Monadic chain logic over iterations and applications to pushdown systems

Dietrich Kuske
Institut für Informatik, Universität Leipzig, Germany
kuske@informatik.uni-leipzig.de

Markus Lohrey
Universität Stuttgart, FMI, Germany
lohrey@informatik.uni-stuttgart.de

Abstract

Logical properties of iterations of relational structures are studied and these decidability results are applied to the model checking of a powerful extension of pushdown systems. It is shown that the monadic chain theory of the iteration of a structure $\mathcal{A}$ (in the sense of Shelah and Stupp) is decidable in case the first-order theory of the structure $\mathcal{A}$ is decidable. This result fails if Muchnik's clone-predicate is added. A model of pushdown automata, where the stack alphabet is given by an arbitrary (possibly infinite) relational structure, is introduced. If the stack structure has a decidable first-order theory with regular reachability predicates, then the same holds for the configuration graph of this pushdown automaton. This result follows from our decidability result for the monadic chain theory of the iteration.

1. Introduction

In this paper, we study iterations of relational structures, their logical properties, and apply our results to the model checking of a powerful extension of pushdown systems.

The *local full iteration* $\mathcal{A}^*_{\mathrm{loc}}$ of a relational base structure $\mathcal{A}$ with universe A consists of the set A^* of finite words over A. One of its relations is the immediate successor relation son. The sons of a word w carry the relations of the base structure $\mathcal{A}$. Furthermore, Muchnik's unary clone predicate collects all words whose final two letters are identical. Semenov [20] sketched a proof of what is now known as Muchnik's preservation theorem: The monadic second order (MSO for short) theory of the local full iteration $\mathcal{A}^*_{\mathrm{loc}}$ can be reduced to the MSO-theory of the base structure $\mathcal{A}$, and, if two base structures have the same MSO-theory, then the same holds for their iterations. Hence, if the MSO-theory of a structure $\mathcal{A}$ is decidable, then also the MSO-theory of the local full iteration $\mathcal{A}^*_{\mathrm{loc}}$ is decidable. A full proof of this result was given by Walukiewicz in [25]. A first-order variant of Muchnik's theorem for first-order logic follows from [14]. For modulo counting extensions of MSO and for guarded second order logic, a preservation theorem was shown by Blumensath and Kreutzer [3].

The *full iteration* $\mathcal{A}^*_{\mathrm{fu}}$ differs from the local full iteration only in as far as it contains the prefix relation on A^* instead of the immediate successor relation son. Since this prefix relation is the transitive closure of son, there is an MSO-interpretation of the full iteration in the local full iteration. As an immediate consequence from [25, 3], one obtains a preservation theorem for MSO and its modulo counting extensions for this full iteration in place of the local full iteration. Since one can express in first-order logic that an element of the full iteration (i.e., a word over the base structure) represents a path in the base structure, both parts of the preservation theorem fail for the full iteration and first-order logic (Propositions 3.4 and 3.5).

To overcome this problem, Section 4 is devoted to the study of the *basic iteration* $\mathcal{A}^*_{\mathrm{ba}}$ where one omits Muchnik's clone predicate but keeps the prefix order. For basic iterations, the preservation theorem for MSO was proved by Stupp [22] (cf. [21]). Rabin's seminal result on the decidability of the MSO-theory of the complete infinite binary tree [18] is an immediate corollary of this preservation theorem. For this basic iteration we are able to prove the preservation theorem for first-order logic. In fact, we can show even more: If a structure has a decidable first-order theory, then its basic iteration has a decidable $\mathrm{MSO}^{\mathrm{ch}}$-theory (Thm. 4.10). $\mathrm{MSO}^{\mathrm{ch}}$ is the fragment of MSO where second-order quantification is restricted to chains (i.e., ordered subsets) with respect to the tree structure of the iteration. $\mathrm{MSO}^{\mathrm{ch}}$ on trees was investigated in [23]. To reduce the $\mathrm{MSO}^{\mathrm{ch}}$-theory of the basic iteration to the first-order theory of the base structure, we proceed as follows: First, we show that quantification over chains can be restricted to ultimately periodic chains of bounded offset and period length (Thm. 4.7). Truth of $\mathrm{MSO}^{\mathrm{ch}}$-formulas with bounded quantification can be determined in a bounded prefix of the basic iteration. Finally, this bounded prefix can be interpreted in the base structure. Since all these bounds can be computed effectively, our preservation theorem follows.

Roughly speaking, the results from Section 3 and Section 4 show that, in order to have a first-order preservation theorem for the iteration, we are not allowed to copy an infinite amount of information between the levels of the tree structure — this is in some sense the essence of the clone-predicate. Thus, the clone-predicate has an immense effect on the expressive power of the basic iteration although it looks quite innocent at first glance. It should be also noted that the clone-predicate allows to define the unravelling of a graph G within the full iteration of G (cf. [6]).

In Section 5 we present an application of our decidability result for $\mathrm{MSO}^{\mathrm{ch}}$ over basic iterations to pushdown systems. Pushdown systems were used to model the state space of sequential programs with nested procedure calls, see e.g. [9]. Model-checking problems for pushdown systems were studied for various temporal logics (LTL, CTL, modal μ-calculus) [1, 9, 13, 24]. When modeling recursive sequential programs via pushdown systems, it is necessary to abstract local variables (which have to be stored on the stack) with an infinite range (like for instance integers) to some finite range, in order to obtain a finite pushdown alphabet. This abstraction may lead to so called spurious counterexamples [8]. Here, we introduce pushdown systems where the stack alphabet is the (possibly infinite) universe of an arbitrary stack structure $\mathcal{A}$. With any change of the control state, our pushdown model associates one of three basic operations: (i) replacing the topmost symbol of the stack by another one according to some binary predicate of the stack structure, (ii) pushing or (iii) popping a symbol from some unary predicate of the stack structure. Such a pushdown system can model programs with nested procedure calls, where procedures use variables with an infinite domain. The configuration graph of such a pushdown system is defined as for finite stack alphabets. We study the logic FOREG for these configuration graphs. FOREG is the extension of first-order logic which allows to define new binary predicates by regular expressions over the binary predicates of the base structure $\mathcal{A}$. Variants of FOREG were studied in [15, 19, 26]. FOREG is a suitable language for the specification of reachability properties of reactive systems; its expressive power is between first-order logic and MSO. Based on our decidability result Thm. 4.10 we show that if FOREG is decidable for the base structure $\mathcal{A}$ of a pushdown system, then FOREG remains decidable for the configuration graph of the pushdown system (Thm. 5.1). For this result, it is important that in our pushdown model procedure calls and returns cannot transfer an infinite amount of information to another call level. This reflects our undecidability result Proposition 3.4 for the clone predicate.

2. Preliminaries

Let Σ be a (not necessarily finite) alphabet. With Σ^+ we denote the set of all finite non-empty words over Σ. Then $\Sigma^* = \Sigma^+ \cup \{\varepsilon\}$ with ε the empty word. With $\preceq$ we denote the prefix relation on finite words and $\prec$ is its non-reflexive part. For a subalphabet $\Gamma \subseteq \Sigma$ and a word $u \in \Sigma^*$ we denote with $|u|_\Gamma$ the number of occurrences of symbols from Γ in u. In case Σ is finite, $\mathrm{REG}(\Sigma)$ denotes the set of all regular languages over the alphabet Σ.

2.1. Iterations

Let $\mathcal{A} = (A, (R^{\mathcal{A}})_{R\in\sigma})$ be a relational structure over the finite relational signature σ. The *basic iteration* $\mathcal{A}^*_{\mathrm{ba}}$ *of* $\mathcal{A}$ is the structure

$$\mathcal{A}^*_{\mathrm{ba}} = (A^*, \preceq, (\widehat{R})_{R\in\sigma}, \varepsilon) \text{ where}$$
$$\widehat{R} = \{(ua_1, \ldots, ua_n) \mid u \in A^*, (a_1, \ldots, a_n) \in R^{\mathcal{A}}\}.$$

Example 2.1 *Suppose the structure $\mathcal{A}$ has two elements a and b and two unary relations $R_1 = \{a\}$ and $R_2 = \{b\}$. Then $\widehat{R}_1 = \{a, b\}^*a$ and $\widehat{R}_2 = \{a, b\}^*b$. Hence the basic iteration $\mathcal{A}^*_{\mathrm{ba}}$ can be visualized as a complete binary tree with unary predicates telling whether the current node is the first or the second son of its father. In addition, the root ε is a constant of the structure $\mathcal{A}^*_{\mathrm{ba}}$.*

In the *full iteration* $\mathcal{A}^*_{\mathrm{fu}}$ of $\mathcal{A}$, we have the additional unary *clone predicate* $\mathrm{cl} = \{uaa \mid u \in A^*, a \in A\}$, i.e.,

$$\mathcal{A}^*_{\mathrm{fu}} = (A^*, \preceq, \mathrm{cl}, (\widehat{R})_{R\in\sigma}, \varepsilon) \ .$$

We will also consider a relaxation of the full iteration where the prefix relation is replaced by the direct successor relation $\mathrm{son} = \{(u, ua) \mid a \in A^*, a \in A\}$, i.e.,

$$\mathcal{A}^*_{\mathrm{loc}} = (A^*, \mathrm{son}, \mathrm{cl}, (\widehat{R})_{R\in\sigma}, \varepsilon) \ .$$

We refer to this iteration as *local iteration*. Note that $\mathcal{A}^*_{\mathrm{fu}}$ is MSO-definable (but not first-order definable) in $\mathcal{A}^*_{\mathrm{loc}}$.

2.2. Logics

Let σ be some signature. *Atomic formulas* are $R(x_1, \ldots, x_n)$, $x_1 = x_2$, and $x_1 \in X$ where $x_1, \ldots, x_n$ are *individual variables*, $R \in \sigma$ is an n-are relational symbol, and X is a *set variable*. *Monadic second-order formulas* are obtained from atomic formulas by conjunction, negation, and quantification $\exists x$ and $\exists X$ for x an individual and X a set variable. The satisfaction relation $(\mathcal{A}, \bar{a}, \bar{C}) \models \varphi(\bar{x}, \bar{X})$ is defined as usual with the understanding that set variables range over subsets of A. A *first-order formula* is a monadic second-order formula without set variables.

Now let $\preceq$ be a designated binary relation symbol in σ. A *monadic second-order chain formula* or $\mathrm{MSO}^{\mathrm{ch}}$-formula is just a monadic second-order formula. For these $\mathrm{MSO}^{\mathrm{ch}}$-formulas, we define a new satisfaction relation $\models^{\mathrm{ch}}$: it is defined as $\models$ with the only difference that set variables range over chains (i.e., sets whose elements are mutually comparable) in $(A, \preceq)$. Note that if φ is a first-order formula, then $\mathcal{A} \models \varphi$ if and only if $\mathcal{A} \models^{\mathrm{ch}} \varphi$.

Let $\mathcal{A}$ and $\mathcal{B}$ be two σ-structures. Then we write $\mathcal{A} \equiv_m^{\mathrm{MSO}} \mathcal{B}$ if, for any MSO-formula φ of quantifier depth at most m, we have $\mathcal{A} \models \varphi \iff \mathcal{B} \models \varphi$. This relation is an equivalence relation. If we only consider first-order formulas ($\mathrm{MSO}^{\mathrm{ch}}$-formulas, resp.) φ of quantifier depth at most m, then we write $\mathcal{A} \equiv_m^{\mathrm{FO}} \mathcal{B}$ ($\mathcal{A} \equiv_m^{\mathrm{ch}} \mathcal{B}$, resp.).

3. The theory of the full iteration

We first deal with MSO-theories. Muchnik's theorem sharpens an earlier result of Stupp. Its full proof can be found in [25] (cf. also [2]).[1]

Theorem 3.1 *Let σ be some finite relational signature. There exists a computable function* $\mathrm{red} : \mathrm{MSO}(\sigma, \mathrm{cl}, \mathrm{son}) \to \mathrm{MSO}(\sigma)$ *such that, for any σ-structure $\mathcal{A}$, we have $\mathcal{A} \models \mathrm{red}(\varphi)$ if and only if $\mathcal{A}^*_{\mathrm{loc}} \models \varphi$.*

One infers immediately:

Corollary 3.2 *If the MSO-theory of a structure $\mathcal{A}$ is decidable, then the MSO-theory of its local iteration $\mathcal{A}^*_{\mathrm{loc}}$ is decidable as well.*

To derive another corollary, let $m \in \mathbb{N}$ be arbitrary. Then, there is a finite set Φ of MSO-formulas such that any MSO-sentence of quantifier depth at most m is logically equivalent to some sentence from Φ. Let n be an upper bound for the quantifier depth of $\mathrm{red}(\varphi)$ for $\varphi \in \Phi$. This observation yields:

Corollary 3.3 *For any $m \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that, for any two σ-structures $\mathcal{A}$ and $\mathcal{B}$ with $\mathcal{A} \equiv_n^{\mathrm{MSO}} \mathcal{B}$, we have $\mathcal{A}^*_{\mathrm{loc}} \equiv_m^{\mathrm{MSO}} \mathcal{B}^*_{\mathrm{loc}}$.*

Note that the MSO-theories of the local and the full iteration can be reduced onto each other. Hence Muchnik's Thm. 3.1 and Corollaries 3.2 and 3.3 hold for the full iteration $\mathcal{A}^*_{\mathrm{fu}}$ in place of the local iteration $\mathcal{A}^*_{\mathrm{loc}}$ equally well. Surprisingly, this is not the case for first-order logic as we show next.

Proposition 3.4 *There exists a structure $\mathcal{A}$ with a decidable first-order theory such that the full iteration $\mathcal{A}^*_{\mathrm{fu}}$ has an undecidable first-order theory.*

Proof. Let $\mathcal{M}$ be a Turing machine that accepts a non-recursive set L (we assume that $\mathcal{M}$ accepts with empty tape). Let Σ be the set of tape symbols and states of $\mathcal{M}$. Then consider the following structure $\mathcal{A} = (A, E, (E_a)_{a \in \Sigma})$ where A is the set of configurations of $\mathcal{M}$ and, for any configurations c_1, c_2 and any $a \in \Sigma$, we have

- $(c_1, c_2) \in E$ if and only if c_2 can be obtained from c_1 by one step of the Turing machine.
- $(c_1, c_2) \in E_a$ if and only if $c_2 = c_1 a$.

The first-order theory of $\mathcal{A}$ is decidable since $\mathcal{A}$ is automatic [12]. There is a formula α with one free variable x such that $(\mathcal{A}, c) \models \alpha$ if and only if c is a configuration with empty tape. Furthermore, from a state q and an input word w, we can write a first-order formula φ_{qw} with one free variable x such that, for any configuration c, $(\mathcal{A}, c) \models \varphi_{qw}$ if and only if $c = qw$.

Now consider the full iteration of $\mathcal{A}$. The formulas $\widehat{\alpha}$ and $\widehat{\varphi_{qw}}$ are obtained by restricting the quantification to siblings of the free variable x. Furthermore, let w be some input word and let q_0 be the initial state of $\mathcal{M}$. Then w is accepted if and only if there exists a sequence of configurations $u = c_0 c_1 \ldots c_n \in A^*$ such that the following hold in the full iteration of $\mathcal{A}$: (i) the minimal nonempty prefix c_0 of u satisfies $\widehat{\varphi_{q_0 w}}$, (ii) u satisfies $\widehat{\alpha}$, and (iii) for all proper and non-empty prefixes v of u, we have

$$\exists v', v'' : v \lessdot v' \preceq u \land v \lessdot v'' \land \mathrm{cl}(v'') \land \widehat{E}(v'', v'),$$

where $x \lessdot y$ is shorthand for $x \prec y \land \forall z (x \preceq z \prec y \to x = z)$. Since the language of the Turing machine $\mathcal{M}$ is non-recursive, this proves that the first-order theory of the full iteration of $\mathcal{A}$ is undecidable. $\square$

Hence, Corollary 3.2 and therefore Thm. 3.1 with the full iteration taking the place of the local iteration and first-order logic replacing MSO do not hold. A similar problem arises with respect to Corollary 3.3.

Proposition 3.5 *For every $n \in \mathbb{N}$ there exist structures $\mathcal{A}_n$ and $\mathcal{B}_n$ such that $\mathcal{A}_n \equiv_n^{\mathrm{FO}} \mathcal{B}_n$ but $(\mathcal{A}_n)^*_{\mathrm{fu}} \not\equiv_6^{\mathrm{FO}} (\mathcal{B}_n)^*_{\mathrm{fu}}$.*

Proof. For $n \in \mathbb{N}$, let $\mathcal{A}_n$ denote the structure $\mathcal{A}_n = (\mathbb{Z}, \mathrm{succ}, 0, 2^{n+1})$ that consists of a copy of the integers with successor relation and two constants a and b. Note that in $\mathcal{A}_n$ there is a path of length 2^{n+1} from a to b. We will also consider the structure $\mathcal{B}_n = (\mathbb{Z}, \mathrm{succ}, 2^{n+1}, 0)$ that differs from $\mathcal{A}_n$ only in the values of the constants (that are exchanged). Then the structures $\mathcal{A}_n$ and $\mathcal{B}_n$ cannot be distinguished by any first-order sentence of quantifier rank at most n, i.e., $\mathcal{A}_n \equiv_n^{\mathrm{FO}} \mathcal{B}_n$.

[1] Thm. 3.1 and its two corollaries also hold for counting extensions of MSO and for guarded second-order logic [3].

Now consider the following sentence φ in the language of the full iteration of $\mathcal{A}_n$ and $\mathcal{B}_n$:

$$\exists x \in \widehat{a}\ \exists z \in \widehat{b} : x \preceq z \wedge$$
$$\forall y : x \preceq y \prec z \to \exists y' \exists y'' \left\{ \begin{array}{l} y \lessdot y' \preceq z \wedge y \lessdot y'' \wedge \\ \mathrm{cl}(y'') \wedge \widehat{E}(y'', y') \end{array} \right\}$$

To show that $\mathcal{A}_n$ satisfies φ, take $x = 0$ and $z = 0\,1\,2 \ldots 2^{n+1}$. Since the last letters of these words are a and b, resp., they belong to $\widehat{a}$ and $\widehat{b}$, resp. Any word y with $x \preceq y \prec z$ has the form $0\,1\,\ldots i$ for some $0 \leq i < 2^{n+1}$. Then $y' = y\,(i+1)$ and $y'' = y\,i$ ensure that φ indeed holds.

On the other hand, $\mathcal{B}_n^*$ does not satisfy φ: Suppose it would, i.e., there are $x = x'a$ and $z = xa_1a_2 \ldots a_k b$ satisfying the second line of the formula φ. Then $a\,a_1\,a_2\,\ldots\,a_k\,b$ is a path in $\mathcal{B}$ from a to b - but such a path does not exist. Since φ has quantifier rank 6, we obtain $(\mathcal{A}_n)^*_{\mathrm{fu}} \not\equiv^{\mathrm{FO}}_6 (\mathcal{B}_n)^*_{\mathrm{fu}}$. □

In Proposition 3.4 and 3.5, the interplay between the prefix relation and the clone-predicate is crucial. If just one of these two relations is present, a first-order version of Muchnik's theorem and its corollaries holds. For the clone-predicate, this follows from a more general result on so called factorized unfoldings from our earlier paper [14] (see Theorem 3.6 below). For the prefix relation, we prove the result in this paper (Theorem 4.10 and Corollary 4.11).

Theorem 3.6 ([14]) *Let σ be a finite relational signature.*

- *Let σ be some finite relational signature. There exists a computable function* $\mathrm{red} : \mathrm{FO}(\sigma, \mathrm{cl}, \mathrm{son}) \to \mathrm{FO}(\sigma)$ *such that, for any σ-structure $\mathcal{A}$, we have $\mathcal{A} \models \mathrm{red}(\varphi)$ if and only if $\mathcal{A}^*_{\mathrm{loc}} \models \varphi$.*
- *If the first-order theory of a structure $\mathcal{A}$ is decidable, then the first-order theory of its local iteration $\mathcal{A}^*_{\mathrm{loc}}$ is decidable as well.*
- *For any $m \in \mathbb{N}$, there exists $n \in \mathbb{N}$ such that, for any two σ-structures $\mathcal{A}$ and $\mathcal{B}$ with $\mathcal{A} \equiv^{\mathrm{FO}}_n \mathcal{B}$, we have $\mathcal{A}^*_{\mathrm{loc}} \equiv^{\mathrm{FO}}_m \mathcal{B}^*_{\mathrm{loc}}$.*

4. The $\mathrm{MSO}^{\mathrm{ch}}$-theory of the basic iteration

In this section, we will show that statements analogous to Muchnik's Thm. 3.1 and Corollaries 3.2 and 3.3 hold for basic iterations and first-order logic. In doing so, it turns out that we can even consider the $\mathrm{MSO}^{\mathrm{ch}}$-theory of the basic iteration. Let us fix a base structure $\mathcal{A} = (A, (R)_{R\in\sigma})$ over a signature σ. In the rest of Section 4, we write

$$t = \mathcal{A}^*_{\mathrm{ba}}. \qquad (1)$$

4.1. Preliminaries

For $i, \ell \in \mathbb{N}$, let $\tau_{i,\ell}$ be the extension of the signature $(\sigma, \preceq)$ by i individual and ℓ chain constants. We write τ_i for $\tau_{i,0}$. From ℓ and m, one can effectively compute a finite upper bound $N_i(\ell, m)$ for the number of equivalence classes of $\equiv^{\mathrm{ch}}_m$ on the class of all $\tau_{i,\ell}$-structures, see [10].

For $u \in A^*$, let t_u be the structure $(uA^*, \sqsubseteq, (\bar{R})_{R\in\sigma}, u)$ over the signature τ_1, where (i) the relation $\sqsubseteq$ is the restriction of $\preceq$ to uA^* and (ii) $\bar{R}$ is the restriction of $\widehat{R}$ to uA^+ (the restriction to uA^* could contain tuples of the form $(u, u, \ldots, u)$ which are excluded from $\bar{R}$). For any $u, v \in A^*$, the mapping $f : t_u \to t_v$ with $f(ux) = vx$ is an isomorphism – this is the reason to consider $\bar{R}$ and not the restriction of $\widehat{R}$ to uA^*. Similarly, the τ_2-structure $t_{u,v} = (uA^* \setminus vA^+, \sqsubseteq, (\bar{R})_{R\in\sigma}, u, v)$ is defined for $u, v \in A^*$ with $u \preceq v$. Here, again, $\bar{R}$ is the restriction of $\widehat{R}$ to $uA^+ \setminus vA^+$.

Example 2.1 (continued). In the case of Example 2.1, t_u is just the subtree rooted at the node u. On the other hand, $t_{u,v}$ is obtained from t_u by deleting all descendents of v and marking the node v as a constant. Thus, we can think of $t_{u,v}$ as a tree with a marked leaf. These *special trees* are fundamental in the work of Gurevich and Shelah [11] and in Thomas' study of the monadic second-order chain theory of the complete binary tree [23]. The following constructions generalize those from [11, 23] to the more general context of basic iterations as considered here.

In the following, fix some $\ell \in \mathbb{N}$. We then define the operations of product and infinite product of $\tau_{i,\ell}$-structures: If $\mathcal{A}$ is a $\tau_{2,\ell}$-structure with second individual constant v and $\mathcal{B}$ a disjoint $\tau_{i,\ell}$-structure with first individual constant u, then their *product* $\mathcal{A} \cdot \mathcal{B}$ is a $\tau_{i,\ell}$-structure. It is obtained from the union of $\mathcal{A}$ and $\mathcal{B}$ by identifying v and u and erasing it from the list of constants. In other words, the individual constants in $\mathcal{A} \cdot \mathcal{B}$ are the first constant from $\mathcal{A}$ and all but the first constant from $\mathcal{B}$. Furthermore, the chains from $\mathcal{A}$ and $\mathcal{B}$ are united. Now let $\mathcal{A}_n$ be disjoint $\tau_{2,\ell}$-structures with individual constants u_n and v_n for $n \in \mathbb{N}$. Then the *infinite product* $\prod_{n\in\mathbb{N}} \mathcal{A}_n$ is a $\tau_{1,\ell}$-structure. It is obtained from the union of the structures $\mathcal{A}_n$ by identifying v_n and u_{n+1} for any $n \in \mathbb{N}$. The only individual constant of this infinite product is u_0. If $\mathcal{A}_n \cong \mathcal{A}_{n+1}$ for all $n \in \mathbb{N}$, then we write simply $\mathcal{A}_0^\omega$ for the infinite product of the structures $\mathcal{A}_n$. Standard applications of Ehrenfeucht-Fraïssé-games (cf. [7]) yield:

Proposition 4.1 *Let $m, \ell \in \mathbb{N}$, $\mathcal{A}_n, \mathcal{A}'_n$ be $\tau_{2,\ell}$-structures for $n \in \mathbb{N}$ and let $\mathcal{B}, \mathcal{B}'$ be some $\tau_{i,\ell}$-structures such that $\mathcal{A}_n \equiv^{\mathrm{ch}}_m \mathcal{A}'_n$ for $n \in \mathbb{N}$ and $\mathcal{B} \equiv^{\mathrm{ch}}_m \mathcal{B}'$. Then*

$$\mathcal{A}_0 \cdot \mathcal{B} \equiv^{\mathrm{ch}}_m \mathcal{A}'_0 \cdot \mathcal{B}' \quad \text{and} \quad \prod_{n\in\mathbb{N}} \mathcal{A}_n \equiv^{\mathrm{ch}}_m \prod_{n\in\mathbb{N}} \mathcal{A}'_n.$$

4.2. Ultimately periodic chains and their combinatorics

For a word $u \in A^\infty = A^* \cup A^\omega$, let $\downarrow u \subseteq A^*$ denote the set of finite prefixes of u, and $\Downarrow u = \downarrow u \setminus \{u\}$. Similarly, $\downarrow C = \bigcup\{\downarrow u \mid u \in C\}$ for $C \subseteq A^*$. Finally, $u^{-1}C = \{v \in A^* \mid uv \in C\}$ for $u \in A^*$ and $C \subseteq A^*$.

A chain $C \subseteq A^*$ is *ultimately periodic* if it can be written as $E \cup uv^*F$ with $E, F \subseteq A^*$ finite and $u, v \in A^*$. If $E \subseteq \Downarrow u$ and $F \subseteq \Downarrow v$, it is *ultimately $|v|$-periodic with offset* $|u|$. Since $v = \varepsilon$ is possible, finite chains are ultimately 0-periodic. Furthermore, if C is ultimately p-periodic with offset q, then it is also ultimately xp-periodic with offset $q+y$ for any $x, y \in \mathbb{N}$ with $x \geq 1$.

In the following, we will consider the structure t from (1) together with $\ell + 1$ chains $C_1, \ldots, C_\ell, C$. To make the presentation more concise, write $\bar{C}$ for the ℓ-tuple $(C_1, \ldots, C_\ell)$. We will also meet structures t_u and $t_{u,v}$ together with the restriction of $\bar{C}, C$ to their domain. Again for simplicity, we write, e.g., (t_u, C) for $(t_u, C \cap uA^*)$.

4.3. Shortening ultimately periodic chains

Suppose we are in the realm of Example 2.1 and let $C_i \subseteq A^*$ be regular and let $u_i \in A^*$. Then, as a corollary from Rabin's tree theorem, for any $C \subseteq A^*$, there exists a *regular* set $D \subseteq A^*$ that satisfies the same MSO-formulas of quantifier depth m in the structure $(t, C_1, \ldots, C_\ell, u_1, \ldots, u_n)$ as C does. In this section, we want to prove a similar result for basic iterations. For this, "regular set" is replaced by "ultimately periodic chain". In addition, we want to bound the offset and the period of the chain D.

We start showing that some ultimately periodic chain D exists that can take the role of C (Proposition 4.2). Proposition 4.6 will allow to bound the period of D (thereby possibly enlarging the offset). Finally, Lemma 4.3 bounds the size of the offset (without changing the period). Finally, Thm. 4.7 shows that we succeeded in our attempt to find an equivalent ultimately periodic chain D of small period and offset.

4.3.1. Existence of ultimately periodic chains

Proposition 4.2 *Let $m \in \mathbb{N}$, $C_1, \ldots, C_\ell \subseteq A^*$ be ultimately periodic chains and let $C \subseteq A^*$ be any chain. Then there exists an ultimately periodic chain D such that $(t, \bar{C}, C) \equiv_m^{\mathrm{ch}} (t, \bar{C}, D)$.*

Proof. Assume C not to be ultimately periodic (and therefore infinite) and let $\alpha \in A^\omega$ with $C \subseteq \downarrow\alpha$. By Ramsey's theorem (see [16] for this application), there is a strictly increasing sequence $u_1 \prec u_2 \prec u_3 \cdots$ of non-empty prefixes of α such that,

(a) $|u_1|$ exceeds the offset of all the chains $C_1, C_2, \ldots, C_\ell$,

(b) for any $1 \leq i \leq \ell$ and for any $n \geq 1$, the period length of C_i divides $|u_{n+1}| - |u_n|$, and

(c) for any $1 \leq i < j$, we have $(t_{u_1,u_2}, \bar{C}, C) \equiv_m^{\mathrm{ch}} (t_{u_i,u_j}, \bar{C}, C)$.

This implies

$$\begin{aligned}(t, \bar{C}, C) &= (t_{\varepsilon,u_1}, \bar{C}, C) \cdot \prod_{n>0} (t_{u_n,u_{n+1}}, \bar{C}, C)\\ &\equiv_m^{\mathrm{ch}} (t_{\varepsilon,u_1}, \bar{C}, C) \cdot (t_{u_1,u_2}, \bar{C}, C)^\omega\end{aligned}$$

Now let $v \in A^+$ with $u_1 v = u_2$ and consider $E = C \cap \downarrow u_1$, $F = u_1^{-1}(C \cap \downarrow u_2) = u_1^{-1}C \cap \downarrow v$, and $D = E \cup u_1 v^* F$. Because of (a) and (b) we can continue as follows:

$$\begin{aligned}&= (t_{\varepsilon,u_1}, \bar{C}, E) \cdot (t_{u_1,u_2}, \bar{C}, F)^\omega\\ &\cong (t_{\varepsilon,u_1}, \bar{C}, D) \cdot \prod_{n\geq 0} (t_{u_1 v^n, u_1 v^{n+1}}, \bar{C}, D)\\ &= (t, \bar{C}, D).\end{aligned}$$

Since $E \subseteq \downarrow u_1 v^\omega$ and $F \subseteq \downarrow v^\omega$, the set D is linearly ordered and therefore ultimately periodic. $\square$

4.3.2. Ultimately periodic chains with small offset

Lemma 4.3 *Let $m > 0$, $C_i \subseteq A^*$ be an ultimately p_i-periodic chain with offset q_i for $1 \leq i \leq \ell$ and let $C \subseteq A^*$ be an ultimately p-periodic chain with offset $q > \max(q_1, \ldots, q_\ell) + \mathrm{lcm}(p_1, \ldots, p_\ell) \cdot (N_1(\ell+1, m) + 2)$. Then there exists an ultimately p-periodic chain D with offset $q - \mathrm{lcm}(p_1, \ldots, p_\ell)$ such that $(t, \bar{C}, C) \equiv_m^{\mathrm{ch}} (t, \bar{C}, D)$.*

Proof. Let $C = E \cup uv^*F$ with $E \subseteq \Downarrow u$, $F \subseteq \Downarrow v$, $|u| = q$ and $|v| = p$. Then we can write $u = u'xyz$ such that $|u'| = \max(q_1, \ldots, q_\ell)$, $|x|, |y| > 0$ are multiples of $\mathrm{lcm}(p_1, \ldots, p_\ell)$, $z \neq \varepsilon$, and $(t_{u'x}, \bar{C}, C) \equiv_m^{\mathrm{ch}} (t_{u'xy}, \bar{C}, C)$.

When deleting in the structure $(t, \bar{C}, C)$ all nodes from $u'xA^+ \setminus u'xyA^*$, we end up with $(t_{\varepsilon,u'x}, \bar{C}, C) \cdot (t_{u'xy}, \bar{C}, C)$. Since u' is long enough and the lengths of x and y are multiples of $\mathrm{lcm}(p_1, \ldots, p_\ell)$, the structures $(t, \bar{C})$ and $(t_{\varepsilon,u'x}, \bar{C}) \cdot (t_{u'xy}, \bar{C})$ are isomorphic. Hence there is a chain $D \subseteq A^*$ such that

$$\begin{aligned}(t, \bar{C}, C) &= (t_{\varepsilon,u'x}, \bar{C}, C) \cdot (t_{u'x}, \bar{C}, C)\\ &\equiv_m^{\mathrm{ch}} (t_{\varepsilon,u'x}, \bar{C}, C) \cdot (t_{u'xy}, \bar{C}, C)\\ &\cong (t, \bar{C}, D).\end{aligned}$$

This chain has the same period as C, but the offset is reduced by $|y| \geq \mathrm{lcm}(p_1, \ldots, p_\ell)$. $\square$

A similar proof yields the following lemma.

Lemma 4.4 *Let $m > 0$, $C_i \subseteq A^*$ be an ultimately p_i-periodic chain with offset q_i for $1 \le i \le k$ and let $u_i \in A^*$ be words with $|u_i| = q_i$ for $k < i \le \ell$ and let $u \in A^*$ with $|u| \ge \max(q_1, \dots, q_\ell) + \operatorname{lcm}(p_1, \dots, p_k) \cdot (N_1(\ell + 1, m) + 2)$. Then there exists a word $v \in A^*$ with $|v| \le |u| - \operatorname{lcm}(p_1, \dots, p_k)$ and $(t, \bar{C}, \bar{u}, u) \equiv_m^{\mathrm{ch}} (t, \bar{C}, \bar{u}, v)$.*

4.3.3. Ultimately periodic chains with small period

Lemma 4.5 *Let $m \in \mathbb{N}$, $C_i \subseteq A^*$ be an ultimately p_i-periodic chain with offset q_i for $1 \le i \le \ell$ and let $C \subseteq A^*$ be an ultimately p-periodic chain with offset q. Suppose furthermore $p > 2\operatorname{lcm}(p_1, \dots, p_\ell)(N_2(\ell + 1, m) + 2)$ is a multiple of $\operatorname{lcm}(p_1, \dots, p_\ell)$. Then there exists an ultimately p'-periodic chain $D \subseteq A^*$ such that $p' < p$ is a multiple of $\operatorname{lcm}(p_1, \dots, p_\ell)$ and $(t, \bar{C}, C) \equiv_m^{\mathrm{ch}} (t, \bar{C}, D)$.*

Proof. It is sufficient to consider an infinite chain C. In this case, one first shows the existence of $u, w \in A^*$, $v \in w^+$ and $F, F_1, \dots, F_\ell \subseteq A^*$ such that the following hold:

- $|w| = \operatorname{lcm}(p_1, \dots, p_\ell)$ and $p = |v|$
- $\emptyset \ne F \subseteq {\Downarrow}v$ and $v^* F = u^{-1}C =: C'$
- $\emptyset \ne F_i \subseteq {\Downarrow}w$ and $w^* F_i = u^{-1}C_i =: C'_i$ for $1 \le i \le \ell$

The word v can be factorized as $x_1 x_2 x_3$ such that $x_1, x_2 \in w^+, x_3 \in w^*$, and

$$(t_{x_1, v x_1}, \bar{C}', C') \equiv_m^{\mathrm{ch}} (t_{x_1 x_2, v x_1}, \bar{C}', C').$$

For $n > 0$, we have $v^n x_1 = x_1 (x_2 x_3 x_1)^n$. One can show that

$$(t_{x_1, v x_1}, \bar{C}', C') \cong (t_{v^n x_1, v^{n+1} x_1}, \bar{C}', C') \text{ and}$$
$$(t_{x_1 x_2, v x_1}, \bar{C}', C') \cong (t_{v^n x_1 x_2, v^{n+1} x_1}, \bar{C}', C').$$

Hence we have

$$\begin{aligned}
(t, \bar{C}', C') &= (t_{\varepsilon, v x_1}, \bar{C}', C') \cdot \prod_{n>0} (t_{v^n x_1, v^{n+1} x_1}, \bar{C}', C') \\
&\cong (t_{\varepsilon, v x_1}, \bar{C}', C') \cdot (t_{x_1, v x_1}, \bar{C}', C')^\omega \\
&\equiv_m^{\mathrm{ch}} (t_{\varepsilon, v x_1}, \bar{C}', C') \cdot (t_{x_1 x_2, v x_1}, \bar{C}', C')^\omega \\
&\cong (t_{\varepsilon, v x_1}, \bar{C}', C') \cdot \prod_{n>0} (t_{v^n x_1 x_2, v^{n+1} x_1}, \bar{C}', C') \\
&\cong (t, \bar{C}', D')
\end{aligned}$$

for some $|x_3 x_1|$-periodic chain D'.

Then $p' := |x_3 x_1|$ is a multiple of $\operatorname{lcm}(p_1, \dots, p_\ell)$. Furthermore

$$\begin{aligned}
(t, \bar{C}, C) &= (t_{\varepsilon, u}, \bar{C}, C) \cdot (t_u, \bar{C}, C) \\
&\cong (t_{\varepsilon, u}, \bar{C}, C) \cdot (t, \bar{C}', C') \\
&\equiv_m^{\mathrm{ch}} (t_{\varepsilon, u}, \bar{C}, C) \cdot (t, \bar{C}', D') \\
&\cong (t, \bar{C}, D)
\end{aligned}$$

with $D = (C \cap {\Downarrow}u) \cup u D'$. Since the period length of D' equals p', the chain D is ultimately p'-periodic. □

Proposition 4.6 *Let $m \in \mathbb{N}$, $C_i \subseteq A^*$ be an ultimately p_i-periodic chain with offset q_i for $1 \le i \le \ell$ and let $C \subseteq A^*$ be an ultimately p-periodic chain. Then there exists an ultimately p'-periodic chain D such that $(t, \bar{C}, C) \equiv_m^{\mathrm{ch}} (t, \bar{C}, D)$ and $p' \le 2\operatorname{lcm}(p_1, \dots, p_\ell) \cdot (N_2(\ell + 1, m) + 2)$.*

Proof. Set $p'_0 = p \cdot 2\operatorname{lcm}(p_1, \dots, p_\ell) \cdot (N_2(\ell + 1, m) + 2)$. This allows to apply Lemma 4.5 iteratively. The result is a sequence of ultimately p'_i-periodic chains with $p'_0 > p'_1 > \cdots p'_n$. This process terminates once $p'_n \le 2\operatorname{lcm}(p_1, \dots, p_\ell) \cdot (N_2(\ell + 1, m) + 2)$. □

Now we can finally prove that any chain C can be replaced by an ultimately periodic chain D of small period and offset without changing the $\mathrm{MSO}^{\mathrm{ch}}$-properties:

Theorem 4.7 *Let $m \in \mathbb{N}$, $C_i \subseteq A^*$ be ultimately p_i-periodic chains with offset q_i for $1 \le i \le \ell$ and let $C \subseteq A^*$ be a chain. Then there exists an ultimately p'-periodic chain D with offset $q' \le \max(q_1, \dots, q_\ell) + \operatorname{lcm}(p_1, \dots, p_\ell) \cdot (N_1(\ell + 1, m) + 2)$ and $p' \le 2\operatorname{lcm}(p_1, \dots, p_\ell) \cdot (N_2(\ell + 1, m) + 2)$ such that $(t, \bar{C}, C) \equiv_m^{\mathrm{ch}} (t, \bar{C}, D)$.*

Proof. By Prop. 4.2, we can assume C to be ultimately periodic. Prop. 4.6 allows to bound its period by $2\operatorname{lcm}(p_1, \dots, p_\ell) \cdot (N_2(\ell + 1, m) + 2)$. Although this increases the offset, an iterative application of Lemma 4.3 shortens the offset again to a value of at most $\max(q_1, \dots, q_\ell) + \operatorname{lcm}(p_1, \dots, p_\ell) \cdot (N_1(\ell + 1, m) + 2)$ without increasing the period. □

4.4. Bounded $\mathrm{MSO}^{\mathrm{ch}}$-theory

For an $\mathrm{MSO}^{\mathrm{ch}}$-formula ψ and $q, p \in \mathbb{N}$ let $\exists C \le (q, p) : \psi$ stand for "there exists an ultimately p'-periodic chain C with offset at most q and $p' \le p$ such that ψ holds". Similarly, $\exists x \le q : \psi$ means "there exists a word x of length at most q such that ψ holds". The formulas $\forall C \le (q, p) : \psi$ and $\forall x \le q : \psi$ should be understood similarly. A *bounded* $\mathrm{MSO}^{\mathrm{ch}}$*-sentence* is an expression of the form

$$Q_1 C_1 \le (q_1, p_1) \cdots Q_\ell C_\ell \le (q_\ell, p_\ell) \quad Q'_1 x_1 \le r_1 \cdots Q'_k x_k \le r_k : \psi$$

where ψ is a Boolean combination of atomic formulas and $Q_i, Q'_j \in \{\exists, \forall\}$. Standard techniques allow to shift set quantifiers to the front in a prenex normalform formula (at the expense of additional quantifiers). Hence Thm. 4.7 and Lemma 4.4 imply:

Proposition 4.8 *From an* $\mathrm{MSO}^{\mathrm{ch}}$*-sentence* φ*, one can effectively compute a bounded* $\mathrm{MSO}^{\mathrm{ch}}$*-sentence* ψ *such that, for any structure* $\mathcal{A}$*, we have* $\mathcal{A}^*_{\mathrm{ba}} \models^{\mathrm{ch}} \varphi$ *if and only if* $\mathcal{A}^*_{\mathrm{ba}} \models^{\mathrm{ch}} \psi$*.*

Remark 4.9 *If we restrict set quantification in* $\mathrm{MSO}^{\mathrm{ch}}$*-sentences further to ultimately periodic chains, we obtain a new satisfaction relation* $\models^{\mathrm{period}}$ *(that only makes sense for iterations). The above proposition implies in particular that this seemingly new satisfaction relation equals* $\models^{\mathrm{ch}}$*. This consequence parallels Rabin's result [17] where he restricts set quantifications to run over regular sets.*

4.5. Reduction of the $\mathrm{MSO}^{\mathrm{ch}}$-theory to the first-order theory

Theorem 4.10 *From an* $\mathrm{MSO}^{\mathrm{ch}}$*-sentence* φ*, one can effectively compute a first-order sentence* φ' *such that, for any structure* $\mathcal{A}$*, we have* $\mathcal{A}^*_{\mathrm{ba}} \models^{\mathrm{ch}} \varphi$ *if and only if* $\mathcal{A} \models \varphi'$*.*

Proof. Let $\mathcal{A}$ be some structure. For $n \in \mathbb{N}$ let $\mathcal{A}^{\leq n} = (A^{\leq n}, \preceq, (\widehat{R})_{R\in\sigma}, \mathrm{eq})$ where (i) $A^{\leq n}$ is the set of words in A^* of length at most n and (ii) $(u, v) \in \mathrm{eq}$ if and only if there exist $a \in A$ and $u', v' \in A^*$ with $u = u'a$ and $v = v'a$. Now let φ be a bounded $\mathrm{MSO}^{\mathrm{ch}}$-sentence with first-order kernel ψ and let $n \in \mathbb{N}$ be the maximal number appearing in the bounds in φ. Note that ψ does not relate the chains C_i directly, but only indirectly via the individual variables x_j. This allows to write a first-order formula α in the language of $\mathcal{A}^{\leq n}$ such that $\mathcal{A}^*_{\mathrm{ba}} \models^{\mathrm{ch}} \varphi$ if and only if $\mathcal{A}^{\leq n} \models \alpha$. Here, the predicate eq is necessary in order to express the periodicity of a chain.

Note that the first-order theory of $\mathcal{A}^{\leq n}$ can be reduced to that of $\mathcal{A}$. There is even such a reduction that works uniformly in n and $\mathcal{A}$. Hence the proof is complete. $\square$

Since Thm. 4.10 parallels Muchnik's Theorem 3.1, we can derive similar corollaries:

Corollary 4.11 *Let* σ *be some finite relational signature.*

- *If the first-order theory of a* σ*-structure* $\mathcal{A}$ *is decidable, then the* $\mathrm{MSO}^{\mathrm{ch}}$*-theory of its basic iteration* $\mathcal{A}^*_{\mathrm{ba}}$ *is decidable as well.*
- *For any* $m \in \mathbb{N}$*, there exists* $n \in \mathbb{N}$ *such that, for any two* σ*-structures* $\mathcal{A}$ *and* $\mathcal{B}$ *with* $\mathcal{A} \equiv^{\mathrm{FO}}_n \mathcal{B}$*, we have* $\mathcal{A}^*_{\mathrm{ba}} \equiv^{\mathrm{ch}}_m \mathcal{B}^*_{\mathrm{ba}}$*.*

5. FOREG over pushdown systems

In this section we apply our decidability result for $\mathrm{MSO}^{\mathrm{ch}}$ over basic iterations to pushdown systems. We introduce pushdown systems where the stack alphabet is the (possibly infinite) universe of an arbitrary base structure G. Push- and pop operations are triggered via the relations of the base structure G and a finite set of control states, but are independent from the topmost stack symbol. The configuration graph of such a pushdown system is defined as for finite stack alphabets. We study the logic FOREG for these configuration graphs. FOREG is the extension of first-order logic which allows to define new binary predicates by regular expressions over the binary predicates of the base structure $\mathcal{A}$. Based on our decidability result Corollary 4.11 we show that if FOREG is decidable for the base structure G of a pushdown system, then FOREG remains decidable for the configuration graph of the pushdown system (Thm. 5.1).

5.1. The logic FOREG

Let Σ be a finite alphabet of labels and let $G = (A, (E_\sigma)_{\sigma\in\Sigma}, R_1, \dots, R_m)$ be a relational structure, where $E_\sigma \subseteq A \times A$ is a binary relation and $R_1, \dots, R_m$ are additional non-binary relations. For a word $w = \sigma_1 \cdots \sigma_n$ with $\sigma_i \in \Sigma$ we define the binary relation $\xrightarrow{w}_G = E_{\sigma_1} \circ \cdots \circ E_{\sigma_n}$. We have $\xrightarrow{\varepsilon}_G = \mathrm{id}_A$ and $\xrightarrow{\sigma}_G = E_\sigma$ for $\sigma \in \Sigma$. For a regular language $L \subseteq \Sigma^*$ we define $\mathrm{reach}_L = \bigcup_{w\in L} \xrightarrow{w}_G$. An *FOREG-formula* over the structure G is simply a first-order formula over the extended structure $(A, (\mathrm{reach}_L)_{L\in\mathrm{REG}(\Sigma)}, R_1, \dots, R_m)$.

5.2. Pushdown systems over infinite stack alphabets

A *pushdown system* $S = (Q, G, \tau)$ over a *stack structure* G is given by the following data:

- G is a relational structure of the form $G = (A, (\mathrm{eq}_\alpha)_{\alpha\in\Sigma_1}, (\mathrm{push}_\beta)_{\beta\in\Sigma_2}, (\mathrm{pop}_\gamma)_{\gamma\in\Sigma_3}, \perp)$, where Σ_1, Σ_2, Σ_3 are finite and mutually disjoint alphabets (let $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \Sigma_3$ in the following), $\mathrm{eq}_\alpha \subseteq A \times A$, $\mathrm{push}_\beta, \mathrm{pop}_\gamma \subseteq A$, and $\perp \in A$.
- Q is a finite set of states such that $Q \cap A = \emptyset$.
- $\tau : \Sigma \to Q \times Q$

With S we associate the configuration graph $\mathcal{C}(S) = (A^+Q, (E_\sigma)_{\sigma\in\Sigma})$, where:

- $E_\alpha = \{(wap, wbq) \mid w \in A^*, (a, b) \in \mathrm{eq}_\alpha, \tau(\alpha) = (p, q)\}$ for $\alpha \in \Sigma_1$
- $E_\beta = \{(wp, waq) \mid w \in A^+, a \in \mathrm{push}_\beta, \tau(\beta) = (p, q)\}$ for $\beta \in \Sigma_2$
- $E_\gamma = \{(wap, wq) \mid w \in A^+, a \in \mathrm{pop}_\gamma, \tau(\gamma) = (p, q)\}$ for $\gamma \in \Sigma_3$

The following theorem is the main result of this section:

Theorem 5.1 *Let G be a stack structure with decidable FOREG-theory. Then the configuration graph $\mathcal{C}(S)$ has a decidable FOREG-theory.*

The rest of this section is devoted to the proof of this theorem. The idea is to define, using the logic FOREG, a suitable structure $\mathcal{A}$ in the stack structure G. Since we assume that the FOREG-theory of G is decidable, it follows that the first-order theory of $\mathcal{A}$ is decidable. Thus, by Corollary 4.11 the $\mathrm{MSO}^{\mathrm{ch}}$-theory of the basic iteration $\mathcal{A}^*_{\mathrm{ba}}$ is decidable. To obtain Thm. 5.1, we give an $\mathrm{MSO}^{\mathrm{ch}}$-interpretation of the configuration graph $\mathcal{C}(S)$ in $\mathcal{A}^*_{\mathrm{ba}}$.

For a finite automaton T and states μ and ν of T, let $L(T,\mu,\nu)$ be the set of words that label some path from μ to ν in T.

Fix regular languages $L_1,\dots,L_k \subseteq \Sigma^*$. Then there exists a finite automaton $T = (\Theta,\Sigma,\delta)$ with state set Θ such that every language L_i is a union of languages of the form $L(T,\mu,\nu)$ for certain states $\mu,\nu \in \Theta$. For $\mu,\nu \in \Theta$ let

$$\mathrm{reach}_{\mu,\nu} = \{(up,vq) \in A^+Q \times A^+Q \mid \exists w \in L(T,\mu,\nu) : up \xrightarrow{w}_{\mathcal{C}(S)} vq\}.$$

Thus, $\mathrm{reach}_{\mu,\nu} = \mathrm{reach}_L$ for $L = L(T,\mu,\nu)$. We will show that the first-order theory of the structure

$$\mathcal{B} = (A^+Q, (\mathrm{reach}_{\mu,\nu})_{\mu,\nu\in\Theta}) \qquad (2)$$

is decidable. Since the decision procedure for the first-order theory of $\mathcal{B}$ will be uniform in the automaton T, this proves Thm. 5.1.

Let $\mu,\nu \in \Theta$, $p,q \in Q$, $u \in A^+$, and $a \in A$. We write $(up,uaq) \in \mathrm{reach}^{(+)}_{\mu,\nu}$ if and only if there exist $\beta \in \Sigma_2$ and $x \in \Sigma^*$ such that $\beta x \in L(T,\mu,\nu)$, $up \xrightarrow{\beta x}_{\mathcal{C}(S)} uaq$, $|y|_{\Sigma_2} \geq |y|_{\Sigma_3}$ for every prefix y of x, and $|x|_{\Sigma_2} = |x|_{\Sigma_3}$. Thus, (up,uaq) belongs to the relation $\mathrm{reach}^{(+)}_{\mu,\nu}$ if there exists a path from up to uaq in the configuration graph $\mathcal{C}(S)$ whose label belongs to $L(T,\mu,\nu)$ such that all the configurations along this path except the very first one up are of the form uvr for some $v \in A^+$, and $r \in Q$. Note that $(up,uaq) \in \mathrm{reach}^{(+)}_{\mu,\nu}$ implies $(vp,vaq) \in \mathrm{reach}^{(+)}_{\mu,\nu}$ for all $v \in A^+$. Symmetrically, we write $(uap,uq) \in \mathrm{reach}^{(-)}_{\mu,\nu}$ if and only if there exist $\gamma \in \Sigma_3$ and $x \in \Sigma^*$ such that $x\gamma \in L(T,\mu,\nu)$, $uap \xrightarrow{x\gamma}_{\mathcal{C}(S)} uq$, $|y|_{\Sigma_2} \geq |y|_{\Sigma_3}$ for every prefix y of x, and $|x|_{\Sigma_2} = |x|_{\Sigma_3}$. Finally, for $\mu,\nu \in \Theta$, $p,q \in Q$, $u,v \in A^+$ we write $(up,vq) \in \mathrm{reach}^{(=)}_{\mu,\nu}$ if and only if there exists $w \in L(T,\mu,\nu)$ such that $up \xrightarrow{w}_{\mathcal{C}(S)} vq$, $|y|_{\Sigma_2} \geq |y|_{\Sigma_3}$ for every prefix y of w, and $|w|_{\Sigma_2} = |w|_{\Sigma_3}$. Thus, (up,vq) belongs to the relation $\mathrm{reach}^{(+)}_{\mu,\nu}$ if there exists a path from up to vq in the configuration graph $\mathcal{C}(S)$ whose label belongs to $L(T,\mu,\nu)$ such that all the configurations along this path are of the form wr for some $r \in Q$ and $w \in A^+$ with $|w| = |u| = |v|$. Thus, $(uap,ubq) \in \mathrm{reach}^{(=)}_{\mu,\nu}$ for some (and hence all) $u \in A^*$ if and only if $(ap,bq) \in \mathrm{reach}_{\mu,\nu}$.

Lemma 5.2 *For $c,d \in A^+Q$ and $\mu,\nu \in \Theta$ we have $(c,d) \in \mathrm{reach}_{\mu,\nu}$ if and only if there exist $m,n \geq 0$, $\mu_m,\dots,\mu_0,\nu_0,\dots,\nu_n \in \Theta$, and configurations $c_m,\dots,c_0,d_0,\dots,d_n \in A^+Q$ such that:*

- $c_m = c$, $d_n = d$, $\mu_m = \mu$, $\nu_n = \nu$
- $(c_i,c_{i-1}) \in \mathrm{reach}^{(-)}_{\mu_i,\mu_{i-1}}$ *for* $1 \leq i \leq m$
- $(c_0,d_0) \in \mathrm{reach}^{(=)}_{\mu_0,\nu_0}$
- $(d_{j-1},d_j) \in \mathrm{reach}^{(+)}_{\nu_{j-1},\nu_j}$ *for* $1 \leq j \leq n$

Proof. The if-direction is obvious. For the other direction take for c_0 (resp. d_0) the leftmost (resp. rightmost) occurrence of a configuration of minimal height along a path in $\mathcal{C}(S)$ from the configuration c to the configuration d. □

For all $p,q \in Q$ and $\mu,\nu \in \Theta$ define a binary predicate

$$H(p,\mu,q,\nu) = \{(a,b) \in A \times A \mid (ap,bq) \in \mathrm{reach}_{\mu,\nu}\}.$$

Lemma 5.3 *The relation $H(p,\mu,q,\nu)$ is effectively* FOREG*-definable over the stack structure G.*

Proof. We will construct effectively a finite automaton B with state set $Q \times \Theta$ and alphabet Σ_1 such that $(ap,bq) \in \mathrm{reach}_{\mu,\nu}$ if and only if $a \xrightarrow{u}_G b$ for some $u \in L(B,(p,\mu),(q,\nu))$, which proves the lemma. For this, we will construct a finite sequence of automata B_i ($i \geq 0$) with state set $Q \times \Theta$ and alphabet Σ_1, which converges to the automaton B. The finite state automaton B_0 contains the transition $(p,\mu) \xrightarrow{\alpha} (q,\nu)$ if and only if $\tau(\alpha) = (p,q)$ and $\mu \xrightarrow{\alpha}_T \nu$ for some $\alpha \in \Sigma_1$. Now assume that B_i is already constructed and assume that

- in T there are transitions $\mu \xrightarrow{\beta}_T \mu'$, ($\beta \in \Sigma_2$) and $\nu' \xrightarrow{\gamma}_T \nu$ ($\gamma \in \Sigma_3$),
- $\tau(\beta) = (p,p')$, $\tau(\gamma) = (q',q)$, and
- there exist $a \in \mathrm{push}_\beta, b \in \mathrm{pop}_\gamma$, and $u \in L(B_i,(p',\mu'),(q',\nu'))$ such that $a \xrightarrow{u}_G b$.

Note that the last point is decidable, since the FOREG-theory of G is decidable. In this situation we add the ε-transition $(p,\mu) \xrightarrow{\varepsilon} (q,\nu)$ to B_i and call the resulting automaton B_{i+1}. We repeat this process as long as we will add new ε-transitions. Note that in each step the state set is not changed. Let B be the resulting automaton. It is not difficult to prove $(ap,bq) \in \mathrm{reach}_{\mu,\nu}$ if and only if $a \xrightarrow{u}_G b$, for some $u \in L(B,(p,\mu),(q,\nu))$, which proves the lemma. □

Define for all $p, q \in Q$ and $\mu, \nu \in \Theta$ unary predicates $D(p,\mu,q,\nu), U(p,\mu,q,\nu) \subseteq A$ as follows:

$$a \in D(p,\mu,q,\nu) \;\Leftrightarrow$$
$$\bigvee_{\substack{p' \in Q, \mu' \in \Theta, \\ \gamma \in \Sigma_3}} \exists b \in \mathrm{pop}_\gamma : \left\{ \begin{array}{l} (a,b) \in H(p,\mu,p',\mu') \wedge \\ \tau(\gamma) = (p',q) \;\wedge\; \mu' \xrightarrow{\gamma}_T \nu \end{array} \right\}$$
$$a \in U(p,\mu,q,\nu) \;\Leftrightarrow$$
$$\bigvee_{\substack{p' \in Q, \mu' \in \Theta, \\ \beta \in \Sigma_2}} \exists b \in \mathrm{push}_\beta : \left\{ \begin{array}{l} (b,a) \in H(p',\mu',q,\nu) \wedge \\ \tau(\beta) = (p,p') \;\wedge\; \mu \xrightarrow{\beta}_T \mu' \end{array} \right\}$$

By Lemma 5.3, the unary predicates $D(p,\mu,q,\nu)$ and $U(p,\mu,q,\nu)$ are FOREG-definable in the stack structure G. The next lemma follows directly from the definition of the predicates $D(p,\mu,q,\nu)$, $H(p,\mu,q,\nu)$, and $U(p,\mu,q,\nu)$.

Lemma 5.4 *We have:*

$$D(p,\mu,q,\nu) = \{a \in A \mid \exists u \in A^+ : (uap, uq) \in \mathrm{reach}^{(-)}_{\mu,\nu}\}$$
$$U(p,\mu,q,\nu) = \{a \in A \mid \exists u \in A^+ : (up, uaq) \in \mathrm{reach}^{(+)}_{\mu,\nu}\}$$

Note that in Lemma 5.4, one might replace the quantifier $\exists u \in A^+$ by $\forall u \in A^+$. Lemma 5.2 and Lemma 5.4 imply:

Lemma 5.5 *We have $(up, vq) \in \mathrm{reach}_{\mu,\nu}$ if and only if there exist $m, n \geq 0$, $w \in A^*$, $a_i \in A, p_i \in Q, \mu_i \in \Theta$ ($0 \leq i \leq m$), and $b_j \in A, q_j \in Q, \nu_j \in \Theta$ ($0 \leq j \leq m$) such that:*

- $\mu = \mu_m$, $\nu = \nu_n$, $p = p_m$, $q = q_n$,
- $u = wa_0 \cdots a_m$, $v = wb_0 \cdots b_n$,
- $a_i \in D(p_i, \mu_i, p_{i-1}, \mu_{i-1})$ *for all* $1 \leq i \leq m$,
- $(a_0, b_0) \in H(p_0, \mu_0, q_0, \nu_0)$,
- $b_j \in U(q_{j-1}, \nu_{j-1}, q_j, \nu_j)$ *for all* $1 \leq j \leq n$

Define the binary relations $D(p,\mu,q,\nu)^*, U(p,\mu,q,\nu)^* \subseteq A^+ \times A^+$ as follows:

$(u,v) \in D(p,\mu,q,\nu)^*$ if and only if there exist $m \geq 0$, $a_1, \ldots, a_m \in A$, $p_0, \ldots, p_m \in Q$, $\mu_0, \ldots, \mu_m \in \Theta$ with

- $u = va_1 \cdots a_m$,
- $p = p_m, q = p_0, \mu = \mu_m, \nu = \mu_0$, and
- $a_i \in D(p_i, \mu_i, p_{i-1}, \mu_{i-1})$ for $1 \leq i \leq m$.

$(u,v) \in U(p,\mu,q,\nu)^*$ if and only if there exist $n \geq 0$, $b_1, \ldots, b_n \in A$, $q_0, \ldots, q_n \in Q$, $\nu_0, \ldots, \nu_n \in \Theta$ with

- $v = ub_1 \cdots b_n$,
- $p = q_0, q = q_n, \mu = \nu_0, \nu = \nu_n$, and
- $b_j \in U(q_{j-1}, \nu_{j-1}, q_j, \nu_j)$ for $1 \leq j \leq n$.

These definitions and Lemma 5.5 imply:

Lemma 5.6 *We have $(up, vq) \in \mathrm{reach}_{\mu,\nu}$ if and only if there exist $u', v' \in A^+$, $p', q' \in Q$, $\mu', \nu' \in \Theta$ such that:*

$$(u,u') \in D(p,\mu,p',\mu')^*$$
$$(u',v') \in \{(xa, xb) \mid x \in A^*, (a,b) \in H(p',\mu',q',\nu')\}$$
$$(v',v) \in U(q',\nu',q,\nu)^*$$

Now, let us consider the structure

$$\begin{aligned}\mathcal{A} = (A \cup Q, (q)_{q\in Q}, &(D(p,\mu,q,\nu))_{p,q\in Q,\mu,\nu\in\Theta},\\ &(H(p,\mu,q,\nu))_{p,q\in Q,\mu,\nu\in\Theta},\\ &(U(p,\mu,q,\nu))_{p,q\in Q,\mu,\nu\in\Theta}).\end{aligned}$$

Since each of its relations is FOREG-definable in the stack structure G, and G has a decidable FOREG-theory, $\mathcal{A}$ has a decidable first-order theory. Thus, by Thm. 4.10,

$$\begin{aligned}\mathcal{A}^*_{\mathrm{ba}} = ((A \cup Q)^*, \preceq, (\widehat{q})_{q\in Q}, &\\ &(\widehat{D(p,\mu,q,\nu)})_{p,q\in Q,\mu,\nu\in\Theta},\\ &(\widehat{H(p,\mu,q,\nu)})_{p,q\in Q,\mu,\nu\in\Theta},\\ &(\widehat{U(p,\mu,q,\nu)})_{p,q\in Q,\mu,\nu\in\Theta})\end{aligned}$$

has a decidable $\mathrm{MSO}^{\mathrm{ch}}$-theory. We finally show that the structure $\mathcal{B}$ from (2), for which we have to show decidability of the first-order theory, is $\mathrm{MSO}^{\mathrm{ch}}$-interpretable in $\mathcal{A}^*_{\mathrm{ba}}$, which proves Thm. 5.1. Clearly, the universe A^+Q of $\mathcal{B}$ is first-order definable in $\mathcal{A}^*_{\mathrm{ba}}$ using the prefix relation $\preceq$ and the unary relations $\widehat{q} = (A \cup Q)^* q$ for $q \in Q$. In order to define the relations $\mathrm{reach}_{\mu,\nu}$ of $\mathcal{B}$ it suffices by Lemma 5.6 to define the binary relations $D(p,\mu,q,\nu)^*, U(p,\mu,q,\nu)^*, \{(xa, xb) \mid x \in A^*, (a,b) \in H(p,\mu,q,\nu)\} \subseteq A^+ \times A^+$ in $\mathcal{A}$ using $\mathrm{MSO}^{\mathrm{ch}}$. The relation $\{(xa, xb) \mid x \in A^*, (a,b) \in H(p,\mu,q,\nu)\}$ can be defined as $\widehat{H(p,\mu,q,\nu)} \cap A^+ \times A^+$; note that A^+ is first-order definable in $\mathcal{A}$ using the prefix relation $\preceq$ and the relations $\widehat{q}$ for $q \in Q$. Finally, the $\mathrm{MSO}^{\mathrm{ch}}$-definitions of $D(p,\mu,q,\nu)^*$ and $U(p,\mu,q,\nu)^*$ follow Büchi's technique for expressing the existence of a successful run of an automaton on a finite word in MSO. This completes the proof of Thm. 5.1.

6. Open problems

We have shown that the $\mathrm{MSO}^{\mathrm{ch}}$-theory (i.e., the fragment of the full MSO-theory where set quantification is restricted to chains) of the basic iteration $\mathcal{A}^*_{\mathrm{ba}}$ of a structure $\mathcal{A}$ can be reduced to the first-order theory of $\mathcal{A}$. Using this result, we have shown that the FOREG-theory of the configuration graph $\mathcal{C}(S)$ of a pushdown system S over an infinite stack structure $\mathcal{A}$ can be reduced to the FOREG-theory of $\mathcal{A}$. We plan to investigate whether similar preservation theorems can be shown also for temporal logics like CTL, CTL*, or the modal μ-calculus. Another interesting candidate for investigations of this kind is TC^2 [26], i.e., first-order logic extended by the transitive closure operator for binary relations.

References

[1] R. Alur, M. Benedikt, K. Etessami, P. Godefroid, T. W. Reps, and M. Yannakakis. Analysis of recursive state machines. *ACM Trans. Program. Lang. Syst*, 27(4):786–818, 2005.

[2] D. Berwanger and A. Blumensath. The monadic theory of tree-like structures. In *Automata, logics, and infinite games*, LNCS 2500, pages 285–301. Springer, 2002.

[3] A. Blumensath and S. Kreutzer. An extension of muchnik's theorem. *J. Log. Comput.*, 15:59–64, 2005.

[4] A. Carayol and S. Wöhrle. The Caucal hierarchy of infinite graphs in terms of logic and higher-order pushdown automata. In *FSTTCS 2003*, LNCS 2914, pages 112–123. Springer, 2003.

[5] D. Caucal. On the transition graphs of Turing machines. *Theor. Comput. Sci.*, 296(2):195–223, 2003.

[6] B. Courcelle and I. Walukiewicz. Monadic second-order logic, graph coverings and unfoldings of transition systems. *Ann. Pure Appl. Logic*, 92(1):35–62, 1998.

[7] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer, 1991.

[8] J. Esparza, S. Kiefer, and S. Schwoon. Abstraction refinement with Craig interpolation and symbolic pushdown systems. In *TACAS 06*, LNCS 3920, pages 489–503. Springer, 2006.

[9] J. Esparza, A. Kucera, and S. Schwoon. Model checking LTL with regular valuations for pushdown systems. *Inf. Comput.*, 186(2):355–376, 2003.

[10] Y. Gurevich. Monadic second-order theories. In J. Barwise and S. Feferman, editors, *Model-Theoretic Logics*, pages 479–506. Springer, 1985.

[11] Y. Gurevich and S. Shelah. Rabin's uniformization problem. *J. of Symb. Logic*, 48:1105–1119, 1983.

[12] B. Khoussainov and A. Nerode. Automatic presentations of structures. In *Logic and Computational Complexity*, LNCS 960, pages 367–392. Springer, 1995.

[13] O. Kupferman and M. Y. Vardi. An automata-theoretic approach to reasoning about infinite-state systems. In *CAV 2000*, LNCS 1855, pages 36–52. Springer, 2000.

[14] D. Kuske and M. Lohrey. Logical aspects of Cayley-graphs: The monoid case. *Int. J. Algebra Comput.*, 16(2):307–340, 2006.

[15] F. Neven and T. Schwentick. Expressive and efficient pattern languages for tree-structured data. In *PODS'00*, pages 145–156. ACM Press, 2000.

[16] D. Perrin and J.-E. Pin. *Infinite Words. Automata, semigroups, logic and games*. Pure and Applied Mathematics vol. 141. Elsevier, 2004.

[17] M. Rabin. *Automata on infinite objects and Church's problem*. American Mathematical Society, Providence, R.I., 1972. Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 13.

[18] M. O. Rabin. Decidability of second-order theories and automata on infinite trees. *Trans. Am. Math. Soc.*, 141:1–35, 1969.

[19] T. Schwentick. On diving in trees. In *MFCS'00*, LNCS 1893, pages 660–669. Springer, 2000.

[20] A. L. Semenov. Decidability of monadic theories. In *MFCS'84*, LNCS 176, pages 162–175. Springer, 1984.

[21] S. Shelah. The monadic theory of order. *Ann. Math. (2)*, 102:379–419, 1975.

[22] J. Stupp. The lattice-model is recursive in the original model. The Hebrew University, Jerusalem, 1975.

[23] W. Thomas. On chain logic, path logic, and first-order logic over infinite trees. In *LICS'87*, pages 245–256. IEEE Computer Society Press, 1987.

[24] I. Walukiewicz. Pushdown processes: Games and model-checking. *Inf. Comput.*, 164(2):234–263, 2001.

[25] I. Walukiewicz. Monadic second-order logic on tree-like structures. *Theor. Comput. Sci.*, 275(1–2):311–346, 2002.

[26] S. Wöhrle and W. Thomas. Model checking synchronized products of infinite transition systems. In *LICS'04*, pages 2–11. IEEE Computer Society Press, 2004.