \mathbb{N}$ by $\mathcal{D}$ is a mapping $t \colon \mathbb{N} \times \mathbb{N} \to D$ such that for all $n, m \in \mathbb{N}$: $\big(t(n,m), t(n+1,m)\big) \in R_H$ and $\big(t(n,m), t(n,m+1)\big) \in R_V$.

(iii) A tiling t is *periodic* if there are $p, q \geqslant 1$ such that $t(n+p,m) = t(n,m)$ for all m and all sufficiently large n, and similarly $t(n,m+q) = t(n,m)$ for all n and all sufficiently large m.

(iv) A tiling t is *recurrent for* $d_0 \in D$ if $t^{-1}(d_0) \subseteq \mathbb{N} \times \mathbb{N}$ is infinite.

Dominoes are a classical route to undecidability proofs through reduction, see [9].

Theorem 3.5 (Berger [8], Gurevich-Koryakov [17]) *The class of dominoes that admit a periodic tiling is recursively inseparable from the class of dominoes that admit no tiling at all. In particular both the tiling problem and the periodic tiling problem are undecidable, in fact the tiling problem is co-r.e.-complete and the periodic tiling problem is r.e.-complete.*

Theorem 3.6 (Harel [19, 20]) *The class of dominoes that admit a recurrent tiling with respect to a designated piece d_0 is Σ^1_1-complete.*

Lemma 3.7 *There are recursive translations from domino systems $\mathcal{D}$ (with distinguished piece d_0) to formulae $\psi_{\mathcal{D}}$ and $\psi_{\mathcal{D},d_0}$ of L^2_μ such that*

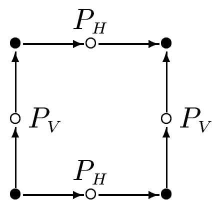
(i) $\mathcal{D}$ admits a tiling if and only if $\psi_{\mathcal{D}} \in \mathrm{sat}(\mathrm{L}^2_\mu)$.

(ii) $\mathcal{D}$ admits a periodic tiling if and only if $\psi_{\mathcal{D}} \in$ fin-sat(L^2_μ).

(iii) $\mathcal{D}, d_0$ admits a recurrent tiling if and only if $\psi_{\mathcal{D},d_0} \in \mathrm{sat}(\mathrm{L}^2_\mu)$.

This lemma proves Theorem 3.1. Indeed by virtue of the Gurevich-Koryakov Theorem, (i) and (ii) show *fin-sat*(L^2_μ) to be recursively inseparable from the complement of *sat*(L^2_μ), so that in the terminology of the classical decision problem, L^2_μ is a conservative reduction class, see [9].

For the proof of the lemma we consider descriptions of valid tilings for $\mathcal{D}$ as Kripke structures. Think of the underlying grid $\mathbb{N} \times \mathbb{N}$ as a relational structure with horizontal and vertical successor relations H and V. A placement of domino pieces is encoded by unary predicates (basic propositions) P_d for $d \in D$, so that compatibility of the tiling with the allowed horizontal and vertical adjacencies may be asserted in modal conditions along H- and V-edges. We want to use just one accessibility relation E, however, and therefore encode both H- and V-edges by means of E. Let to this end P_H and P_V be two extra unary predicates and think of H-edges as split into two consecutive E-edges with an extra world in which P_H is true put in the middle, and similarly for V and P_V.



Let $\mathcal{D} = (D, R_H, R_V)$ be a domino, $t\colon \mathbb{N} \times \mathbb{N} \to D$ a tiling. With this tiling we associate the following Kripke structure $(\mathfrak{A}_t, o)$ (with distinguished world) over basic propositions P_d for $d \in D$ and P_H, P_V.

Let $\mathbb{N} + 1/2$ stand for the set of odd multiples of $1/2$. The universe of $\mathfrak{A}_t$ is

$$\mathbb{N} \times \mathbb{N} \quad \dot{\cup} \quad (\mathbb{N} + 1/2) \times \mathbb{N} \quad \dot{\cup} \quad \mathbb{N} \times (\mathbb{N} + 1/2)$$

The distinguished world is $o := (0, 0)$. P_H and P_V mark the in-between points on horizontal and vertical edges:

$$\begin{array}{rcl} P_H & = & (\mathbb{N} + 1/2) \times \mathbb{N} \\ P_V & = & \mathbb{N} \times (\mathbb{N} + 1/2) \\ E & = & \big\{ \big((r,s),(r+1/2,s)\big) \bigm| s \in \mathbb{N}, r \in \mathbb{N} \cup (\mathbb{N} + 1/2) \big\} \\ & & \cup \big\{ \big((r,s),(r,s+1/2)\big) \bigm| r \in \mathbb{N}, s \in \mathbb{N} \cup (\mathbb{N} + 1/2) \big\}. \end{array}$$

For a description of the mapping t put, for each $d \in D$,

$$P_d \quad = \quad t^{-1}(d) \subseteq \mathbb{N} \times \mathbb{N}.$$

Observation 3.8 *A tiling t is periodic if and only if $\mathrm{can}(\mathfrak{A}_t, o)$ is finite.*

We next isolate some L^2_μ-statements that are obviously true in $(\mathfrak{A}_t, o)$ if t is a valid tiling, and of which we then show that their satisfiability in turn implies the existence of tilings.

Let Ux_i, Hx_1x_2 and Vx_1x_2 be abbreviations for the following L^2_μ-formulae:

$$\begin{array}{rcl} Ux_i & := & \neg P_H x_i \wedge \neg P_V x_i, \\ Hx_1x_2 & := & \Diamond_1\big(P_H x_1 \wedge \Diamond_1 x_1 \sim x_2\big), \\ Vx_1x_2 & := & \Diamond_1\big(P_V x_1 \wedge \Diamond_1 x_1 \sim x_2\big), \end{array}$$

where $x_1 \sim x_2$ is shorthand for the L^2_μ-formula from Lemma 1.9 that defines bisimulation equivalence (with respect to basic propositions $(P_d)_{d \in D}, P_H, P_V$) as a global relation.

We regard a unary predicate $U^\sim$ and binary relations $H^\sim$ and $V^\sim$ as interpreted over canonical structures $\mathrm{can}(\mathfrak{A}, a)$ through the above formulae.

We introduce derived modalities $\Diamond^{\mathrm{H}}$, $\Diamond^{\mathrm{V}}$ (with duals $\Box^{\mathrm{H}}$, $\Box^{\mathrm{V}}$) and their 2-dimensional indexed variants. Formally we regard these as abbreviations according to, for instance

$$\begin{array}{rcl} \Box^{\mathrm{H}}_i \varphi(x_1, x_2) & \equiv & \Box_i\big(P_H x_i \to \Box_i \varphi(x_1, x_2)\big), \\ \Diamond^{\mathrm{H}}_i \varphi(x_1, x_2) & \equiv & \Diamond_i\big(P_H x_i \wedge \Diamond_i \varphi(x_1, x_2)\big). \end{array}$$

There are L^2_μ-formulae φ_h and φ_v which over canonical models express that $H^\sim$ and $V^\sim$ are the graphs of unary functions $h, v\colon U^\sim \to U^\sim$, and which are clearly satisfied in $(\mathfrak{A}_t, o)$:

$$\begin{aligned}\varphi_h = \ & [\forall x_1]\Big(\big(Ux_1 \to (\Diamond^{\mathrm{H}}_1 Ux_1 \wedge \Box^{\mathrm{H}}_1 Ux_1)\big) \wedge (\Diamond^{\mathrm{H}}_1 Ux_1 \to Ux_1)\Big)\\ & \wedge [\forall x_1 \forall x_2]\big((Ux_1 \wedge Ux_2 \wedge x_1 \sim x_2) \to \Box^{\mathrm{H}}_1 \Box^{\mathrm{H}}_2\, x_1 \sim x_2\big),\end{aligned}$$

φ_v strictly analogous. Here and in the following $[\forall x_1]\psi(x_1)$ and $[\forall x_1 \forall x_2]\psi(x_1, x_2)$ are used as shorthand for the natural translations of these universal quantifications into L^2_μ which are sound over $\mathcal{CAN}$. For instance (cf. the proof of Proposition 2.10):

$$\begin{aligned}[\forall x_1]\psi(x_1) \ &\equiv\ \nu_X\big(\psi(x_1) \wedge \Box_1 X x_1 x_2\big)(x, x),\\ [\forall x_1 \forall x_2]\psi(x_1, x_2) \ &\equiv\ \nu_X\big(\psi(x_1, x_2) \wedge \Box_1 X x_1 x_2 \wedge \Box_2 X x_1 x_2\big)(x, x).\end{aligned}$$

The following expresses commutativity of h and v (i.e. $h \circ v = v \circ h$) over canonical models of $\varphi_h \wedge \varphi_v$, and at the same time is clearly satisfied in $(\mathfrak{A}_t, o)$:

$$\varphi_{\mathrm{com}} := [\forall x_1 \forall x_2]\big((Ux_1 \wedge Ux_2 \wedge x_1 \sim x_2) \to \Box^{\mathrm{H}}_1 \Box^{\mathrm{V}}_1 \Box^{\mathrm{V}}_2 \Box^{\mathrm{H}}_2\, x_1 \sim x_2\big).$$

Similarly, completeness of the tiling and compatibility with the adjacencies prescribed in $\mathcal{D}$ is expressed by

$$\varphi_{\mathcal{D}} := [\forall x_1]\Big(Ux_1 \to \big(\dot{\bigvee_{d \in D}} P_d x_1 \ \wedge \bigvee_{(d,d') \in R_H} (P_d x_1 \wedge \Box^{\mathrm{H}}_1 P_{d'} x_1) \ \wedge \bigvee_{(d,d') \in R_V} (P_d x_1 \wedge \Box^{\mathrm{V}}_1 P_{d'} x_1)\big)\Big).$$

We claim that the following formula of L^2_μ (in standard monadic semantics) is in $\mathit{sat}(\mathrm{L}^2_\mu)$ if and only if $\mathcal{D}$ admits a tiling, and in $\mathit{fin\text{-}sat}(\mathrm{L}^2_\mu)$ if and only if $\mathcal{D}$ admits a periodic tiling. This formula therefore serves to prove (i) and (ii) of Lemma 3.7.

$$\psi_{\mathcal{D}}(x) := Ux \wedge \ \varphi_h \wedge \varphi_v \wedge \varphi_{\mathrm{com}} \wedge \varphi_{\mathcal{D}}.$$

Now $(\mathfrak{A}_t, o) \models \psi_{\mathcal{D}}$ if t is a tiling, and $\mathrm{can}(\mathfrak{A}_t, o) \models \psi_{\mathcal{D}}$ if and only if $(\mathfrak{A}_t, o) \models \psi_{\mathcal{D}}$. Also $\mathrm{can}(\mathfrak{A}_t, o)$ is finite if t is periodic. Therefore the existence of (periodic) tilings implies the existence of (finite) models of $\psi_{\mathcal{D}}$. It remains to argue for the converse.

Assume that $(\mathfrak{B}, b) \models \psi_{\mathcal{D}}$. We may assume that $(\mathfrak{B}, b) = \mathrm{can}(\mathfrak{B}, b)$ is canonical, so that in particular $\sim$ is the identity relation on B. By $\varphi_h \wedge \varphi_v \wedge \varphi_{\mathrm{com}}$ it follows that $H^\sim$ and $V^\sim$ are the graphs of two commuting functions h and v on $U^\sim$ over $\mathfrak{B}$. Note that $b \in \mathrm{dom}(h) = \mathrm{dom}(v) = U^\sim$. Commutativity of h and v implies that for all $n, m \in \mathbb{N}$, $h(v^m h^n(b)) = v^m h^{n+1}(b)$. It follows that the standard grid on $\mathbb{N} \times \mathbb{N}$ with its successor functions is homomorphically mapped into $(U^\sim, h, v)$ through $(n, m) \mapsto v^m h^n(b)$. It is easy to check using $\psi_{\mathcal{D}}$ that putting $t(n, m)$ to be that d for which $v^m h^n(b) \in P_d^{\mathfrak{B}}$ defines a valid tiling of $\mathbb{N} \times \mathbb{N}$. Clearly t is periodic if $\mathfrak{B}$ is finite.

To settle also the Σ^1_1-hardness claim of Lemma 3.7 it remains to produce a formula $\varphi^{d_0}_{\mathrm{inf}}$ which asserts of (canonical) models $(\mathfrak{B}, b)$ of $\psi_{\mathcal{D}}$ that $v^m h^n(b) \in P_{d_0}$ for infinitely many pairs $(n, m) \in \mathbb{N}^2$. Consider the following formulae of $\mathrm{L}_\mu \subseteq \mathrm{L}^2_\mu$:

$$\begin{aligned}\varphi_1 \ &:=\ Ux \wedge \mu_X\big(P_{d_0} \vee \Diamond^{\mathrm{V}} Xx\big),\\ \varphi_2 \ &:=\ Ux \wedge \neg\mu_Y\big(\varphi_1 \vee \Diamond^{\mathrm{H}} Yx\big),\\ \varphi_3 \ &:=\ Ux \wedge \mu_Z\big(\varphi_2 \vee \Diamond^{\mathrm{H}} Zx\big).\end{aligned}$$

For $a \in U^{\sim}$ over $(\mathfrak{B}, b) = \mathrm{can}(\mathfrak{B}, b) \models \psi_{\mathcal{D}}$,

$$\begin{array}{rcl} (\mathfrak{B}, a) \models \varphi_1 & \Leftrightarrow & \langle a \rangle^{V^{\sim}} \cap P_{d_0} \neq \emptyset, \\ (\mathfrak{B}, a) \models \varphi_2 & \Leftrightarrow & \langle a \rangle^{V^{\sim} \cup H^{\sim}} \cap P_{d_0} = \emptyset, \\ (\mathfrak{B}, b) \models \varphi_3 & \Leftrightarrow & \exists a \in \langle b \rangle^{H^{\sim}} \text{ s.t. } \langle a \rangle^{V^{\sim} \cup H^{\sim}} \cap P_{d_0} = \emptyset. \end{array}$$

Thinking in terms of the standard grid $\mathbb{N} \times \mathbb{N}$ as embedded through $(n, m) \mapsto v^m h^n(b)$, $(\mathfrak{B}, b) \models \varphi_3$ says that there are only finitely many n for which $\{m \,|\, v^m h^n(b) \in P_{d_0}\}$ is non-empty. If φ_3' is built like φ_3 but with the roles of H and V exchanged, then φ_3' similarly says that there are only finitely many m for which $\{n \,|\, v^m h^n(b) \in P_{d_0}\}$ is non-empty. It follows that $\varphi_{\mathrm{inf}}^{d_0} := \neg\varphi_3 \vee \neg\varphi_3'$ is as required and that

$$\psi_{\mathcal{D}, d_0} := \psi_{\mathcal{D}} \wedge \varphi_{\mathrm{inf}}^{d_0}$$

satisfies the requirements of (iii) in Lemma 3.7.

Conclusions

The search for capturing results in the absence of order is a very central issue in finite model theory. The open question in particular whether a logical characterization can be given for the class of all PTIME queries continues to be an outstanding incentive for descriptive complexity research.

In this paper we have seen one more weak fragment of PTIME to have such a logical characterization. In terms of the full PTIME problem this capturing result concerns a weaker fragment even than those other two fragments that have yet been captured, namely the two fragments PTIME $\cap\, L^2_{\infty\omega}$ and its extension with counting quantifiers PTIME $\cap\, C^2_{\infty\omega}$ [28]. The abstract capturing result, i.e. the mere existence of a logic (or of recursive syntax) for bisimulation-invariant PTIME, moreover is an immediate consequence of the effective canonization according to Proposition 2.12.

On the other hand our capturing result here concerns a particularly natural fragment of PTIME, because bisimulation invariance and the modal scenario capture very natural logical concerns. Bisimulation invariance is the adequate notion of invariance for many applications of logic to the analysis of processes and programs. L_μ^ω is the logic which is complete for all efficient tasks in this scenario. Beyond the abstract capturing result the major point of the present result is the *explicit presentation* in terms of this very natural extension of the well-known μ-calculus.

This capturing result thereby establishes a new close link between bisimulation invariance and L_μ, a logic that has been studied in its own right for many other reasons. To mention one of the more recent results in the study of L_μ, Janin and Walukiewicz [22] show that (over general Kripke structures) L_μ exactly corresponds to the bisimulation-invariant fragment of monadic second-order logic. Even though it is not L_μ itself but rather its *vectorizations* L_μ^k that come up here, the present capturing result illustrates the naturalness of L_μ from yet another angle.

The strong undecidability result for L_μ^ω shows that L_μ is much more expressive than L_μ even in terms of the standard monadic semantics. This is even more noteworthy as the effect of vectorization is rather trivial for both ML and ML_∞ ($\mathrm{ML} \equiv \mathrm{ML}^k$ and $\mathrm{ML}_\infty^k \equiv \mathrm{ML}_\infty$ for the standard monadic semantics). Vectorization at the intermediate level of L_μ has strong effects in terms of decidability, and leads to a logic that is complete for PTIME within ML_∞.

Acknowledgement I am grateful to Jörg Flum for suggesting a conceptual improvement in the proof of the main theorem.

References

1. S. Abiteboul and V. Vianu, *Generic computation and its complexity*, Proc. 23rd ACM Symp. on Theory of Computing, 1991, pp. 209–219.
2. J. Balcazar, J. Gabarro and M. Santha, *Deciding bisimilarity is P-complete*, Formal Aspects of Computing, 4, 1992, pp. 638–648.
3. J. Barwise and J. van Benthem, *Interpolation, preservation, and pebble games*, to appear in Journal of Symbolic Logic, 1998.
4. J. Barwise and L. Moss, *Vicious Circles*, CSLI Lecture Notes Series, vol. 60, 1996.
5. ——, *Modal correspondence for models*, to appear in Journal of Philosophical Logic, 1998.
6. J. van Benthem, *Modal correspondence theory*, PhD thesis, University of Amsterdam, 1976.
7. ——, *Modal Logic and Classical Logic*, Bibliopolis, 1985.
8. R. Berger, *The undecidability of the domino problem*, Memors of the AMS, 66, 1966.
9. E. Börger, E. Grädel, and Y. Gurevich, *The Classical Decision Problem*, Perspectives in Mathematical Logic, Springer, 1997.
10. A. Chandra and D. Harel, *Structure and complexity of relational queries*, Journal of Computer and System Sciences, 25, 1982, pp. 99–128.
11. A. Dawar, S. Lindell, and S. Weinstein, *Infinitary logic and inductive definability over finite structures*, Information and Computation, 119, 1995, pp. 160–175.
12. H.-D. Ebbinghaus and J. Flum, *Finite Model Theory*, Perspectives in Mathematical Logic, Springer, 1995.
13. E. Emerson, *Temporal and modal logic*, in J. van Leeuwen (ed.), *Handbook of Theoretical Computer Science*, vol. B: Formal Models and Semantics, Elsevier, 1990, pp. 995–1072.
14. E. Grädel and M. Otto, *Inductive definability with counting on finite structures*, In E. Börger et al., editors, Selected Papers of CSL '92, Lecture Notes in Computer Science 702, Springer Verlag 1993, pp. 231–247.
15. E. Grädel, M. Otto, and E. Rosen, *Undecidability results on logics with two variables*, to appear in Archive for Mathematical Logic, 1998. See also: Proc. of 14th Symposium on Theoretical Aspects of Computer Science, 1997, Lecture Notes in Computer Science, vol. 1200, Springer, 1997, pp. 249–260.
16. Y. Gurevich, *Logic and the challenge of computer science*, in E. Börger (ed.), *Current Trends in Theoretical Computer Science*, Computer Science Press, 1988, pp. 1–57.
17. Y. Gurevich and I. Koryakov, *Remarks on Berger's paper on the domino problem*, Siberian Math. Journal, 13, 1972, pp. 319–321.
18. Y. Gurevich and S. Shelah, *Fixed-point extensions of first-order logic*, Annals of Pure and Applied Logic, 32, 1986, pp. 265–280.
19. D. Harel, *Recurring dominoes: Making the highly undecidable highly understandable*, Annals of Discrete Mathematics, 24, 1985, pp. 51–72.

20. ———, *Effective transformations on infinite trees, with applications to high undecidability, dominoes and fairness*, Journal of the ACM, 33, 1986, pp. 224–248.

21. N. Immerman, *Relational queries computable in polynomial time*, Information and Control, 68, 1986, pp. 86–104.

22. D. Janin and I. Walukiewicz, *On the expressive completeness of the propositional mu-calculus with respect to monadic second-order logic*, in: *Proc. 7th Int. Conf. on Concurrency Theory, Concur '96*, Lecture Notes in Computer Science, vol. 1119, Springer, 1996, pp. 263–277.

23. D. Kozen, *Results on the propositional μ-calculus*, Theoretical Computer Science, 27, 1983, pp. 333–354.

24. D. Kozen, *A finite model theorem for the propositional μ-calculus*, Studia Logica, 47, 1987, pp. 233–241.

25. D. Kozen and R. Parikh, *A decision procedure for the propositional μ-calculus*, in: E. Clarke and D. Kozen (eds.), *Logics of Programs*, Lecture Notes in Computer Science, vol. 164, Springer, 1984, pp. 313–325.

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

27. M. Otto, *The expressive power of fixed-point logic with counting*, Journal of Symbolic Logic, 61, 1996, pp. 147–176.

28. ———, *Canonization for two variables and puzzles on the square*, Annals of Pure and Applied Logic, 85, 1997, pp. 243–283.

29. ———, *Bounded Variable Logics and Counting — A Study in Finite Models*, Lecture Notes in Logic, vol. 9, Springer, 1997.

30. ———, *Bounded variable logics: two, three, and more*, to appear in Archive for Mathematical Logic, 1998.

31. D. Park, *Concurrency and automata on infinite sequences*, in: P. Deussen (ed.) *Proc. 5th GI-Conf. on Theoretical Computer Science*, Lecture Notes in Computer Science, vol. 104, Springer, 1981, pp. 167–183.

32. V. Sazonov and A. Lisitsa, *On linear oerdering of strongly extensional finitely-branching graphs and non-well-founded sets*, in: *Proc. 4th Symp. on Logicaal Foundations of Computer Science*, Lecture Notes in Computer Science, vol. 1234, Springer, 1997, pp. 178–188.

33. R. Streett and E. Emerson, *An automata theoretic decision procedure for the propositional μ-calculus*, Information and Computation, 81, 1989, pp. 249–264.

34. M. Vardi, *The complexity of relational query languages*, in Proc. 14th ACM Symp. on Theory of Computing, 1982, pp. 137–146.