On First-Order Fragments for Mazurkiewicz Traces

Volker Diekert   Martin Horsch
Manfred Kufleitner

{diekert, kufleitner}@fmi.uni − stuttgart.de

July 24, 2007

Abstract

Mazurkiewicz traces form a model for concurrency. Temporal logic and first-order logic are important tools in order to deal with the abstract behavior of such systems. Since typical properties can be described by rather simple logical formulas one is interested in logical fragments.

One focus of this paper is unary temporal logic and first-order logic in two variables. Over words, this corresponds to the variety of finite monoids called **DA**. However, over Mazurkiewicz traces it is crucial whether traces are given as dependence graphs or as partial orders (over words these notions coincide). The main technical contribution is a generalization of important characterizations of **DA** from words to dependence graphs, whereas the use of partial orders leads to strictly larger classes. As a consequence we can decide whether a first-order formula over dependence graphs is equivalent to a first-order formula in two variables. The corresponding result for partial orders is not known.

This difference between dependence graphs and partial orders also affects the complexity of the satisfiability problems for the fragments under consideration: for first-order formulas in two variables we prove an NEXPTIME upper bound, whereas the corresponding problem for partial orders leads to EXPSPACE.

Furthermore, we give several separation results for the alternation hierarchy for first-order logic. It turns out that even for those levels at which one can express the partial order relation in terms of dependence graphs, the fragments over partial orders have more expressive power.

1 Introduction

According to one of the first sentences in the trend-setting 1977 Aarhus paper [16] of Mazurkiewicz a *trace is a partially ordered set of symbol occurrences*. He proposed *trace theory* as an algebraic framework for studying concurrent processes and observed that the behavior of a concurrent process is not described by a string, but more accurately by some labeled partial order. The partial order relation of a trace is defined via a dependence alphabet so that the set of traces forms a *free partially commutative monoid* in the sense of Cartier and Foata [1]. There is a natural extension to infinite objects which lead to the notion of *real* *trace*. However in this paper we focus on finite traces, only. For an overview on trace theory we refer to *The Book of Traces* [4].

One advantage of trace theory is that formal specifications of concurrent systems by temporal logic formulas have a direct interpretation for Mazurkiewicz traces. Moreover, if the underlying alphabet is fixed, satisfiability of all local temporal logics over traces where the modalities are definable in monadic second order logic are decidable in PSPACE [10]. For some local temporal logics this result has been generalized to the case where the alphabet is also part of the input [11]. This is optimal since the PSPACE-hardness occurs already for words (over a two letter alphabet).

On the algebraic side temporal logics corresponds to aperiodic monoids, this is known as *Kamp's Theorem*, see e.g. [9]. Kamp's Theorem has been generalized to traces in [2, 3, 21]. However typical formulas do not require the full expressive power of temporal logic. So one is interested in fragments. The probably most interesting fragment is given on the algebraic side by the variety **DA**. Over words, **DA** admits many different characterizations, which led to the title *Diamonds are Forever* in [19]. In this paper we show that some of these characterizations can be generalized to trace monoids, but it turns out that we have to distinguish whether traces are given as dependence graphs or as partial orders. This is rather unexpected since for words and also for the full first-order theory over traces there is no such difference. But for the expressive power of fragments as well as for the complexity of the satisfiability problem this distinction is rather crucial.

The paper is organized as follows: In Section 2 we introduce basic notions and facts that are needed in the remainder of the paper. In Section 3 we give a language theoretic characterization of the first order fragment Σ_2 over dependence graphs. Section 4 introduces (local) unary temporal logic and exposes its connection to first-order logic with two variables FO^2. The interest in the fragment FO^2 arises from the fact that first-order logic with only three variables has already the same expressive power as the full first-order theory over traces [12]. Section 5 carries together several results connected to the variety **DA**. Theorem 5.1 gives characterizations of **DA** over traces such as unary temporal logic, FO^2 over dependence graphs, and first-order logic with one quantifier alternation over dependence graphs. Theorem 5.2 shows that as soon as a logical fragment is capable to express concurrency, this fragment exceeds the expressive power of **DA**. Section 6 contains several complexity results for the satisfiability problem of the fragments under consideration. In Theorems 6.1 and 6.2 we show that satisfiability for unary temporal logic without an operator for concurrency is NP-complete whereas additionally allowing a modality for concurrency yields PSPACE-completeness. In Theorems 6.3, 6.4, and 6.5 we give some complexity results for FO^2 over traces. In Section 7 we give several separation results. We start by introducing Ehrenfeucht-Fraïssé games for the respective logical fragments. Those games are the main tool in this section. Theorem 7.1 says that in general, in FO^2 over partial orders one cannot express whether there are three pairwise concurrent actions. The remainder of this section discusses the alternation hierarchy for first-order logic. Theorem 7.2 shows that for every $n \in \mathbb{N}$ there exists a trace monoid $\mathbb{M}$ and a language $L \subseteq \mathbb{M}$ such that L is expressible in FO^2 over partial orders whereas L is not expressible at level n of the alternation hierarchy. Therefore, restricting the number of variables on one hand and restricting the number of quantifier alternations on the other hand yields incomparable fragments. Theorem 7.3 depicts the relation of the alternation hierarchy over partial orders and the alternation hierarchy over dependence graphs.

2 Preliminaries

A *dependence alphabet* is a pair (Γ, D) where the alphabet Γ is a finite set (of actions) and the *dependence relation* $D \subseteq \Gamma \times \Gamma$ is reflexive and symmetric. The *independence relation* I is the complement of D. It is irreflexive and symmetric. A *Mazurkiewicz trace* is an isomorphism class of a node-labeled directed acyclic graph $t = [V, E, \lambda]$, where V is a finite set of vertices labeled by $\lambda : V \to \Gamma$ and $E \subseteq (V \times V) \setminus \mathrm{id}_V$ is the edge relation such that

$$E \cup E^{-1} \cup \mathrm{id}_V = \{ (x, y) \in V \times V \mid (\lambda(x), \lambda(y)) \in D \}$$

We call $[V, E, \lambda]$ a *dependence graph*. By E^+ we mean the transitive closure of E, frequently we write $<$ instead of E^+, and $\leq$ means the induced partial order, i.e., $\leq$ is the union of the relations $<$ and id_V. We write $x \parallel y$ if the vertices x and y are incomparable with respect to $\leq$. Note that node labeled graphs (V, E, λ) and (V', E', λ') are isomorphic if and only if the corresponding labeled partial orders $(V, \leq, \lambda)$ and $(V', \leq', \lambda')$ are isomorphic. Hence as mentioned above, a trace is indeed a partially ordered set of symbol occurrences.

Let $t_1 = [V_1, E_1, \lambda_1]$ and $t_2 = [V_2, E_2, \lambda_2]$ be traces. Then we define the concatenation of t_1 and t_2 to be $t_1 \cdot t_2 = [V, \leq, \lambda]$ where $V = V_1 \cup V_2$ is a disjoint union, $\lambda = \lambda_1 \cup \lambda_2$, and $E = E_1 \cup E_2 \cup \{ (x, y) \in V_1 \times V_2 \mid (\lambda(x), \lambda(y)) \in D \}$. The set $\mathbb{M}$ of traces becomes a monoid with the empty trace $1 = (\emptyset, \emptyset, \emptyset)$ as unit. It is generated by Γ, where a letter a is viewed as a graph with a single vertex labeled by a. Thus, we obtain a canonical surjective homomorphism

$$\pi : \Gamma^* \to \mathbb{M}$$

The effect of the mapping π can be made explicit as follows. We start with a word $w = a_1 \cdots a_n$ where all a_x are letters in Γ. Each x is viewed as an element in $\{ 1, \ldots, n \}$, with label $\lambda(x) = a_x$. We draw an arc from a_x to a_y if and only if both, $x < y$ and $(a_x, a_y) \in D$. This dependence graph is $\pi(w)$. Note that $\mathbb{M}$ is also canonically isomorphic to the quotient monoid $\Gamma^* / \{ ab = ba \mid (a, b) \in I \}$.

A *subtrace* s of a trace $t = [V, E, \lambda]$ is defined by a subset V' of V. Formally, we let $s = [V', E', \lambda']$ be the induced labeled subgraph. This means that $E' = E \cap V' \times V'$ and $\lambda' : V' \to \Gamma$ is the restriction of λ. It is clear that s is a dependence graph, however in general it is no factor. A *factor* of t is a subtrace $s = [V', E', \lambda']$ such that $(x, y), (y, z) \in E^+$ with $x, z \in V'$ implies $y \in V'$, too. Every factor s of t yields a factorization $t = psq$ (possibly more than one). For the other direction, every factorization $t = psq$ of $t = [V, E, \lambda]$ yields a unique factor $s = [V', E', \lambda']$ with $V' \subseteq V$.

If $t = [V, E, \lambda]$ then throughout we write $x \in t$ instead of $x \in V$. By $d(\Gamma, D)$ we denote the length of a longest simple path in the undirected graph (Γ, D).

There is a basic observation which holds for all $x, y \in t \in \mathbb{M}$ with $d = d(\Gamma, D)$:

$$(x,y) \in E \;\Leftrightarrow\; (x,y) \in E^+ \wedge (\lambda(x), \lambda(y)) \in D \tag{1}$$

$$(x,y) \in E^+ \;\Leftrightarrow\; \exists x_1 \cdots \exists x_d \colon \left\{ \begin{array}{c} x_d = y \,\wedge\, (x, x_1) \in E \,\wedge \\ \displaystyle\bigwedge_{1 \le i < d} (x_i, x_{i+1}) \in E \cup \mathrm{id}_V \end{array} \right\} \tag{2}$$

There are some standard notations we adopt here. The *length* of a trace t is denoted by $|t|$ and its *alphabet* is $\mathrm{alph}(t)$. By $\min(t)$ ($\max(t)$ resp.) we denote the set of minimal (maximal resp.) vertices of t. Note that $\min(t)$ and $\max(t)$ are factors of t. A *trace language* L is a subset of $\mathbb{M}$. It is called *recognizable* if $L = h^{-1}h(L)$ for some homomorphism $h : \mathbb{M} \to M$, where M is a finite monoid.

We are interested in first-order definable languages, hence we restrict our attention to aperiodic monoids. A finite monoid M is called *aperiodic* if there is some $n \ge 0$ such that $v^n = v^{n+1}$ for all $v \in M$. By $\mathbf{A}$ we denote the variety of all finite aperiodic monoids. The variety $\mathbf{DA}$ is defined as all finite monoids M satisfying the equation $(uvw)^n v (uvw)^n = (uvw)^n (uvw)^n$ for all $u, v, w \in M$ and some $n \ge 0$. Setting $u = 1$ and $w = 1$ we see that $\mathbf{DA}$ contains aperiodic monoids, only. By $\mathbf{DA}(\mathbb{M})$ we mean the class of trace languages which are recognized by some finite monoid in $\mathbf{DA}$. We have

$$L \in \mathbf{DA}(\mathbb{M}) \;\Leftrightarrow\; \pi^{-1}(L) \in \mathbf{DA}(\Gamma^*) \tag{3}$$

This is a special case of a well-known result due to the fact that π induces an isomorphism between the syntactic monoids of L and $\pi^{-1}(L)$, see e.g. [5]. In fact, if $\pi^{-1}(L)$ is recognized by some finite monoid M, then L is recognized by some quotient of a submonoid of M.

The syntax of first-order logic formulas $\mathrm{FO}[E]$ is built upon atomic formulas of type

$$\top, \;\; \lambda(x) = a, \;\; \text{and} \;\; (x,y) \in E$$

where $\top$ means *true*, x, y are variables and $a \in \Gamma$ is a letter. If φ, ψ are first-order formulas, then $\neg\varphi$, $\varphi \vee \psi$, $\exists x\, \varphi$ are first-order formulas, too. We use the usual shorthands as $\bot = \neg\top$ meaning *false*, $\varphi \wedge \psi = \neg(\neg\varphi \vee \neg\psi)$, and $\forall x\, \varphi = \neg\exists x\, \neg\varphi$. Given $\varphi \in \mathrm{FO}[E]$ the semantics is defined as usual for node labeled graphs. In particular, if all free variables in φ belong to a set $\{ x_1, \dots, x_m \}$, then for all $t \in \mathbb{M}$ and all $x_1, \dots, x_m \in t$ we can write $t, x_1, \dots, x_m \models \varphi$ and this has a well-defined truth value. We identify formulas by semantic equivalence. Hence, if φ and ψ are formulas with m free variables, then we write $\varphi = \psi$ as soon as $t, x_1, \dots, x_m \models (\varphi \leftrightarrow \psi)$ for all trace monoids $\mathbb{M}$, all traces $t \in \mathbb{M}$, and all positions $x_1, \dots, x_m \in t$. There are completely analogous definitions for the first-order logic $\mathrm{FO}[<]$. The only difference is that instead of $(x,y) \in E$ we have an atomic predicate $x < y$. We let $\mathrm{FO}^m[E]$ be the set of all formulas with at most m different names for variables. Moreover, we allow $(\lambda(x), \lambda(y)) \in D$ as an additional atomic formula (which has an interpretation for all $x, y \in t$ where t is a trace). With this convention $\mathrm{FO}^m[E]$ becomes a fragment of $\mathrm{FO}^m[<]$ due to (1).

A *first-order sentence* is a formula in $\mathrm{FO}[E]$ or $\mathrm{FO}[<]$ without free variables. For a first-order sentence φ we define $L(\varphi) = \{ t \in \mathbb{M} \mid t \models \varphi \}$. A trace language $L \subseteq \mathbb{M}$ is called *first-order definable* if $L = L(\varphi)$ for some first-order sentence φ and we let $\mathrm{FO}(\mathbb{M}) = \{ L(\varphi) \mid \varphi \in \mathrm{FO}[E] \}$. We do not write

$\mathrm{FO}[E](\mathbb{M})$, because $\mathrm{FO}(\mathbb{M}) = \{ L(\varphi) \mid \varphi \in \mathrm{FO}[<] \}$ as well, due to (2). So, in first-order it is not necessary to distinguish between E and $<$. However, for subclasses of FO we need this distinction. We define the following classes for $E' = E$ and $E' = <$ respectively. We let $\Sigma_0[E'] = \Pi_0[E']$ be the set of all Boolean combinations of atomic formulas. For $n > 0$ the classes $\Sigma_n[E']$ and $\Pi_n[E']$ are inductively defined by the following conditions:

i. $\varphi \in \Sigma_{n-1}[E'] \cup \Pi_{n-1}[E']$ implies $\varphi \in \Sigma_n[E']$ and $\varphi \in \Pi_n[E']$.

ii. $\varphi \in \Sigma_n[E']$ implies $\exists x\, \varphi \in \Sigma_n[E']$ and $\forall x\, \varphi \in \Pi_{n+1}[E']$.

iii. $\varphi \in \Pi_n[E']$ implies $\exists x\, \varphi \in \Sigma_{n+1}[E']$ and $\forall x\, \varphi \in \Pi_n[E']$.

The fragments $\Sigma_n[E']$ and $\Pi_n[E']$ are closed under conjunctions and disjunctions. According to our convention to identify equivalent formulas, it makes sense to write e.g.:

$$\varphi \in \Sigma_n[E'] \;\Leftrightarrow\; \neg\varphi \in \Pi_n[E']$$

We have $\bigcup_{0 \le n} \Sigma_n[E] = \mathrm{FO}[<]$ due to the following observation:

Lemma 2.1 *We have:*

$$\Sigma_n[E] \subseteq \Sigma_n[<] \subseteq \Sigma_{n+1}[E]$$

Proof: The first inclusion follows from (1) and the second inclusion is a consequence of (2). □

For $E' = E$ and $E' = <$ we define the following language classes:

i. $\mathrm{FO}^m[E'](\mathbb{M}) = \{ L(\varphi) \mid \varphi \in \mathrm{FO}^m[E'] \}$.

ii. $\Sigma_n[E'](\mathbb{M}) = \{ L(\varphi) \mid \varphi \in \Sigma_n[E'] \}$.

iii. $\Pi_n[E'](\mathbb{M}) = \{ L(\varphi) \mid \varphi \in \Pi_n[E'] \}$.

iv. $\Delta_n[E'](\mathbb{M}) = \Sigma_n[E'](\mathbb{M}) \cap \Pi_n[E'](\mathbb{M})$.

Remark 2.1 *i. The inclusion $\Sigma_n[<] \subseteq \Sigma_{n+1}[E]$ in Lemma 2.1 is not purely syntactic, since strictly speaking it depends on the alphabet size.*

ii. The classes $\mathrm{FO}^m[E'](\mathbb{M})$ and $\Delta_n[E'](\mathbb{M})$ are Boolean algebras.

iii. The classes $\Sigma_n[E'](\mathbb{M})$ and $\Pi_n[E'](\mathbb{M})$ are closed under union and intersection.

There is a direct way to transform a first-order formula of $\varphi \in \mathrm{FO}[E]$ over $\mathbb{M}$ into a corresponding first-order formula $\varphi^* \in \mathrm{FO}[<]$ over Γ^*: We simply replace all atomic subformulas of φ of the form $(x, y) \in E$ by the conjunction $x < y \wedge (\lambda(x), \lambda(y)) \in D$. Notice that for a fixed dependence alphabet $(\lambda(x), \lambda(y)) \in D$ can be read as a macro for $\bigvee_{(a,b)\in D} (\lambda(x) = a \wedge \lambda(y) = b)$.

Lemma 2.2 *Let $L \in \Sigma_n[E](\mathbb{M}) \cap \mathrm{FO}^m[E](\mathbb{M})$. Then we have $\pi^{-1}(L) \in \Sigma_n[<](\Gamma^*) \cap \mathrm{FO}^m[<](\Gamma^*)$.*

Proof: Straightforward by using the transformation of $\varphi \in \mathrm{FO}[E]$ to $\varphi^* \in \mathrm{FO}[<]$ as above. □

3 Polynomials

The language class of polynomials $\mathrm{Pol}(\mathbb{M})$ is the smallest family of languages over $\mathbb{M}$ which contains the singletons $\{a\}$ for $a \in \Gamma$, languages A^* for $A \subseteq \Gamma$, and which is closed under finite union and concatenation. The class co-$\mathrm{Pol}(\mathbb{M})$ contains all languages $L \subseteq \mathbb{M}$ such that $\mathbb{M} \setminus L \in \mathrm{Pol}(\mathbb{M})$.

Lemma 3.1 *i. Let $a \in \Gamma$ and $A \subseteq \Gamma$. Then we have $\{a\}$ and $A^* \in \Sigma_2[E](\mathbb{M})$.*

ii. The classes $\Sigma_2[E](\mathbb{M})$ and $\Sigma_2[<](\mathbb{M})$ are closed under concatenation.

iii. We have $\mathrm{Pol}(\mathbb{M}) \subseteq \Sigma_2[E](\mathbb{M})$.

Proof: Assertion (iii.) follows from (i.) and (ii.) since $\Sigma_2[E](\mathbb{M})$ is closed under union.
(i.): The language $\{a\}$ for $a \in \Gamma$ is given by the $\Sigma_2[E]$ formula

$$\exists x \forall y \colon \lambda(x) = a \wedge x = y$$

and the language A^* for $A \subseteq \Gamma$ corresponds to

$$\forall y \colon \bigvee_{a \in A} \lambda(y) = a$$

(ii.): Let $L_0, L_1 \in \Sigma_2[E](\mathbb{M})$. We show $L_0 \cdot L_1 \in \Sigma_2[E](\mathbb{M})$. The case of $\Sigma_2[<](\mathbb{M})$ is even simpler. We construct a formula $\varphi \in \Sigma_2[E]$ expressing $L = L_0 \cdot L_1$. The construction is rather standard. The aim is to check whether there exists a factorization $t = t_0 t_1$ with $t_0 \in L_0$ and $t_1 \in L_1$. This will be done by determining a cut through t and relativization of the formulas defining L_0 and L_1. We first introduce some macro formulas:

$$\begin{aligned}
\mathrm{path}(x_1, \ldots, x_n) &\equiv \bigwedge_{1 \le i < n} (x_i = x_{i+1} \vee (x_i, x_{i+1}) \in E) \\
x \parallel y &\equiv \forall x_2 \cdots \forall x_{|\Gamma|-1} \colon \\
&\qquad \neg \mathrm{path}(x, x_2, \ldots, x_{|\Gamma|-1}, y) \wedge \neg \mathrm{path}(y, x_2, \ldots, x_{|\Gamma|-1}, x) \\
x < y &\equiv x \neq y \wedge \exists x_2 \cdots \exists x_{|\Gamma|-1} \colon \mathrm{path}(x, x_2, \ldots, x_{|\Gamma|-1}, y)
\end{aligned}$$

Note that these $\Sigma_2[E]$ macros can be viewed as $\Sigma_2[<]$ formulas. For alphabets $A_0 = \{ a_1, \ldots, a_n \}$ and $A_1 = \{ b_1, \ldots, b_m \}$ we define a $\Sigma_2[E]$ formula $\alpha_{A_0,A_1}(x_1, \ldots, x_n, y_1, \ldots, y_m)$ with $n + m$ free variables. Consider a factorization $t = t_0 t_1$. The idea is that α_{A_0,A_1} is true if each variable x_i is interpreted at the last occurrence of the letter a_i in t_0 and the alphabet of t_0 is A_0. The variables y_i represent the first positions of letters b_i of t_1 and the alphabet of t_1 is A_1. We set:

$$\begin{aligned}
&\alpha_{A_0,A_1}(x_1, \ldots, x_n, y_1, \ldots, y_m) = \\
&\bigwedge_{1 \le i \le n} \lambda(x_i) = a_i \;\wedge \bigwedge_{1 \le i \le m} \lambda(y_i) = b_i \;\wedge \bigwedge_{\substack{1 \le i \le n \\ 1 \le j \le m}} (x_i \parallel y_j \;\vee\; x_i < y_j) \;\wedge \\
&\forall z \colon \left\{ \begin{array}{l} \displaystyle\bigvee_{1 \le i \le n} \big(\lambda(z) = \lambda(x_i) \;\wedge (z = x_i \;\vee (z, x_i) \in E)\big) \;\vee \\ \displaystyle\bigvee_{1 \le j \le m} \big(\lambda(z) = \lambda(y_j) \;\wedge (y_j = z \;\vee (y_j, z) \in E)\big) \end{array} \right\}
\end{aligned}$$

Let $\overline{x} = (x_1, \ldots, x_n)$ be a tuple of variables. Suppose that we have $L_0 = L(\varphi_0)$, where $\varphi_0 = \exists \overline{y} \forall \overline{z}\, \psi_0(\overline{y}, \overline{z}) \in \Sigma_2[E]$ with $\overline{y} = (y_1, \ldots, y_m)$ and $\overline{z} = (z_1, \ldots, z_\ell)$. The restriction of φ_0 to the past of $x_1, \ldots, x_n$ is

$$\begin{aligned} \overleftarrow{\varphi_0}(\overline{x}) \;=\; & \exists y_1 \cdots \exists y_m \forall z_1 \cdots \forall z_\ell \colon \\ & \bigwedge_{1 \le j \le m} \left(\bigvee_{1 \le i \le n} \big(y_j = x_i \;\vee\; (y_j, x_i) \in E\big) \right) \;\wedge \\ & \left(\bigvee_{1 \le k \le \ell} \left(\bigwedge_{1 \le i \le n} \neg\big(z_k = x_i \;\vee\; (z_j, x_i) \in E\big) \right) \;\vee\; \psi_0(\overline{y}, \overline{z}) \right) \end{aligned}$$

Similarly, suppose L_1 is expressed by $\varphi_1 \in \Sigma_2[E]$, We can define the restriction $\overrightarrow{\varphi_1}$ of φ_1 to the future of $\overline{y} = (y_1, \ldots, y_m)$. Using these restrictions, we define

$$\varphi = \bigvee_{A_0, A_1 \subseteq \Gamma} \exists x_1 \cdots \exists x_{|A_0|}\, \exists y_1 \cdots \exists y_{|A_1|} \colon \big(\alpha_{A_0, A_1}(\overline{x}, \overline{y}) \;\wedge\; \overleftarrow{\varphi_0}(\overline{x}) \;\wedge\; \overrightarrow{\varphi_1}(\overline{y})\big)$$

Note that φ is a $\Sigma_2[E]$ formula. It expresses the language L. $\square$

As a final remark in this section we notice that polynomials are closed under homomorphic images. Therefore we can state:

Remark 3.1 *Let* $L \in \mathrm{Pol}(\Gamma^*)$. *Then we have* $\pi(L) \in \mathrm{Pol}(\mathbb{M})$.

Corollary 3.1 *We have:*

$$\mathrm{Pol}(\mathbb{M}) = \Sigma_2[E](\mathbb{M})$$

Proof: The inclusion from left to right is Lemma 3.1. For the other inclusion let $L \in \Sigma_2[E](\mathbb{M})$. Then by Lemma 2.2 we have $\pi^{-1}(L) \in \Sigma_2[<](\Gamma^*)$. Now, in the word case it follows $\pi^{-1}(L) \in \mathrm{Pol}(\Gamma^*)$, see [17]. With Remark 3.1 we conclude $L = \pi\pi^{-1}(L) \in \mathrm{Pol}(\mathbb{M})$. $\square$

4 Unary temporal logic

In [12] it is shown that $\mathrm{FO}^3[<](\mathbb{M}) = \mathrm{FO}(\mathbb{M})$. The binary atomic predicate $x < y$ can be expressed by an $\mathrm{FO}^3[E]$ formula. Let x, y, z be the sole variables. For $n = 1$ we set $E^1 = E$ and for $n > 1$ we inductively define the formula $(x, y) \in E^n$ by

$$(x, y) \in E^n \;\Leftrightarrow\; \exists z \colon \big(x = z \vee (x, z) \in E\big) \wedge (z, y) \in E^{n-1}$$

When using the hypothesis we interchange the roles of x and z in order to express $(z, y) \in E^{n-1}$. By equation (2) it follows that $x < y$ is equivalent to $(x, y) \in E^{|\Gamma|}$ and hence $\mathrm{FO}^3[E](\mathbb{M}) = \mathrm{FO}^3[<](\mathbb{M}) = \mathrm{FO}(\mathbb{M})$. Since three variables (over $<$ as well as over E) are sufficient to express all first-order properties, it is natural to consider the fragments $\mathrm{FO}^2[<]$ and $\mathrm{FO}^2[E]$. In this section we characterize them in terms of temporal logic.

Local temporal logic formulas are defined by first-order formulas having at most one free variable. In this paper we focus on unary operators, only. In temporal logic we write $a(x)$ for the atomic formula $\lambda(x) = a$. Inductively, we define $\mathsf{SF}\varphi(x)$ (S*trict* F*uture*), $\overleftarrow{\mathsf{SF}}\varphi(x)$ (*strict past*), $\mathsf{M}\varphi(x)$ (*so*M*ewhere*), $\mathsf{Eco}\,\varphi(x)$ (E*xists* co*ncurrently*) as follows.

$$\begin{aligned}
\mathsf{SF}\varphi(x) &= \exists y\colon x < y \wedge \varphi(y)\\
\overleftarrow{\mathsf{SF}}\varphi(x) &= \exists y\colon y < x \wedge \varphi(y)\\
\mathsf{M}\varphi(x) &= \exists y\colon \varphi(y)\\
\mathsf{Eco}\,\varphi(x) &= \exists y\colon x \parallel y \wedge \varphi(y)
\end{aligned}$$

It is common to write φ instead of $\varphi(x)$. We define various macros:

$$\begin{aligned}
\mathsf{F}\,\varphi &= \varphi \vee \mathsf{SF}\varphi && \mathsf{F}\textit{uture}\\
\mathsf{G}\,\varphi &= \neg\,\mathsf{F}\,\neg\varphi && \mathsf{G}\textit{lobally (in the future)}\\
\mathsf{SG}\varphi &= \neg\mathsf{SF}\neg\varphi && \mathsf{S}\textit{trict } \mathsf{G}\textit{lobally (in the future)}\\
\overleftarrow{\mathsf{SG}}\varphi &= \neg\overleftarrow{\mathsf{SF}}\neg\varphi && \mathsf{S}\textit{trict } \mathsf{G}\textit{lobally (in the past)}\\
\mathsf{Aco}\,\varphi &= \neg\mathsf{Eco}\,\neg\varphi && \mathsf{A}\textit{ll } \mathsf{co}\textit{ncurrently}\\
\mathsf{Ev}\varphi &= \neg\mathsf{M}\neg\varphi && \mathsf{Ev}\textit{erywhere}
\end{aligned}$$

The *depth* (or *operator depth*) of a formula $\varphi \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathbb{M}, \mathsf{Eco}\,]$ is the maximal number of nested temporal operators occurring within φ (cf. [7]). Let $\mathcal{C}$ be a subset of temporal operators from the set above, then $\mathrm{TL}[\mathcal{C}]$ means the formulas where all operators are from $\mathcal{C}$. Clearly, if $\mathcal{C}$ consists of unary operators as above, then

$$\mathrm{TL}[\mathcal{C}] \subseteq \left\{\, \varphi \in \mathrm{FO}^2[<] \;\middle|\; \varphi \text{ has at most one free variable} \,\right\}$$

This can be made more precise and reveals a difference between $\mathrm{FO}^2[E]$ and $\mathrm{FO}^2[<]$ for traces. As we will see later, the fragment $\mathrm{FO}^2[<]$ is more powerful than $\mathrm{FO}^2[E]$, in general.

Lemma 4.1 *We have:*

i. $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{Eco}\,] = \left\{\, \varphi \in \mathrm{FO}^2[<] \;\middle|\; \varphi \textit{ has at most one free variable} \,\right\}$

ii. $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}] = \left\{\, \varphi \in \mathrm{FO}^2[E] \;\middle|\; \varphi \textit{ has at most one free variable} \,\right\}$

Proof: The inclusion from left to right in (i.) holds by definition of the temporal operators. For the analogous inclusion in (ii.) we can use equation (2). For example, $\mathsf{SF}\psi$ corresponds to the FO^2-formula $\varphi(x) \equiv \exists y : x < y \wedge \psi(y)$ where by induction on the number of operators $\psi \in \mathrm{FO}^2[E]$. Now, $\varphi(x)$ is equivalent to $\widehat{\varphi^{|\Gamma|}}(x) \in \mathrm{FO}^2[E]$ defined by

$$\begin{aligned}
\widehat{\varphi^{1}}(x) &\equiv \exists y\colon (x,y) \in E \,\wedge\, \psi(y)\\
\widehat{\varphi^{i+1}}(x) &\equiv \exists y\colon \big(x = y \vee (x,y) \in E\big) \,\wedge\, \widehat{\varphi^{i}}(y)
\end{aligned}$$

where in $\widehat{\varphi^{i}}(y)$ the roles of x and y are interchanged. Note that this step would not be possible if we had replaced $\psi(y)$ by some formula $\psi(x,y)$ that also depends on x. The proof of the inclusions from right to left follows the same lines as it has been shown in word case by Etessami-Vardi-Wilke [7]. To keep this paper self-contained we give a complete proof for this direction. Let $\varphi(x) \in \mathrm{FO}^2[<]$ be a formula with one free variable. We show by induction on the quantifier depth and the size of the formula that there exists $\widetilde{\varphi} \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{E}\infty]$ such that

$$t, x \models \varphi(x) \;\Leftrightarrow\; t, x \models \widetilde{\varphi}$$

Atomic formulas and Boolean operators are translated as follows:

$$\begin{aligned} \varphi(x) &\equiv\; a(x) \;\text{ for } a \in \Gamma & &\rightsquigarrow & \widetilde{\varphi} &\equiv a \\ \varphi(x) &\equiv\; \psi_1(x) \vee \psi_2(x) & &\rightsquigarrow & \widetilde{\varphi} &\equiv \widetilde{\psi_1} \vee \widetilde{\psi_2} \\ \varphi(x) &\equiv\; \neg\psi(x) & &\rightsquigarrow & \widetilde{\varphi} &\equiv \neg\widetilde{\psi} \end{aligned}$$

If $\varphi(x)$ is of the form $\exists x \colon \psi(x)$, in an intermediate step it is transformed into $\varphi'(x) \equiv \exists y \colon \psi(y)$ by interchanging x and y. The formula $\varphi(x) \equiv \exists y \colon \psi(y)$ can be interpreted as $\varphi(x) \equiv \exists y \colon \psi(x, y)$ where x is a dummy variable. We now consider the general case $\varphi(x) \equiv \exists y \colon \psi(x, y)$. First, we transform $\varphi(x)$ into an equivalent formula $\varphi''(x)$ of the same quantifier depth. Let

$$\psi(x, y) \;\equiv\; \beta\big(\, x = y, x < y, y < x, \xi_1(x), \ldots, \xi_n(x), \zeta_1(y), \ldots, \zeta_m(y)\,\big)$$

where β is a propositional formula and $\xi_i(x)$, $\zeta_j(y)$ are atomic formulas or existential formulas with smaller quantifier depth. The first step in the transformation of $\varphi(x)$ is to *guess* the values of $\xi_i(x)$ before the quantification of y. We set $\varphi'(x) \equiv$

$$\bigvee_{\overline{\gamma} \in \{\top, \bot\}^n} \Bigg(\bigwedge_{1 \le i \le n} \big(\xi_i(x) \leftrightarrow \gamma_i\big) \;\wedge\; \exists y \colon \\ \beta\big(x = y, x < y, y < x, \overline{\gamma}, \zeta_1(y), \ldots, \zeta_m(y)\,\big)\Bigg)$$

The next step is to *guess* the order-type τ that holds between x and y in advance. The possible relations are $x = y$, $x < y$, $x > y$ or (other than in the word case) none of them and then x and y correspond to parallel positions, $x \parallel y$. We choose τ from the set $\{=, <, >, \parallel\}$ and define $\varphi''(x) \equiv$

$$\bigvee_{\overline{\gamma} \in \{\top, \bot\}^n} \Bigg(\bigwedge_{1 \le i \le n} (\xi_i(x) \leftrightarrow \gamma_i) \;\wedge \bigvee_{\tau \in \{=,<,>,\parallel\}} \exists y \colon (x \,\tau\, y \,\wedge\, \beta(y)) \Bigg)$$

with $\beta(y) = \beta((x = y)^\tau, (x < y)^\tau, (y < x)^\tau, \overline{\gamma}, \zeta_1(y), \ldots, \zeta_m(y)\,)$. Note that the first $3 + n$ arguments are constant Boolean values at this point. After this transformation of $\varphi(x)$ it remains to show how to translate formulas of the form $\exists y \colon (x \,\tau\, y \wedge \beta(y))$ with $\tau \in \{=, <, >, \parallel\}$:

$$\begin{aligned} \exists y \colon (x = y \wedge \beta(y)) &\quad\rightsquigarrow\quad \widetilde{\beta} \\ \exists y \colon (x < y \wedge \beta(y)) &\quad\rightsquigarrow\quad \mathsf{SF}\widetilde{\beta} \\ \exists y \colon (x > y \wedge \beta(y)) &\quad\rightsquigarrow\quad \overleftarrow{\mathsf{SF}}\widetilde{\beta} \\ \exists y \colon (x \parallel y \wedge \beta(y)) &\quad\rightsquigarrow\quad \mathsf{E}\infty\,\widetilde{\beta} \end{aligned}$$

Note that φ'' is at most exponentially bigger than φ. Therefore, the size of $\widetilde{\varphi}$ is (poly-)exponentially bounded by the size of φ. The transformation of

$\varphi(x) \in \mathrm{FO}^2[E]$ into an equivalent formula $\widetilde{\varphi} \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}]$ is similar. The difference arises only in the set of order types and in the last step. Note that

$$x \neq y \ \wedge (x,y) \notin E \ \wedge (y,x) \notin E \ \Leftrightarrow \ \big(\lambda(x), \lambda(y)\big) \in I$$

Therefore, the order-type $x \parallel y$ in $\mathrm{FO}^2[<]$ corresponds to independence of the labels of x and y in the $\mathrm{FO}^2[E]$ setting and we get the following transformations

$$\begin{aligned}
\exists y \colon \big(\, (x,y) \in E \ \wedge \beta(y)\,\big) &\quad\rightsquigarrow\quad \bigvee_{(a,b)\in D} a \wedge \mathsf{SF}(b \wedge \widetilde{\beta}) \\
\exists y \colon \big(\, (y,x) \in E \ \wedge \beta(y)\,\big) &\quad\rightsquigarrow\quad \bigvee_{(a,b)\in D} a \wedge \overleftarrow{\mathsf{SF}}(b \wedge \widetilde{\beta}) \\
\exists y \colon \big(\, (\lambda(x),\lambda(y)) \in I \ \wedge \beta(y)\,\big) &\quad\rightsquigarrow\quad \bigvee_{(a,b)\in I} a \wedge \mathsf{M}(b \wedge \widetilde{\beta})
\end{aligned}$$

This completes the proof of the inclusions from right to left. $\square$

In order to pass form formulas in temporal logic to languages we would like to define $L(\varphi) \subseteq \mathbb{M}$, even if φ has a free variable. There is however no canonical choice. We use an existential variant; and we define here:

$$L_\exists(\varphi) = \{\, t \in \mathbb{M} \ \mid \exists x \in t \colon t, x \models \varphi \,\} = L(\mathsf{M}\varphi)$$

Clearly, $L_\exists(\mathsf{M}\varphi) = L_\exists(\varphi) = L(\mathsf{M}\varphi)$. Define $\mathrm{TL}[\mathcal{C}](\mathbb{M})$ as the Boolean closure of languages defined by $L_\exists(\varphi)$ with $\varphi \in \mathrm{TL}[\mathcal{C}]$. As a consequence of the existential choice we never have $1 \in L_\exists(\varphi)$. But this will be no obstacle in the following, because $L_\exists(\top) = \mathbb{M} \setminus \{1\}$ implies that $L \in \mathrm{TL}[\mathcal{C}](\mathbb{M})$ if and only if $L \setminus \{1\} \in \mathrm{TL}[\mathcal{C}](\mathbb{M})$. In the following lemma, we show that the operator M does not add any expressiveness beyond Boolean operations.

Lemma 4.2 *We have:*

$$\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}](\mathbb{M}) = \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}](\mathbb{M})$$

Proof: The inclusion from left to right is trivial. For the other direction, let $\varphi \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}]$ and $\mathsf{M}\psi$ be a subformula of φ. For $\tau \in \{\top, \bot\}$ let $\varphi[\mathsf{M}\psi \leftarrow \tau]$ be the formula where $\mathsf{M}\psi$ is replaced by τ. Now,

$$\varphi = \mathsf{M}\psi \wedge \varphi[\mathsf{M}\psi \leftarrow \top] \ \vee \ \neg\mathsf{M}\psi \wedge \varphi[\mathsf{M}\psi \leftarrow \bot]$$

because $\mathsf{M}\psi$ has no free variables. The result follows by induction since $L(\mathsf{M}\psi) = L_\exists(\psi)$ and $\varphi[\mathsf{M}\psi \leftarrow \tau]$ has one M less, and $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}](\mathbb{M})$ is by definition a Boolean algebra. $\square$

Essentially the converse holds as well, up to the empty trace, Boolean operations can be replaced by the use of the operator M. This fact is stated in the following lemma in order to complete the picture, but not used henceforth.

Lemma 4.3 $\{\, L_\exists(\varphi) \ \mid \varphi \in \mathrm{TL}[\mathcal{C} \cup \{\mathsf{M}\}] \,\}$ *is a Boolean algebra with respect to* $\mathbb{M} \setminus \{1\}$.

Proof: We have $\mathbb{M} \setminus \big(L_\exists(\varphi) \cup \{1\}\,\big) = L_\exists(\neg\mathsf{M}\varphi)$. $\square$

Remark 4.1 *If $\mathbb{M}$ is a free monoid, then* $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}] = \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}]$*, because* $\mathsf{M}\varphi = \varphi \vee \mathsf{SF}\varphi \vee \overleftarrow{\mathsf{SF}}\varphi$ *in this case. Hence* $\left\{ L_{\exists}(\varphi) \;\middle|\; \varphi \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}] \right\}$ *forms a Boolean algebra with respect to* Γ^+*.*

5 Characterizing DA over traces

Various characterizations of $\mathbf{DA}(\Gamma^*)$ are known [19]. In the following we extend some of them to traces. In particular, we are interested in the corresponding first-order fragments and its temporal logic counterpart. The following is one of the main results of the paper. The crucial step is $\mathbf{DA}(\mathbb{M}) \subseteq \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}](\mathbb{M})$. This has been established first in the Ph.D. thesis [14]. In Lemma 5.3 we give a new and self-contained proof of this fact. This will require some arithmetic in finite monoids and is postponed to the end of this section.

Theorem 5.1 *We have:*

$$\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}](\mathbb{M}) = \mathrm{FO}^2[E](\mathbb{M}) = \mathrm{Pol}(\mathbb{M}) \cap \text{co-Pol}(\mathbb{M}) = \Delta_2[E](\mathbb{M}) = \mathbf{DA}(\mathbb{M})$$

Proof: For words, i.e., for $\mathbb{M} = \Gamma^*$, the result follows by [7, 17, 20]. For arbitrary trace monoids we have

$L \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}](\mathbb{M})$	
$\Rightarrow L \in \mathrm{FO}^2[E](\mathbb{M})$	(Lemma 4.1 and 4.2)
$\Rightarrow \pi^{-1}(L) \in \mathrm{FO}^2[<](\Gamma^*)$	(Lemma 2.2)
$\Rightarrow \pi^{-1}(L) \in \mathrm{Pol}(\Gamma^*) \cap \text{co-Pol}(\Gamma^*)$	(by word case)
$\Rightarrow L \in \mathrm{Pol}(\mathbb{M}) \cap \text{co-Pol}(\mathbb{M})$	(Remark 3.1)
$\Rightarrow L \in \Delta_2[E](\mathbb{M})$	(Lemma 3.1)
$\Rightarrow \pi^{-1}(L) \in \Delta_2[<](\Gamma^*)$	(Lemma 2.2)
$\Rightarrow \pi^{-1}(L) \in \mathbf{DA}(\Gamma^*)$	(by word case)
$\Rightarrow L \in \mathbf{DA}(\mathbb{M})$	(by equation (3))
$\Rightarrow L \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}](\mathbb{M})$	(by [14, 15] or Lemma 5.3)

$\square$

Note that membership in the variety **DA** is decidable and hence the membership problem for all of the above characterizations is decidable. The theorem above characterizes $\mathrm{FO}^2[E](\mathbb{M})$. However for words there is no difference between E and the relation $<$. In particular, $\mathrm{FO}^2[E](\Gamma^*) = \mathrm{FO}^2[<](\Gamma^*)$. Can we hope that this is always true? The answer is *no*:

Theorem 5.2 *The following assertions are equivalent:*

i. *D is transitive, i.e., $\mathbb{M}$ is a direct product of free monoids.*

ii. $\mathrm{FO}^2[E](\mathbb{M}) = \mathrm{FO}^2[<](\mathbb{M})$.

iii. $\Delta_2[E](\mathbb{M}) = \Delta_2[<](\mathbb{M})$.

Proof: If D is transitive, then E and $<$ are identical. Now, suppose D is not transitive. Then there exist three letters $a, b, c \in \Gamma$ with $(a, b), (b, c) \in D$ and $(a, c) \notin D$. Consider the trace language $L = L(\varphi)$ with $\varphi = \exists x \exists y \colon x \parallel y \in \mathrm{FO}^2[<] \cap \Sigma_1[<]$. The language L contains all traces with two parallel positions. We have $L \in \mathrm{FO}^2[<](\mathbb{M}) \cap \Delta_2[<](\mathbb{M})$.

Assume by contradiction, $L \in \mathbf{DA}(\mathbb{M})$. Then there exists a homomorphism $h : \mathbb{M} \to M$ onto a finite monoid $M \in \mathbf{DA}$ such that $h^{-1}h(L) = L$. Since $M \in \mathbf{DA}$ there exists $n \in \mathbb{N}$ with $\forall u, v, w \in M \colon (uvw)^n v (uvw)^n = (uvw)^n (uvw)^n$. With $u = h(b)$, $v = h(abc)$ and $w = 1$ we conclude $(babc)^n abc (babc)^n \in L$ if and only if $(babc)^n (babc)^n \in L$. This is a contradiction, since $(babc)^n (babc)^n \notin L$ is a sequence without any parallel positions, whereas $(babc)^n abc (babc)^n \in L$ contains two parallel positions labeled by c and a. Therefore $L \notin \mathbf{DA}(\mathbb{M})$. By Theorem 5.1 we have $\mathrm{FO}^2[E](\mathbb{M}) = \Delta_2[E](\mathbb{M}) = \mathbf{DA}(\mathbb{M})$. It follows $L \in \mathrm{FO}^2[<](\mathbb{M}) \setminus \mathrm{FO}^2[E](\mathbb{M})$ and $L \in \Delta_2[<](\mathbb{M}) \setminus \Delta_2[E](\mathbb{M})$. $\square$

Remark 5.1 *In general, it is open whether membership is decidable for classes like* $\mathrm{FO}^2[<](\mathbb{M})$ *or* $\Delta_2[<](\mathbb{M})$. *Theorems 5.1 and 5.2 imply that membership for* $\mathrm{FO}^2[<](\mathbb{M})$ *and for* $\Delta_2[<](\mathbb{M})$ *is decidable if D is transitive.*

In Lemma 5.1 and Lemma 5.2 we give some properties of the variety **DA**. *Green's relations* are one possibility to formulate those properties. We introduce only those Green's relations that will be used in the remainder of this section. Let M be a monoid and $u, v \in M$. We define

$$\begin{array}{lll} u \mathrel{\mathcal{L}} v \iff Mu = Mv & u \leq_{\mathcal{L}} v \iff Mu \subseteq Mv & u <_{\mathcal{L}} v \iff Mu \subsetneq Mv \\ u \mathrel{\mathcal{R}} v \iff uM = vM & u \leq_{\mathcal{R}} v \iff uM \subseteq vM & u <_{\mathcal{R}} v \iff uM \subsetneq vM \end{array}$$

Note that $\mathcal{R}$ and $\mathcal{L}$ are equivalence relations, whereas $\leq_{\mathcal{R}}$ and $\leq_{\mathcal{L}}$ are preorders. For every product $u_\ell u_{\ell-1} \cdots u_1 u_0$ we have $1_M \geq_{\mathcal{L}} u_0 \geq_{\mathcal{L}} u_1 u_0 \geq_{\mathcal{L}} u_2 u_1 u_0 \geq_{\mathcal{L}} \cdots \geq_{\mathcal{L}} u_\ell \cdots u_0$. For every monoid $M \in \mathbf{DA}$ there exists $n \in \mathbb{N}$ such $(uvw)^n v (uvw)^n = (uvw)^{2n}$ for all $u, v, w \in M$. In particular, $v^{2n+1} = v^{2n}$ for all $v \in M$ which shows that every monoid in **DA** is aperiodic.

Lemma 5.1 *Let M be an aperiodic monoid and let $u, v \in M$. If $u \leq_{\mathcal{L}} v$ and $v \leq_{\mathcal{R}} u$ then $u = v$.*

Proof: Choose $m \in \mathbb{N}$ such that $y^m = y^{m+1}$ for all $y \in M$. Since $u \in Mu \subseteq Mv$ there exists $x \in M$ such that $u = xv$. Similarly, there exists $y \in M$ such that $v = uy$. We have $u = xv = xuy = x^m u y^m = x^m u y^{m+1} = uy = v$. $\square$

The crucial properties for monoids in **DA** are aperiodicity (as used in the lemma just above) and the property as given in the next lemma. Of course, there is also a symmetric statement using Green's relation $\mathcal{R}$.

Lemma 5.2 *Let $u, v, a \in M \in \mathbf{DA}$. If $u \mathrel{\mathcal{L}} vu$ and $v \in MaM$ then $u \mathrel{\mathcal{L}} avu$.*

Proof: Choose $n \in \mathbb{N}$ such $(uvw)^n v (uvw)^n = (uvw)^{2n}$ for all $u, v, w \in M$. We have $wu \leq_{\mathcal{L}} u$ for all $u, w \in M$. Therefore, it suffices to show $u \leq_{\mathcal{L}} avu$. Let

$x, y, z \in M$ such that $v = xay$ and $u = zvu$. Then

$$\begin{aligned} yu &= yzvu = yzxa \cdot yu \\ &= (yzxa)^n \cdot yu \\ &= (yzxa)^{2n} \cdot yu \\ &= (yzxa)^n xa (yzxa)^n \cdot yu && \text{since } M \in \mathbf{DA} \\ &= (yzxa)^n \cdot xa \cdot yu \in Mavu \end{aligned}$$

Therefore $u = zxa \cdot vu \in zxa \cdot Mavu$ and hence $Mu \subseteq Mavu$. $\square$

Lemma 5.3 *If $L \in \mathbf{DA}(\mathbb{M})$ then $L \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}](\mathbb{M})$.*

Proof: Let $M \in \mathbf{DA}$ and let $h : \mathbb{M} \to M$ be a homomorphism with $h^{-1}h(L) = L$. We define an equivalence on traces, called *operator-depth-equivalence*, by $\mathrm{OD}_n(u) = \mathrm{OD}_n(v)$ if u and v agree on all formulas of operator depth at most n. Let $n > 2\,|M| \cdot |\Gamma|$ and $u, v \in \mathbb{M}$ with $\mathrm{OD}_n(u) = \mathrm{OD}_n(v)$. We show that this implies $h(u) = h(v)$. From $n \geq 1$ we conclude that $\mathrm{alph}(u) = \mathrm{alph}(v)$. If $\mathrm{alph}(u) = \emptyset$ then $u = 1 = v$ and hence $h(u) = h(v)$. Thus we may assume $\mathrm{alph}(u) \neq \emptyset$ and we perform an induction on the size of this alphabet. For some $\ell < |M|$ we can write $u = u_\ell C_\ell \cdots u_1 C_1 u_0$ where $C_i = \max(u_i C_i)$ and for all choices of letters $a_i \in C_i$ we have:

$$\begin{aligned} 1_M \;\mathcal{L}\; h(u_0) &>_{\mathcal{L}} h(a_1 u_0) \\ &\geq_{\mathcal{L}} h(C_1 u_0) \;\mathcal{L}\; h(u_1 C_1 u_0) >_{\mathcal{L}} \cdots \\ &>_{\mathcal{L}} h(a_\ell u_{\ell-1} \cdots u_1 C_1 u_0) \\ &\geq_{\mathcal{L}} h(C_\ell u_{\ell-1} \cdots u_1 C_1 u_0) \;\mathcal{L}\; h(u_\ell C_\ell u_{\ell-1} \cdots u_1 C_1 u_0) \end{aligned}$$

A suitable factorization $u = u_\ell C_\ell \cdots u_1 C_1 u_0$ can be found as follows. We choose u_0 of maximal length such that $1_M \;\mathcal{L}\; h(u_0)$, then C_1 is the product over the maximal letters in the remaining prefix before u_0; and this leads to the general procedure: Once C_i is defined, choose u_i of maximal length and let C_{i+1} be the maximal letters in the remaining prefix before u_i. The effect is that every $a_i \in C_i$ reduces the level of the $\mathcal{L}$-class. In general, the above factorization is not unique.

By Lemma 5.2 we see $u_{i-1} \notin M a_i M$ for all $a_i \in C_i$ which implies $C_i \cap \mathrm{alph}(u_{i-1}) = \emptyset$. We claim that for each $a \in C_i$ there is a formula $\varphi_{a,i} \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}]$ of operator depth $2i - 1$ such that $u, x \models \varphi_{a,i}$ if and only if x corresponds to the position of the letter a in the factor C_i. For $i = 1$ this is

$$\varphi_{a,1} = a \wedge \neg \mathsf{SF} a$$

and for $i > 1$ we can use induction

$$\varphi_{a,i} \;=\; \left(a \;\wedge \bigvee_{b \in C_{i-1}} \mathsf{SF} \varphi_{b,i-1} \right) \;\wedge\; \neg \mathsf{SF} \left(a \wedge \bigvee_{b \in C_{i-1}} \mathsf{SF} \varphi_{b,i-1} \right)$$

The formula $\varphi_{a,i}$ expresses that the position x has label a and is before some position in C_{i-1} and it is the last position with this property. Note that the operator depth of $\varphi_{a,i}$ is indeed $2i-1$. Using the (syntactical) conventions $\bigvee_{a\in C_0} \mathsf{SF}\varphi_{a,0} = \neg \bigvee_{a\in C_{\ell+1}} \mathsf{SF}\varphi_{a,t+1} = \top$ we can specify the positions in u_i by

$$\psi_i \;=\; \left(\bigvee_{a\in C_i} \mathsf{SF}\varphi_{a,i}\right) \wedge \neg \left(\bigvee_{a\in C_{i+1}} \mathsf{SF}\varphi_{a,i+1}\right)$$

Note that a formula of operator depth of at most 2ℓ can specify that for each position exactly one of the formulas $\varphi_{a,i}$ or ψ_i holds. Since $n \geq 2\ell$ and $\mathrm{OD}_n(u) = \mathrm{OD}_n(v)$ we can factorize $v = v_\ell C_\ell \cdots v_1 C_1 v_0$ where the position of each C_i consists of the maximal letters in $v_i C_i$. By using ψ_i for relativizations we see that $\mathrm{OD}_{n-2\ell}(u_i) = \mathrm{OD}_{n-2\ell}(v_i)$ for $0 \leq i \leq \ell$. Furthermore $\mathrm{alph}(u_i) \subseteq \mathrm{alph}(u) \setminus C_{i+1}$ for $0 \leq i < \ell$. By induction on the alphabet size we obtain $h(u_i) = h(v_i)$ for $0 \leq i < \ell$ (we cannot use the induction hypothesis for $i = \ell$ since we may have $\mathrm{alph}(u_\ell) = \Gamma$). Thus

$$\begin{aligned} h(v) &\leq_{\mathcal{L}} h(C_\ell v_{\ell-1} \cdots C_1 v_0) \\ &= h(C_\ell u_{\ell-1} \cdots C_1 u_0) \;\mathcal{L}\; h(u_\ell C_\ell u_{\ell-1} \cdots C_1 u_0) \\ &= h(u) \end{aligned}$$

This means $h(v) \leq_{\mathcal{L}} h(u)$. Symmetrically, by starting with a factorization of v with respect to Green's $\mathcal{R}$-relation we see that $h(u) \leq_{\mathcal{R}} h(v)$. From Lemma 5.1 we conclude $h(u) = h(v)$. Up to equivalence there are only finitely many formulas of operator depth n. By specifying which of them hold and which of them do not hold we see that for all $m \in M$ the language $h^{-1}(m)$ can be expressed by a $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}]$ formula. The lemma now follows since $L = \bigcup_{m\in h(L)} h^{-1}(m)$. $\square$

6 Complexity

In this section we show that the possibility to express concurrency also affects the complexity of the satisfiability problem.

Theorem 6.1 *The following problem is* NP*-complete:*

Input: A dependence alphabet (Γ, D) *and a formula* $\varphi \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}]$.

Question: Is φ *satisfiable over* $\mathbb{M} = \mathbb{M}(\Gamma, D)$*?*

Proof: The hardness follows from the special case that satisfiability for $\mathrm{TL}[\mathsf{F}]$ is already NP-hard for words [18]. Therefore we show inclusion in NP, only. First, we transform $\varphi \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}]$ by allowing additional temporal operators into some

$$\Phi \in \mathrm{TL}^+[\mathsf{SF}, \mathsf{SG}, \overleftarrow{\mathsf{SF}}, \overleftarrow{\mathsf{SG}}, \mathsf{M}, \mathsf{Ev}]$$

where TL^+ means positive Boolean combinations and the only negations are of the form $\neg a$ for $a \in \Gamma$. Let $|\Phi|$ be the size of Φ. We observe $|\Phi| \leq 2\,|\varphi|$. We show that $\mathrm{TL}^+[\mathsf{SF}, \mathsf{SG}, \overleftarrow{\mathsf{SF}}, \overleftarrow{\mathsf{SG}}, \mathsf{M}, \mathsf{Ev}]$ has a so-called *small model property*: if there exists a trace $t \in \mathbb{M}$ and a position $x \in t$ such that $t, x \models \Phi$ then there also exists a *small* trace $s \in \mathbb{M}$ (of polynomial size) and a position $y \in s$ such that $s, y \models \Phi$ and $|s| \leq 2\,|\Phi| \cdot |\Gamma|$.

Suppose $t, x \models \Phi$. By ψ we mean a subformula of Φ. For each ψ and each letter $a \in \Gamma$ such that there exists $y \in t$ with $t, y \models \psi \wedge a$, we mark a left-most position $\ell(\psi, a) \in t$ and a right-most position $r(\psi, a) \in t$ such that we have both, $t, \ell(\psi, a) \models \psi \wedge a$ and $t, r(\psi, a) \models \psi \wedge a$. Let s be the subtrace of t that consists of all marked positions. We claim that for all $x \in s$ and all ψ we have

$$t, x \models \psi \;\Rightarrow\; s, x \models \psi$$

This holds for $\psi = a$ and $\psi = \neg a$ and positive Boolean connectives. By left-right symmetry, it is enough to consider subformulas of the form $\mathsf{SF}\rho$, $\mathsf{SG}\rho$, $\mathsf{M}\rho$ and $\mathsf{Ev}\rho$. The "for all" situations $\mathsf{SG}\rho$ and $\mathsf{Ev}\rho$ are trivial (because we have no $\mathsf{A}\infty\,\rho$). It remain $\mathsf{SF}\rho$ and $\mathsf{M}\rho$. Now, if $t, x \models \mathsf{M}\rho$ holds then $t, y \models \rho$ for some $y \in t$. Hence $t, \ell(\rho, \lambda(y)) \models \rho$. But $\ell(\rho, \lambda(y))$ is marked, hence $\ell(\rho, \lambda(y)) \in s$ and by induction, $s, \ell(\rho, \lambda(y)) \models \rho$. This implies $s, x \models \mathsf{M}\rho$.

Finally, suppose $t, x \models \mathsf{SF}\rho$. Let $\lambda(x) = a_1$ and $x_1 = r(\mathsf{SF}\rho, a_1)$. For some $d > 1$ there are marked positions $x_1, \ldots, x_d \in t$ with $(x_i, x_{i+1}) \in E$ and $x_i = r(\mathsf{SF}\rho, a_i)$ for $1 \le i < d$ and with $x_d = r(\rho, a_d)$. By structural induction, we have $s, x_2 \models \rho$ for $d = 2$ and, by induction on d, we have $s, x_2 \models \mathsf{SF}\rho$ for $d > 2$. In any case, since $\lambda(x) = a_1$ and $(a_1, \lambda(x_2)) \in D$ we obtain $s, x \models \mathsf{SF}\rho$.

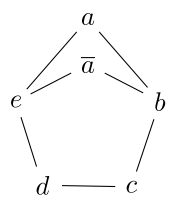
Thus, if φ is satisfiable we can guess $s \in \mathbb{M}$ with $|s| \le 2\,|\Phi| \cdot |\Gamma|$. After that we check in deterministic polynomial time whether $s, x \models \varphi$ for some $x \in s$. $\square$

A class of languages C of $\mathbb{M}$ is called *stutter-invariant* if for every language $L \in C$ we have $t_1 a a t_2 \in L$ if and only if $t_1 a t_2 \in L$ for all traces $t_1, t_2 \in \mathbb{M}$ and every letter $a \in \Gamma$. For marked traces (t, x) we extend this notion by requiring $(t_1 a a t_2, x) \in L$ if and only if $(t_1 a t_2, f(x)) \in L$. The function f maps canonically positions of $t_1 a a t_2$ onto positions of $t_1 a t_2$. Positions of the prefix t_1 and the suffix t_2 are mapped to their counter-parts in $t_1 a t_2$ and the two positions factor aa in the above factorization are both mapped to the position of a in the factorization $t_1 a t_2$. A trace t is called *stutter-free* if there is no factor aa for any $a \in \Gamma$.

Lemma 6.1 $\mathrm{TL}[\mathsf{F}, \overleftarrow{\mathsf{F}}, \mathsf{E}\infty](\mathbb{M})$ *is stutter-invariant.*

Proof: Straightforward by induction. $\square$

Let $\Gamma = \{\, a, \overline{a}, b, c, d, e \,\}$ and let the independence relation I be given by the graph

$I \;=$ 

Theorem 6.2 *Let* $\Gamma = \{\, a, \overline{a}, b, c, d, e \,\}$ *and* $D = \Gamma \times \Gamma \setminus I$ *where* I *is the above independence relation. The following problem is* PSPACE*-hard:*

Input: A formula $\varphi \in \mathrm{TL}[\mathsf{F}, \mathsf{E}\infty]$ *over* Γ.

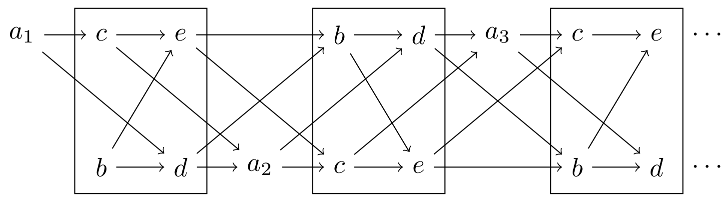
Question: Is φ *satisfiable over* $\mathbb{M} = \mathbb{M}(\Gamma, D)$*?*

Proof: We use a reduction from the following problem which is well-known to be PSPACE-complete, see e.g. [6].

> Input: A formula $\varphi \in \mathrm{TL}[\mathsf{X}, \mathsf{F}]$ over $\{a, \overline{a}\}$.
> Question: Does there exist a word $w \in \{a, \overline{a}\}^+$ with $w, 1 \models \varphi$? Here 1 denotes the first position of w.

For $w = a_1 \cdots a_n \in \{a, \overline{a}\}^+$ we define $\widetilde{w} = a_1(bcde)a_2(bcde) \cdots a_n(bcde) \in \mathbb{M}$ and for a position $i \in \{ 1, \ldots, n \}$ of w we define x_i as the position of a_i in $\widetilde{w}$.

$\widetilde{w} =$



In a first step, we inductively transform φ into a formula $\widetilde{\varphi} \in \mathrm{TL}[\mathsf{F}, \mathsf{Eco}]$ such that

$$w, i \models \varphi \;\Leftrightarrow\; \widetilde{w}, x_i \models \widetilde{\varphi} \tag{4}$$

Boolean combinations and atomic formulas are straightforward: $\widetilde{a} = a$,$\widetilde{\neg\varphi} = \neg\widetilde{\varphi}$, and $\widetilde{\varphi \vee \psi} = \widetilde{\varphi} \vee \widetilde{\psi}$.

$$\widetilde{\mathsf{F}\,\psi} = \mathsf{F}((a \vee \overline{a}) \wedge \widetilde{\psi})$$
$$\widetilde{\mathsf{X}\,\psi} = \mathsf{Eco}_{\mathsf{X}}((a \vee \overline{a}) \wedge \widetilde{\psi})$$

where $\mathsf{Eco}_{\mathsf{X}}$ is defined by $\mathsf{Eco}_{\mathsf{X}}\psi = \mathsf{Eco}\,(b \wedge \mathsf{Eco}\,(c \wedge \mathsf{Eco}\,(d \wedge \mathsf{Eco}\,(e \wedge \mathsf{Eco}\,\psi))))$. By induction, the equivalence (4) holds for all $w = a_1 \cdots a_n \in \{a, \overline{a}\}^+$ and all positions $i \in \{ 1, \ldots, n \}$. In the remainder of the proof we construct a formula ψ such that φ is satisfiable over $\{a, \overline{a}\}^+$ if and only if $\widetilde{\varphi} \wedge \psi$ is satisfiable over $\mathbb{M}$. The formula will ensure some normal form on its models. This normal form is given by the following set of marked traces.

$$K = \left\{ (t, x) \;\middle|\; \begin{array}{c} t = pq \text{ with } \max(p) \subseteq \{d\} \text{ and} \\ q \in ((a^+ \cup \overline{a}^+)b^+c^+d^+e^+)^+ \quad \text{and} \\ x \text{ is a minimal position of } q \\ \text{labeled with either } a \text{ or } \overline{a} \end{array} \right\}$$

We define $\psi \in \mathrm{TL}[\mathsf{F}, \mathsf{Eco}]$ with

$$t, x \models \psi \;\Leftrightarrow\; (t, x) \in K$$

We note that if $(t, x) \in K$ then the factorization $t = pq$ with $\max(p) \subseteq \{d\}$ and $q \in ((a^+ \cup \overline{a}^+)b^+c^+d^+e^+)^+$ is unique since in q every vertex labeled with d is parallel to some c, but $c \notin \max(p) \cup \min(q)$. Since a and $\overline{a}$ have the same behavior we use $\widehat{a}$ as a shorthand for either a or for $\overline{a}$. We define $\psi =$

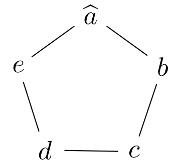
$\psi_{\mathsf{Eco}} \wedge \psi_a \wedge \psi_b \wedge \psi_c \wedge \psi_d \wedge \psi_e \wedge \psi_{\text{start}}$ with

$$
\begin{aligned}
\widehat{a} &= a \vee \overline{a} \\
\psi_{\mathsf{Eco}} &= \mathsf{G}(\neg \mathsf{Eco}\, a \vee \neg \mathsf{Eco}\, \overline{a}) \\
\psi_a &= \mathsf{G}(\widehat{a} \rightarrow \mathsf{Eco}\, b) \\
\psi_b &= \mathsf{G}(b \rightarrow \mathsf{Eco}\, c) \\
\psi_c &= \mathsf{G}(c \rightarrow \mathsf{Eco}\, d) \\
\psi_d &= \mathsf{G}(d \rightarrow \mathsf{Eco}\, e) \\
\psi_e &= \mathsf{G}(e \rightarrow (\mathsf{Eco}\, \widehat{a} \vee \mathsf{G}\, e)) \\
\psi_{\text{start}} &= \widehat{a} \wedge \mathsf{Aco}\left(b \wedge \mathsf{Aco}\,(c \rightarrow \mathsf{G}(\widehat{a} \rightarrow \mathsf{Eco}\, e)) \wedge \mathsf{Eco}\, c \wedge (\neg \mathsf{Eco}\, a \vee \neg \mathsf{Eco}\, \overline{a})\right)
\end{aligned}
$$

One can show that every $(t, x) \in K$ satisfies $t, x \models \psi$, where x is the minimal vertex of q in the factorization $t = pq$ above: Since no vertex of q is parallel to any vertex in p it is enough to verify that $q, x \models \psi$.

For the other direction, suppose $t, x \models \psi$. By stutter-invariance of the fragment $\mathrm{TL}[\mathsf{F}, \mathsf{Eco}](\mathbb{M})$ and $\{K\}$ we can (w.l.o.g.) assume that t is stutter-free, i.e., t does not contain two consecutive positions with the same label. With ψ_{Eco} we ensure that no position in the future of x has both parallel, a and $\overline{a}$. (Remember that 'G' means "globally in the future".) This implies that every position in the future of x is parallel to at most two different letters. Formulas ψ_a, ψ_b, ψ_c, ψ_d and ψ_e assure that starting from x with label $\widehat{a}$ we can always follow a path $\widehat{a}$, b, c, d, e of parallel occurrences by using the Eco operator. Furthermore, if we have not reached a maximal vertex e, we must move from e to some parallel $\widehat{a}$; and we have to start all over again. With ψ_{start} we ensure a well-defined starting configuration. We start from a position x labeled by $\widehat{a}$ and all parallel vertices of x are labeled by b and from none of those b's we can reach a position in the past of x which is labeled by c. All c's that we can reach using two Eco operators are therefore in the future of x. Additionally, we have to repeat the constraints of ψ_b and ψ_{Eco} for concurrent b's.

Using ψ_{Eco} and ψ_a we can treat $\widehat{a}$ as if it were a single letter. From ψ_{start} we conclude $t = p\widehat{a}bq'$ such that every maximal element of p is dependent of both, $\widehat{a}$ and b. Therefore $\max(p) \subseteq \{d\}$. Furthermore there exists some c in q' such that the letter b in $p\widehat{a}bq'$ is concurrent to that c. By stutter-freeness of t it follows that $\{\widehat{a}, b\} \cap \min(q') = \emptyset$ and hence $\min(q') \subseteq \{c, d, e\}$, but $\min(q') \subseteq \{d, e\}$ is not possible because this would contradict the existence of some parallel c for b since d and e are both dependent of b. Hence $q' = cq$ and we can write $t = p\widehat{a}bcq$. Now by ψ_c there exists a d that is parallel to c. Since $(d, \widehat{a}), (\widehat{a}, c) \in D$, this d has to be in q. We have $\min(q) \subseteq \{\widehat{a}, d, e\}$. As before we conclude $q = d\widetilde{q}$ since $\min(q) \subseteq \{\widehat{a}, e\}$ is not possible. We can continue with $t = \widetilde{p}bcd\widetilde{q}$ where $\widetilde{p} = p\widehat{a}$. The situation is analogous to $t = p\widehat{a}bcq$, we only moved one step on the cycle of the independence relation I.

$I =$ 

This process can only stop if $t = p(abcde)^+ r$ and the last e in $(abcde)^+$ satisfies $\mathsf{G}\, e$. It follows that r is empty, since r can only contain the letters $\{a, d, e\}$ but none of them can be minimal.

This shows that if $\varphi \in \mathrm{TL}[\mathsf{X}, \mathsf{F}]$ is satisfiable over $\{a, \overline{a}\}^+$ then $\widetilde{\varphi} \wedge \psi$ is also satisfiable over $\mathbb{M}$. For the other direction, suppose there exists a trace $t \in \mathbb{M}$ and a position $x \in t$ such that $t, x \models \widetilde{\varphi} \wedge \psi$. By construction of ψ we have $t = pq$ with $\max p \subseteq \{d\}$ and $q \in ((a^+ \cup \overline{a}^+)b^+c^+d^+e^+)^+$. Moreover x is a position in q. With the temporal operators F and $\mathsf{E}\mathsf{co}$ we cannot reach positions in p. It follows $q, x \models \widetilde{\varphi} \wedge \psi$. By stutter-invariance of $\mathrm{TL}[\mathsf{F}, \mathsf{E}\mathsf{co}]$ there exists $q' \in ((a \cup \overline{a})bcde)^+$ such that $q', x' \models \widetilde{\varphi} \wedge \psi$ where x' is a minimal position labeled by either a or $\overline{a}$. By (4) we conclude that φ is satisfiable over $\{a, \overline{a}\}^*$. $\square$

Remark 6.1 *It has been shown in [11] that the uniform variant of the above satisfiability problem is in* PSPACE*:*

Input: A dependence alphabet (Γ, D) *and a formula* $\varphi \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{E}\mathsf{co}]$.
Question: Is φ *satisfiable over* $\mathbb{M} = \mathbb{M}(\Gamma, D)$*?*

Theorem 6.3 *The following problem is in* NEXPTIME*:*

Input: A dependence alphabet (Γ, D) *and a formula* $\varphi \in \mathrm{FO}^2[E]$.
Question: Is φ *satisfiable over* $\mathbb{M} = \mathbb{M}(\Gamma, D)$*?*

Proof: Replace in $\varphi \in \mathrm{FO}^2[E]$ each atomic predicate $(x, y) \in E$ by $x < y \wedge (\lambda(x), \lambda(y)) \in D$. We obtain a formula φ^* of size at most $|\varphi| \cdot |\Gamma|$ such that $\emptyset \neq L(\varphi) \subseteq \mathbb{M}$ if and only if $\emptyset \neq L(\varphi^*) \subseteq \Gamma^*$, see also Lemma 2.2. Having done this, we can apply the known NEXPTIME-algorithm on words, [7]. $\square$

Remark 6.2 *It was shown in [23] that if* Γ *is not part of the input then the satisfiability problem for* $\mathrm{FO}^2[<](\Gamma^*)$ *is* NP*-complete. Therefore, for fixed dependence graphs* (Γ, D) *the reduction technique of Theorem 6.3 yields* NP*-completeness of the following problem:*

Input: A formula $\varphi \in \mathrm{FO}^2[E]$.
Question: Is φ *satisfiable over* $\mathbb{M} = \mathbb{M}(\Gamma, D)$*?*

Theorem 6.4 *The following problem is in* EXPSPACE*:*

Input: A dependence alphabet (Γ, D) *and a formula* $\varphi \in \mathrm{FO}^2[<]$.
Question: Is φ *satisfiable over* $\mathbb{M} = \mathbb{M}(\Gamma, D)$*?*

Proof: For membership in EXPSPACE, replace $\varphi \in \mathrm{FO}^2[<]$ according to the proof of Lemma 4.1 in order to obtain an equivalent formula $\psi \in \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{E}\mathsf{co}]$ of exponential size such that $\emptyset \neq L(\varphi) \subseteq \mathbb{M}$ if and only if $\emptyset \neq L(\psi) \subseteq \mathbb{M}$. Having done this, we can apply the PSPACE algorithm for local temporal logic on traces, [11]. $\square$

Theorem 6.5 *The following problem is* NEXPTIME*-hard:*

Input: A dependence alphabet (Γ, D) *and a formula* $\varphi \in \mathrm{FO}^2[<]$.
Question: Is φ *satisfiable over* $\mathbb{M} = \mathbb{M}(\Gamma, D)$*?*

Proof: It is shown in [7] that the following problem is NEXPTIME-hard:

Input: A finite set of unary predicates Σ and a formula $\varphi \in \mathrm{FO}^2[\,]$ using only the predicates in Σ and equality as atomic propositions.

Question: Does there exist a word $w \in \left(2^\Sigma\right)^*$ such that $w \models \varphi$.

We reduce this satisfiability problem over words to the problem over traces above. We simulate the underlying word-alphabet of exponential size by concurrency. We let \$, ¢ and # be fresh letters and we define a dependence alphabet by $\Gamma = \Sigma \cup \{\,\$, ¢, \#\,\}$ such that all letters in $\Sigma \cup \{\$, ¢\}$ are independent of each other and dependent of #. Let $\varphi \in \mathrm{FO}^2[\,]$ and let x and y be the only two variables in φ. Note that the only binary atomic predicate in φ is $x = y$. By $\mathrm{FO}^2[\|]$ we denote the fragment of $\mathrm{FO}^2[<]$ where $x \parallel y$ is the only binary atomic proposition. The formula φ is now transformed into a formula $\widetilde{\varphi} \in \mathrm{FO}^2[\|] \subseteq \mathrm{FO}^2[<]$ using the following rules:

$$\begin{aligned}
\widetilde{a(x)} &= \exists y\,(x \parallel y \wedge a(y))\\
\widetilde{a(y)} &= \exists x\,(x \parallel y \wedge a(x))\\
\widetilde{x = y} &= x \parallel y\\
\widetilde{\exists x\,\psi} &= \exists x\,(\$(x) \wedge \widetilde{\psi})\\
\widetilde{\exists y\,\psi} &= \exists y\,(¢(y) \wedge \widetilde{\psi})\\
\widetilde{\neg\psi} &= \neg\widetilde{\psi}\\
\widetilde{\psi_1 \vee \psi_2} &= \widetilde{\psi_1} \vee \widetilde{\psi_2}
\end{aligned}$$

We see that the purpose of the letter \$ is to suit as position for the variable x and the letter ¢ is for the variable y. We define $\psi \in \mathrm{FO}^2[\|]$ to ensure that models always have suitable positions for x and y:

$$\begin{aligned}
\psi = \forall x\,\big(&\#(x) \vee\\
&\big(\$(x) \wedge \exists y\,(x \parallel y \wedge ¢(y))\big) \vee\\
&\big(¢(x) \wedge \exists y\,(x \parallel y \wedge \$(y))\big) \vee\\
&\big(\exists y\,(x \parallel y \wedge ¢(y)) \wedge \exists y\,(x \parallel y \wedge \$(y))\big)\big)
\end{aligned}$$

A word $w = \left\{a_1^1, \dots, a_k^1\right\} \cdots \left\{a_1^n, \dots, a_\ell^n\right\}$ over the alphabet 2^Σ is transformed into the trace $\widetilde{w} = \$¢a_1^1 \cdots a_k^1 \# \cdots \#\$¢a_1^n \cdots a_\ell^n \in \mathbb{M}$. We have

$$w \models \varphi \;\Leftrightarrow\; \widetilde{w} \models \widetilde{\varphi} \wedge \psi \;\Leftrightarrow\; \#\widetilde{w} \models \widetilde{\varphi} \wedge \psi \;\Leftrightarrow\; \widetilde{w}\# \models \widetilde{\varphi} \wedge \psi \;\Leftrightarrow\; \#\widetilde{w}\# \models \widetilde{\varphi} \wedge \psi$$

The hardness result now follows since $L(\widetilde{\varphi} \wedge \psi)$ is stutter-invariant. $\square$

Remark 6.3 *The proof above shows actually more since* NEXPTIME-*hardness is shown for the fragment* $\mathrm{FO}^2[\|]$ *without equality. However, it remains open whether the problem in Theorem 6.3 is* NEXPTIME-*complete because* $\mathrm{FO}^2[\|]$ *is not a fragment of* $\mathrm{FO}^2[E]$*, see (the proof of) Theorem 5.2. It is also open whether the problem in Theorem 6.4 and 6.5 is* EXPSPACE-*complete.*

7 Ehrenfeucht-Fraïssé games

The central questions in this section are: What is the relation between the fragments $\mathrm{FO}^2[<](\mathbb{M})$ and $\Delta_2[<](\mathbb{M})$? When is $\Sigma_n[E](\mathbb{M})$ a proper subclass of $\Sigma_n[<](\mathbb{M})$? And: Is $\Pi_n[<](\mathbb{M})$ contained in $\Sigma_{n+1}[E](\mathbb{M})$? A main tool in answering these questions are Ehrenfeucht-Fraïssé games. In [8] an Ehrenfeucht-Fraïssé game is defined for the linear temporal logic on words with three operators: *until*, *eventually* and *next*. We adapt this game in order to characterize simple fragments of temporal logic on traces.

Definition 7.1 *(EF game for* $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{Eco}]$*) Ehrenfeucht-Fraïssé (EF) games are played by two persons, called here* Spoiler *and* Duplicator*. For the fragment* $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{Eco}]$ *with n rounds it is played on two traces $t_0 = [V_0, E_0, \lambda_0]$ and $t_1 = [V_1, E_1, \lambda_1]$ using one pebble for each trace. A configuration of the game is a pair of positions $(x_0, x_1) \in V_0 \times V_1$ currently occupied by the pebbles. In each round, Spoiler selects a side $\sigma \in \{0, 1\}$ and one of the moves* SF, $\overleftarrow{\mathsf{SF}}$ *and* Eco*. The rules are simple:*

> SF*: From a position x the pebble is moved to a position y such that $x < y$.*
> $\overleftarrow{\mathsf{SF}}$*: From a position x the pebble is moved to a position y such that $y < x$.*
> Eco*: From a position x the pebble is moved to a position y such that $x \parallel y$.*

First, Spoiler moves the pebble on t_σ and next, Duplicator carries out the same type of move on $t_{1-\sigma}$. This defines one round. Spoiler wins if either Duplicator cannot move his pebble according to the rules or if, after some round the pebbles lie on differently labeled nodes. Duplicator wins if this never occurs within n rounds.

Lemma 7.1 *Let $n \in \mathbb{N}$ and $s, t \in \mathbb{M}$ with $x_s \in s$ and $x_t \in t$. Then the following propositions are equivalent:*

i. Duplicator has a winning strategy for the game in Definition 7.1 with n rounds played on the traces s, t starting at configuration (x_s, x_t).

ii. (s, x_s) and (t, x_t) are models of exactly the same $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{Eco}]$ *formulas with depth n.*

Proof: For $n = 0$ this is true, because at two positions the same formulas of depth 0 hold if and only if they have the same labels. Let $n > 0$. Let $\mathsf{SF}\varphi$ be a formula of depth n such that (without loss of generality) $s, x_s \models \mathsf{SF}\varphi$, whereas $t, x_t \not\models \mathsf{SF}\varphi$. In the game with n rounds starting with the configuration (x_s, x_t), Spoiler can select the SF move and a position $y_s > x_s$ with $s, y_s \models \varphi$, as opposed to Duplicator, who will find no analogous position $y_t > x_t$. Therefore, Spoiler wins this game by induction. The other temporal operators are handled analogously.

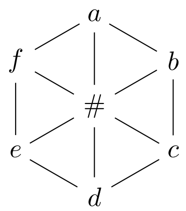
For the other direction let Spoiler win the game starting with the configuration (x_s, x_t) within n rounds. If the positions have different labels, then $s, x_s \models \lambda(x_s)$ whereas $t, x_t \not\models \lambda(x_s)$. Otherwise Spoiler makes his first move. Without loss of generality let this first move be SF on s. Spoiler moves his pebble to y_s. By induction, for every position $y \in t$ with $y > x_t$ there exists a formula φ_y of depth $(n-1)$ such that $s, y_s \models \varphi_y$ and $t, y \not\models \varphi_y$. Let $\varphi = \bigwedge_{y > x_t} \varphi_y$. Then by construction $s, x_s \models \mathsf{SF}\varphi$ and $t, x_t \not\models \mathsf{SF}\varphi$. $\Box$ In the proof above we used the fact that s and t are finite traces, but this is not essential because there are only finitely many formulas of a given depth. Therefore a conjunction $\bigwedge_{y>x_t} \varphi_y$ over formulas of depth $n-1$ is always finite.

Theorem 7.1 *There exists a trace monoid $\mathbb{M}$ such that the $\Sigma_1[<]$ definable language*

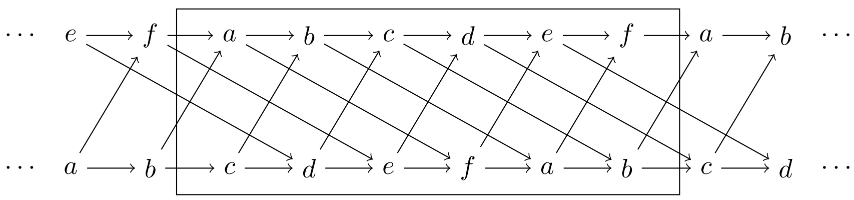
$$L = \{\, t \in \mathbb{M} \mid \exists x, y, z \in t\colon\ (x \parallel y \wedge y \parallel z \wedge z \parallel x) \,\}$$

of traces with three pairwise independent vertices is not in $\mathrm{FO}^2[<](\mathbb{M})$.

Proof: The formula $\exists x, y, z\colon\ (x \parallel y \wedge y \parallel z \wedge z \parallel x) \in \Sigma_1[<]$ expresses L. Let $\mathbb{M}$ be a trace monoid over the alphabet $\Gamma = \{\, a, b, c, d, e, f, \# \,\}$. The graph of the dependence relation D is

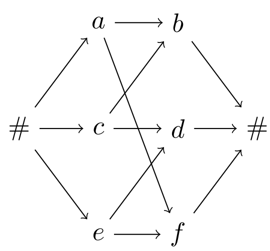
$D =$ 

The maximal sets of pairwise independent letters are $\{\, a, c, e \,\}$ and $\{\, b, d, f \,\}$. Now, we show that L is not in $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{Eco}]$. Let $r = acbdcedfeafb$. For an arbitrarily chosen $n \in \mathbb{N}$ consider the traces $p = r^{2n+1}$ and $q = acebdf$. The trace p looks as follows



r

and the Hasse diagram of $\#q\#$ is



We combine p and q in order to build the larger traces $s = (\#p)^{2n+1} \notin L$ and $t = (\#p)^n \# q (\#p)^n \in L$. We say that p and q are *segments* of s and t. We numerate the segments of s and t with numbers running from $-n$ to n from left to right. All segments of s and t consist of p, except for segment 0 of trace t, which is q. Every segment that corresponds to p is further subdivided into $2n+1$ *blocks* with numbers running from $-n$ to n such that all blocks consist of r.

Consider the n-round Ehrenfeucht-Fraïssé game for $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{Eco}]$ played on the traces s and t. We start on both traces with pebbles on the unique minimal positions. In all blocks, except the outermost ones $-n$ and n, every position has parallel occurrences of all letters that are independent. Although there do not exist three parallel positions in s, we will use this fact to mimic three parallel positions in order to construct a winning strategy for Duplicator.

The main strategy of Duplicator is to copy all moves of Spoiler. The behavior on the segments 0 of s and t is not evident, therefore we will describe the strategy of Duplicator for this case. Whenever Spoiler accesses segment 0 of trace s, Duplicator avoids placing the pebble on segment 0 of trace t by putting his pebble to the corresponding position of either segment -1 or 1 of trace t. In the up to $n-1$ remaining rounds it is possible for Duplicator to keep a maximal difference of 1 between the numbers of the segments. If Spoiler moves his pebble on segment 0 of trace t, Duplicator responds by placing the other pebble on an identically labeled position of block 0 of either segment -1, 0 or 1 of trace s. From this position he can mimic all possible Eco-moves of Spoiler on t. Each Eco-move of Spoiler could force Duplicator to increase or decrease the block number by 1. Since there occur at most $n-1$ of these moves and there are n blocks to the left as well as to the right, Duplicator wins the game. By Lemma 7.1 it follows that $L \notin \mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{Eco}](\mathbb{M}) = \mathrm{FO}^2[<](\mathbb{M})$. □

Since $\Sigma_1[<](\mathbb{M}) \subseteq \Delta_2[<](\mathbb{M})$, it follows from Theorem 7.1 that there exists a trace monoid $\mathbb{M}$ such that $\mathrm{FO}^2[<](\mathbb{M})$ is not a subset of $\Delta_2[<](\mathbb{M})$.

Next we will introduce Ehrenfeucht-Fraïssé games for $\Sigma_n[<]$. In order to characterize this first-order fragment with games, we first have to refine it. We inductively define $\Sigma_n^m[<]$. Intuitively, n describes the number of quantifier blocks and m the number of nested bound variables. More formally: the formulas without quantifiers constitute $\Sigma_0^0[<]$. A formula φ is in $\Sigma_n^m[<]$ if and only if it is a disjunction of formulas of the form

$$\exists x_1 \cdots \exists x_k \neg\psi$$

with $0 \leq k \leq m$ and $\psi \in \Sigma_{n-1}^{m-k}[<]$. We assume that all variables in a $\Sigma_n^m[<]$ formula are distinct. We have $\Sigma_n[<] = \bigcup_{m\in\mathbb{N}} \Sigma_n^m[<]$. For the Boolean closure of $\Sigma_n[<]$, Thomas presented an Ehrenfeucht-Fraïssé game [22]. It can be modified in order to describe $\Sigma_n[<]$. The main difference consists of the fact that $\Sigma_n[<]$ is not closed under complementation. If we want to capture this fragment, it is therefore insufficient to determine whether two traces s and t are equivalent or not. Instead, we ask if s models *at least the same* $\Sigma_n[<]$ formulas as t does. Following [13] it is possible to characterize the $\Sigma_n^m[<]$ fragments by limiting the number of pebbles to $m \in \mathbb{N}$.

Definition 7.2 *(Ehrenfeucht-Fraïssé game for $\Sigma_n^m[<]$) The set of* configurations *for the game played on the traces t_0 and t_1 with nodes V_0 and V_1, respectively, is $V_0^* \times V_1^* \times \{0,1\}$ with the restriction that the size of the first two components is equal and does not exceed m. The first two components of the configuration are interpreted as a distribution of pebbles on the two traces: a pebble labeled with i lies at position $x_i \in V_j$ whenever x_i is the i-th character of the word corresponding to $j \in \{0,1\}$. The third component contains the number of the trace where Spoiler will carry out his next move. Let (w_0, w_1, σ) with $|w_0| = |w_1| \leq m$ and $\sigma \in \{0,1\}$ be the current configuration, then the next turn is carried out as follows:*

- *Spoiler takes $i \le m - |w_0|$ pebbles and distributes them on trace t_σ by assigning them positions $x_1 \cdots x_i \in V_\sigma^i$.*

- *Duplicator places identically labeled pebbles on nodes $y_1 \cdots y_i \in V_{1-\sigma}^i$ of the other trace.*

- *The new configuration is $(w_0 x_1 \cdots x_i, w_1 y_1 \cdots y_i, 1 - \sigma)$.*

The game for $\Sigma_n^m[<]$ consists of n rounds. Duplicator wins if and only if initially and after each of these rounds, the partial mapping $V_0 \to V_1 : w_0(j) \mapsto w_1(j)$ with $1 \le j \le |w_0| = |w_1|$ induces an isomorphism with respect to the labels and the relation $<$.

Lemma 7.2 *Let $n, m, k \in \mathbb{N}$ and $s, t \in \mathbb{M}$ with sequences $w_s \in V_s^k$ and $w_t \in V_t^k$. Then the following propositions are equivalent:*

i. *Duplicator has a winning strategy for the game on s and t from Definition 7.2 with n rounds, maximally m pebbles, and starting with the configuration $(w_s, w_t, 0)$.*

ii. *For all $\varphi \in \Sigma_n^m[<]$ with k free variables we have that $s, w_s \models \varphi$ implies $t, w_t \models \varphi$.*

Proof: The lemma holds for $n = 0$, as without any rounds Duplicator wins the game if and only if in the initial configuration, the k pebbles are isomorphically distributed on both traces, which amounts to saying that the same $\Sigma_0^0[<]$ formulas hold for (s, w_s) and (t, w_t). Suppose $n > 0$. Let $w_s \in V_s^k$, $w_t \in V_t^k$ be interpretations such that the partial mapping $V_s \to V_t : w_s(j) \mapsto w_t(j)$, $1 \le j \le k$ induces an isomorphism with respect to the order relation $<$ and the label function λ.

(i. $\Rightarrow$ ii.) Suppose Duplicator has a winning strategy for the Ehrenfeucht-Fraïssé game with n rounds starting from the configuration $(w_s, w_t, 0)$. The number 0 in the configuration means that Spoiler starts to play on trace s. Let $\varphi \in \Sigma_n^m[<]$ be a formula with the free variables $x_1, \ldots, x_k$ such that (s, w_s) is a model of φ. Without loss of generality, we assume that $\exists$ (and not $\vee$) is the outermost junctor, i.e., $\varphi = \exists x_{k+1} \cdots \exists x_{k+\ell} \neg\psi$ with $\ell \le m$ and $\psi \in \Sigma_{n-1}^{m-\ell}[<]$. Now let Spoiler distribute ℓ pebbles on positions $x_1, \ldots, x_\ell \in V_s$ such that $\neg\psi$ holds on s with the interpretation $v_s = w_s x_1 x_2 \cdots x_\ell$. If Duplicator proceeds according to his winning strategy, he obtains positions $y_1, \ldots, y_\ell \in V_t$. We set $v_t = w_t y_1 y_2 \cdots y_\ell$. Now, Duplicator wins the game with $n-1$ rounds and up to $m - \ell$ pebbles starting from the configuration $(v_s, v_t, 1)$ which is symmetric to $(v_t, v_s, 0)$. By induction hypothesis the implication $t, v_t \models \psi \Rightarrow s, v_s \models \psi$ holds. Hence, from $s, v_s \models \neg\psi$ we can conclude $t, v_t \models \neg\psi$ and therefore $t, w_t \models \varphi$.

(ii. $\Rightarrow$ i.) Assume that Spoiler has a winning strategy starting from the configuration $(w_s, w_t, 0)$. Let his first move according to this strategy consist in placing $\ell \le m$ pebbles on s. Let v_s be the new configuration on s. Now, after every possible response of Duplicator, Spoiler has a winning strategy with at most $n-1$ rounds starting on trace t. By induction, for each $v \in w_t V_t^\ell$ there exists a formula $\psi_v \in \Sigma_{n-1}^{m-\ell}[<]$ such that $t, v \models \psi_v$ and $s, v_s \not\models \psi_v$. Since the range of possible values for v is finite, we can construct $\psi^* = \bigvee_v \psi_v$, which in turn is a $\Sigma_{n-1}^{m-\ell}[<]$ formula. By construction we have

$$s, w_s \models \exists x_1 \cdots \exists x_\ell \neg\psi^* \quad \text{whereas} \quad t, w_t \not\models \exists x_1 \cdots \exists x_\ell \neg\psi^*$$

The lemma now follows by contraposition. $\square$

Theorem 7.2 *For every $n \geq 1$ there exists a trace monoid $\mathbb{M}$ and a trace language $L \subseteq \mathbb{M}$ with*

$$L \in \mathrm{FO}^2[<](\mathbb{M}) \cap \Pi_n[<](\mathbb{M}) \quad \textit{but} \quad L \notin \Sigma_n[<](\mathbb{M})$$

Furthermore, L is the complement of a language $L_\exists(\varphi)$ where φ is a $\mathrm{TL}[\mathsf{Eco}]$ *formula with n occurrences of the operator* Eco*.*

Proof: Let $m \in \mathbb{N}$ be arbitrary, and let a series of dependence alphabets (Γ_n, D_n), $n \geq 1$, be given inductively by $\Gamma_1 = \{a_1\}$ and $\Gamma_n = \Gamma_{n-1} \cup \{a_n, b_n\}$ with the dependence graph $D_n = D_{n-1} — b_n — a_n$ or, more formally, $D_1 = \{(a_1, a_1)\}$ and $D_n = D_{n-1} \cup \Gamma_n \times \{b_n\} \cup \{b_n\} \times \Gamma_n \cup \{(a_n, a_n)\}$. This means that each alphabet Γ_n with $n \geq 2$ introduces a letter a_n, which is independent of all preceding letters, and a letter b_n that depends on all letters. For $n \in \mathbb{N}$, let $\ell_n = (m+1)^n$. We also define the traces $s_1 = 1$, $t_1 = a_1$, and

$$\begin{aligned} s_n &= (b_n a_n t_{n-1})^{\ell_n} \\ t_n &= (b_n a_n t_{n-1})^{\ell_n} \cdot b_n a_n s_{n-1} \cdot (b_n a_n t_{n-1})^{\ell_n} \end{aligned}$$

such that s_n and t_n are traces over the dependence alphabet (Γ_n, D_n). The b_n's partition these traces into *blocks*. We number these blocks from 1 to ℓ_n on s_n and from $-\ell_n$ to ℓ_n on t_n. Within each block, the position labeled with a_n is parallel to all other positions. Below, we depict the blocks -1, 0, and 1 of the trace t_n.

$$\begin{tikzcd}
 & & t_{n-1} \arrow[dr] & & s_{n-1} \arrow[dr] & & t_{n-1} & \\
\cdots & b_n \arrow[ur] \arrow[dr] & & b_n \arrow[ur] \arrow[dr] & & b_n \arrow[ur] \arrow[dr] & & \cdots \\
 & & a_n \arrow[ur] & & a_n \arrow[ur] & & a_n &
\end{tikzcd}$$

We define the $\mathrm{TL}[\mathsf{Eco}]$ formulas $\varphi_1 = a_1$ and $\varphi_n = a_n \wedge \neg \mathsf{Eco}\, \varphi_{n-1}$ for $n > 1$. Let x be the position of a_n in the 0-th block of t_n. For all $n \geq 1$, by induction we have $t_n, x \models \varphi_n$ whereas there does not exist $y \in s_n$ with $s_n, y \models \varphi_n$. It follows $t_n \in L_\exists(\varphi_n)$ and $s_n \notin L_\exists(\varphi_n)$. The formula φ_n has operator depth $n-1$ and hence $\mathsf{M}\varphi_n$ has operator depth n. This implies that the language $L_\exists(\varphi_n) = L(\mathsf{M}\varphi_n)$ can be expressed by a $\Sigma_n[<]$ sentence.

In order to show that the complement of $L_\exists(\varphi_n)$ is not in $\Sigma_n[<](\mathbb{M})$ we consider the Ehrenfeucht-Fraïssé game for $\Sigma_n^m[<]$ played on the traces s_n and t_n, i.e., there are n rounds, m pebbles, and Spoiler places his first pebbles on s_n. Using induction, we describe a winning strategy of Duplicator for this game. The case for $n = 1$ is trivial: Spoiler cannot place any pebbles since s_1 is empty; therefore Duplicator wins by not placing any pebbles. Assume $n > 1$, then in the first round, Spoiler places $m' \leq m$ pebbles on s_n. Because this trace consists of ℓ_n blocks, there must remain a big continuous gap of $\ell' = (m+1)^{n-1}$ blocks without any pebbles. Let such a gap start after the k-th block and consider the following factorizations of s_n and t_n:

$$\begin{aligned} s_n &= p^k \cdot \qquad p^{\ell'} \qquad \cdot p^{\ell_n - k - \ell'} \\ t_n &= p^k \cdot p^{\ell_n - k}\, q\, p^{k+\ell'} \cdot p^{\ell_n - k - \ell'} \end{aligned}$$

where $p = b_n a_n t_{n-1}$ and $q = b_n a_n s_{n-1}$. Duplicator can imitate the move of Spoiler on the outer factors of t_n. For the remaining rounds, we can ignore the outermost factors of both traces because they are identical. By induction we know that Duplicator wins the game on s_{n-1} and t_{n-1}, i.e., for the rest of the game, the blocks p and q cannot be distinguished. Note that the change of ends enforces Spoiler to make his next move on the s_{n-1}-side. Both middle factors consist of at least $(m+1)^{n-1}$ blocks, and there are $n-1$ rounds to play. Therefore, the different number of blocks in s_n and t_n is no advantage for Spoiler and hence, Duplicator wins the game. By Lemma 7.2 we conclude that $s_n \in L(\psi)$ implies $t_n \in L(\psi)$ for all formulas $\psi \in \Sigma_n^m[<]$. This shows, that for all $n \in \mathbb{N}$, no $\Sigma_n^m[<]$ formula can express the complement of $L_\exists(\varphi_n)$ and since $\Sigma_n[<] = \bigcup_m \Sigma_n^m[<]$ we have $L = \mathbb{M} \setminus L_\exists(\varphi_n)$ is not in $\Sigma_n[<](\mathbb{M})$. □

Theorem 7.3 *The following assertions are equivalent:*

i. D is transitive, i.e., $\mathbb{M}$ is a direct product of free monoids.

ii. $\exists n \geq 1\colon \Sigma_n[E](\mathbb{M}) = \Sigma_n[<](\mathbb{M})$.

iii. $\exists n \geq 1\colon \Pi_n[<](\mathbb{M}) \subseteq \Sigma_{n+1}[E](\mathbb{M})$.

Proof: The implications from (i.) to (ii.) and from (i.) to (iii.) are obvious since if (i.) holds then E and $<$ are identical. The implication from (ii.) to (iii.) follows from the basic fact that $\Pi_n[<](\mathbb{M}) \subseteq \Sigma_{n+1}[<](\mathbb{M})$. Note that if $\Sigma_1[E](\mathbb{M}) = \Sigma_1[<](\mathbb{M})$ then the example in Theorem 5.2 shows that D has to be transitive. For the remaining direction (iii.) to (i.) assume that D is not transitive. Then there exist letters a, b, c with the dependencies $a — b — c$. We show that for each $n \geq 1$ there exists a language L_{n+1} in $\Pi_n[<](\mathbb{M})$ that is not in $\Sigma_{n+1}[E](\mathbb{M})$. Define the traces $s_1 = acb$, $t_1 = abcb$, and for $n \geq 1$:

$$\begin{aligned} s_{n+1} &= (t_n b)^\ell \\ t_{n+1} &= (t_n b)^\ell \cdot s_n b \cdot (t_n b)^\ell \end{aligned}$$

where $\ell > 1$ is an arbitrary number. We refer to the factors $t_n b$ and $s_n b$ as *blocks*. We number the blocks from 1 to ℓ in s_n and from $-\ell$ to ℓ in t_n. For $n \geq 2$ every block in s_n and t_n ends with n occurrences of the letter b and this is the only occurrence of n consecutive b's in each block. Additionally, the minimal elements of blocks are not labeled by b. Hence, the factors b^n partition the blocks. Note that $t_2 \models \exists x \exists y\colon x \parallel y$ whereas $s_2 \not\models \exists x \exists y\colon x \parallel y$. Hence for the $\Pi_0[<]$ formula

$$\psi_2(x, y) = x \parallel y$$

we have $s_2 \in L_2 = L(\forall x \forall y\ \neg\psi_2(x, y))$, but $t_2 \notin L_2$. By induction, suppose that we have a formula $\psi_n(x_1, \ldots, x_{2n})$ in $\Pi_{n-2}[<]$ with $2n$ free variables such that

$$\begin{aligned} s_n &\models \forall x_1 \cdots \forall x_{2n}\colon \neg\psi_n(x_1, \ldots, x_{2n}) \\ t_n &\models \exists x_1 \cdots \exists x_{2n}\colon \psi_n(x_1, \ldots, x_{2n}) \end{aligned}$$

and (for $n > 2$) all bound variables of ψ_n are relativized to lie between x_n and x_{2n}. This property is related to the definition of blocks. Using ψ_n we construct a formula $\psi_{n+1} \in \Pi_{n-1}[<]$. Let $y_1, \ldots, y_{n+1}$ and $z_1, \ldots, z_{n+1}$ be new variables. For $n > 2$ the idea is that the formula ψ_{n+1} expresses that $y_1, \ldots, y_{n+1}$ and $z_1, \ldots, z_{n+1}$ are the borders of an s_n-block, i.e., a block that satisfies $\forall \overline{x}\colon \neg\psi_n(\overline{x})$. For $n = 2$ we only express that $y_1, \ldots, y_{n+1}$ is the left border of some block i and that $z_1, \ldots, z_{n+1}$ is the right border of some block j with $i \leq j$ and within these borders there are no parallel positions. We give an informal description of ψ_{n+1}:

$$
\begin{aligned}
&\psi_{n+1}(y_1, \ldots, y_{n+1}, z_1, \ldots, z_{n+1}) \;=\; \\
&\quad y_1 < y_2 < \cdots < y_{n+1} < z_1 < z_2 < \cdots < z_{n+1} \;\wedge \\
&\quad \text{“all } y_i \text{ and all } z_i \text{ are labeled by } b\text{”} \;\wedge \\
&\quad \forall x\colon \text{“} x \text{ does not lie between any two consecutive } y_i\text{'s”} \;\wedge \\
&\quad \forall x\colon \text{“} x \text{ does not lie between any two consecutive } z_i\text{'s”} \;\wedge \\
&\quad \sigma_{n+1}(y_{n+1}, z_1) \;\wedge \\
&\quad \forall x_1 \cdots \forall x_{2n}\colon \text{“all } x_i \text{ lie between } y_{n+1} \text{ and } z_{n+1}\text{”} \;\rightarrow\; \neg\psi_n(x_1, \ldots, x_{2n})
\end{aligned}
$$

where for $n > 2$ we define

$$
\begin{aligned}
\sigma_{n+1}(y, z) \;=\; \forall x_1 \cdots \forall x_{n+1}\colon &\Bigg(y < x_1 < \cdots < x_{n+1} < z \;\wedge \bigwedge_{1 \leq i \leq n+1} \lambda(x_i) = b \Bigg) \\
&\rightarrow \Bigg(\exists x\colon \bigvee_{1 \leq i < n+1} x_i < x < x_{i+1} \Bigg)
\end{aligned}
$$

The formula $\sigma_{n+1}(y, z)$ is true if there are no $n + 1$ consecutive occurrences of the letter b between y and z. Note that $\sigma_{n+1} \in \Pi_2[<]$ which for $n > 2$ yields $\psi_{n+1} \in \Pi_{n-1}[<]$. For $n = 2$ we define $\sigma_3(y, z) = \top$ which yields $\psi_3 \in \Pi_1[<]$. For $n = 2$, block 0 in trace t_3 contains no parallel positions whereas all blocks in trace s_3 contain some parallel positions. The formula ψ_3 uses this property to distinguish between s_3 and t_3. For all $n \geq 2$ the formula ψ_{n+1} has the desired relativization property and for all $n \geq 1$ we have

$$
\begin{aligned}
s_{n+1} &\models \forall x_1 \cdots \forall x_{2(n+1)}\colon \neg\psi_{n+1}(x_1, \ldots, x_{2(n+1)}) \\
t_{n+1} &\models \exists x_1 \cdots \exists x_{2(n+1)}\colon \psi_{n+1}(x_1, \ldots, x_{2(n+1)})
\end{aligned}
$$

where for $n = 1$ and $n = 2$ this holds by construction and therefore for $n > 2$ it follows by induction since t_{n+1} contains an s_n-block whereas s_{n+1} contains only t_n blocks. Altogether, this shows that for all $n \geq 1$ there is a language $L_{n+1} \in \Pi_n[<](\mathbb{M})$ such that $s_{n+1} \in L_{n+1}$ and $t_{n+1} \notin L_{n+1}$ just by taking $L_{n+1} = L(\forall x_1 \cdots \forall x_{2(n+1)}\colon \neg\psi_n)$.

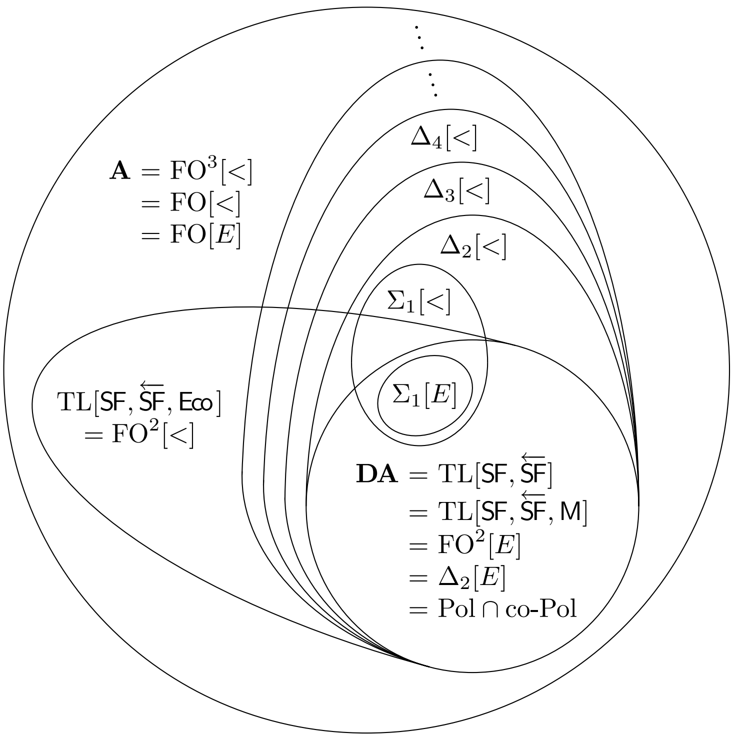
We now show for all $n \geq 2$ that $L_n \notin \Sigma_n[E](\mathbb{M})$ holds. For this purpose we verify that $L \in \Sigma_n[E](\mathbb{M})$ and $s_n \in L$ implies $t_n \in L$ where the exponent ℓ in the construction of s_n and t_n is chosen to be large enough. It is easy to see that one can define an Ehrenfeucht-Fraïssé game for $\Sigma_n^m[E]$ with m pebbles and n rounds similar to the game in Definition 7.2 such that an analogue of Lemma 7.2 holds. We now fix the number m of pebbles. Let the exponent ℓ in the definition of s_n and t_n satisfy $\ell \geq (m+1)^n$. Analogously to the proof of Theorem 7.2 we inductively describe a winning strategy of Duplicator for the game played on s_n and t_n. Again, the idea is to put the big irregular part of t_n into the middle of a big pebble-free gap. For $n = 1$ the strategy of Duplicator is simple. He copies the move of Spoiler and hence does not place any pebbles on the positions labeled by b in $t_1 = abcb$. Let now $n > 1$. In the first round, Spoiler places $m' \leq m$ pebbles on s_n. Due to the large number of blocks, there exists a big continuous sequence of $\ell' \geq (m+1)^{n-1}$ blocks in s_n without any pebbles. Let such a sequence start after the k-th block and consider the following factorizations of s_n and t_n:

$$\begin{aligned} s_n &= p^k \cdot \qquad p^{\ell'} \qquad \cdot p^{\ell-k-\ell'} \\ t_n &= p^k \cdot p^{\ell-k}\, q\, p^{k+\ell'} \cdot p^{\ell-k-\ell'} \end{aligned}$$

where $p = t_{n-1}b$ and $q = s_{n-1}b$. Duplicator imitates the move of Spoiler on the outer factors of t_n. By induction, Duplicator wins the game on s_{n-1} and t_{n-1} and therefore, for the rest of the game, the blocks p and q cannot be distinguished. By aperiodicity, the different number of blocks does not help Spoiler. Therefore, for $n \geq 2$ we have $L_n \notin \Sigma_n[E](\mathbb{M})$ and hence $\Pi_{n-1}[<](\mathbb{M}) \not\subseteq \Sigma_n[E](\mathbb{M})$. $\square$

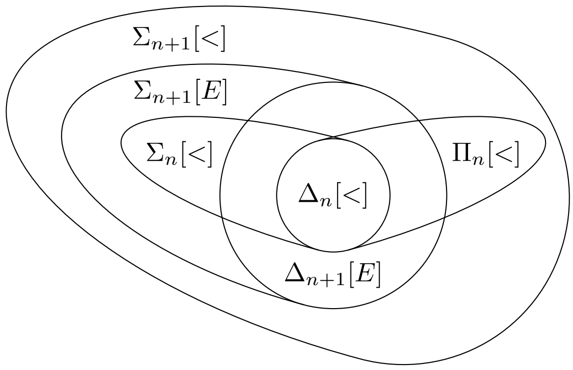
Conclusion

Over words there exist a lot of logical and language-theoretic characterizations of the variety **DA**. In Theorem 5.1 we have shown which of them carry over to traces. In Theorem 5.2 we have shown that the distinction between dependence graphs and partial orders is crucial. Over words, the relations E and $<$ coincide but over traces, as soon as E and $<$ are different, the fragments $\mathrm{FO}^2[<]$ and $\Delta_2[<]$ have strictly more expressive power than their counter-parts $\mathrm{FO}^2[E]$ and $\Delta_2[E]$, respectively. That in general $\mathrm{FO}^2[<](\mathbb{M})$ and $\Delta_2[<](\mathbb{M})$ are incomparable is a consequence of Theorem 7.1 and Theorem 7.2. We summarize the general situation in the following diagram (in which we omitted the suffix $(\mathbb{M})$ in all fragments):



For both, $\mathrm{FO}^2[E]$ and $\mathrm{FO}^2[<]$ we have given equivalent characterizations in terms of temporal logic. For $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{M}]$ – the counter-part of $\mathrm{FO}^2[E]$ – we have shown that the satisfiability problem is in NP whereas the satisfiability problem for $\mathrm{TL}[\mathsf{SF}, \overleftarrow{\mathsf{SF}}, \mathsf{E}\infty]$ – the counter-part of $\mathrm{FO}^2[<]$ – is PSPACE-hard.

Even for fragments in which one can express $<$ by only using quantification over $\mathrm{FO}[E]$-formulas, the distinction between E and $<$ makes a difference. As soon as E and $<$ are different, we have that $\Sigma_n[E](\mathbb{M})$ is a proper subset $\Sigma_n[<](\mathbb{M})$ and that $\Pi_n[<](\mathbb{M})$ is not contained in $\Sigma_{n+1}[E](\mathbb{M})$ for all $n \geq 1$. This is in contrast to the fact that $\Pi_n[<](\mathbb{M}) \subseteq \Sigma_{n+1}[<](\mathbb{M})$. For the fragment $\Sigma_2[E](\mathbb{M})$ we have given a language-theoretic characterization in terms of polynomials in Corollary 3.1.



Several problems remain open. Despite the complexity results in Theorem 6.3 and in Theorem 6.4, completeness results for the satisfiability problems for the fragments $\mathrm{FO}^2[E]$ and $\mathrm{FO}^2[<]$ are still missing. Furthermore, no algebraic characterizations of $\mathrm{FO}^2[<](\mathbb{M})$ and $\Delta_2[<](\mathbb{M})$ are known. Such characterizations could be useful in order to get elementary algorithms for checking membership in those fragments. One of the problems in this task is that $\mathrm{FO}^2[<](\mathbb{M})$ and $\Delta_2[<](\mathbb{M})$ do not form language varieties.

Acknowledgements

We thank Dietrich Kuske and the anonymous referees for various helpful comments.

References

[1] Cartier, P., Foata, D.: *Problèmes combinatoires de commutation et réarrangements*, Lecture Notes in Mathematics, **85**, Springer, 1969. Electronic reedition available at: `http://www-irma.u-strasbg.fr/~foata/paper/ProbComb.pdf`

[2] Diekert, V., Gastin, P.: LTL is expressively complete for Mazurkiewicz traces, *Journal of Computer and System Sciences*, **64**, 2002, 396–418.

[3] Diekert, V., Gastin, P.: Pure future local temporal logics are expressively complete for Mazurkiewicz traces, *Information and Computation*, **204**, 2006, 1597–1619, Conference version in LATIN 2004, LNCS 2976, 170–182, 2004.

[4] Diekert, V., Rozenberg, G., Eds.: *The Book of Traces*, World Scientific, Singapore, 1995.

[5] Eilenberg, S.: *Automata, Languages, and Machines*, vol. B, Academic Press, New York and London, 1976.

[6] Emerson, E. A.: Temporal and Modal Logic, in: *Handbook of Theoretical Computer Science* (J. van Leeuwen, Ed.), vol. B, chapter 16, Elsevier Science Publisher B. V., 1990, 995–1072.

[7] Etessami, K., Vardi, M. Y., Wilke, Th.: First-order logic with two variables and unary temporal logic, *Information and Computation*, 2002, 279–295, Conference version: LICS'97.

[8] Etessami, K., Wilke, Th.: An until hierarchy for temporal logic, *Logic in Computer Science*, 1996.

[9] Gabbay, D., Hodkinson, I., Reynolds, M.: *Temporal Logic: Mathematical Foundations and Computational Aspects*, Clarendon Press, Oxford, 1994.

[10] Gastin, P., Kuske, D.: Satisfiability and model checking for MSO-definable temporal logics are in PSPACE, *Proc. of the 14th Int. Conf. on Concurrency Theory (CONCUR'03)* (R. M. Amadio, D. Lugiez, Eds.), LNCS 2761, Springer, Marseille, France, August 2003.

[11] Gastin, P., Kuske, D.: Uniform satisfiability in PSPACE for local temporal logics over Mazurkiewicz traces, *this volume of Fundamenta Informaticae*, 2007.

[12] Gastin, P., Mukund, M.: An elementary expressively complete temporal logic for Mazurkiewicz traces, *Proc. 29th Int. Colloquium on Automata, Languages and Programming (ICALP'2002)* (P. Widmayer et al., eds.), LNCS 2380, Springer, 2002.

[13] Immerman, N., Kozen, D.: Definability with bounded number of bound variables, *Logic in Computer Science*, 1987.

[14] Kufleitner, M.: *Logical Fragments for Mazurkiewicz Traces: Expressive Power and Algebraic Characterizations*, Dissertation, Institut für Formale Methoden der Informatik, Universität Stuttgart, 2006.

[15] Kufleitner, M.: Polynomials, fragments of temporal logic and the variety DA over traces, *Theoretical Computer Science*, **376**, 2007, 89–100. Conference version in DLT 2006, LNCS 4036, 37–48, 2006.

[16] Mazurkiewicz, A.: *Concurrent program schemes and their interpretations*, DAIMI Rep. PB 78, Aarhus University, Aarhus, 1977.

[17] Pin, J.-É., Weil, P.: Polynominal closure and unambiguous product, *Theory Comput. Syst.*, **30**(4), 1997, 383–422.

[18] Sistla, A. P., Clarke, E.: The complexity of propositional linear time logic, *Journal of the Association for Computing Machinery*, **32**, 1985, 733–749.

[19] Tesson, P., Thérien, D.: Diamonds are forever: The variety DA, *Semigroups, Algorithms, Automata and Languages, Coimbra (Portugal) 2001* (G. M. dos Gomes Moreira da Cunha, P. V. A. da Silva, J.-E. Pin, Eds.), World Scientific, 2002.

[20] Thérien, D., Wilke, Th.: Over words, two variables are as powerful as one quantifier alternation, *STOC*, 1998.

[21] Thiagarajan, P. S., Walukiewicz, I.: An expressively complete linear time temporal logic for Mazurkiewicz traces, *Proc. 12th Annual IEEE Symposium on Logic in Computer Science (LICS'97), Warsaw (Poland)*, 1997.

[22] Thomas, W.: An application of the Ehrenfeucht-Fraïssé game in formal language theory, *Mém. Soc. Math. de France, série 2*, **16**, 1984, 11–21.

[23] Weis, P., Immerman, N.: Structure theorem and strict alternation hierarchy for FO^2 on words, *Proc. 16th EACSL Annual Conference on Computer Science and Logic (CSL'07), Lausanne (Switzerland)* (J. Duparc and Th. Henzinger, Eds.), LNCS 4646, 2007.