Weighted Logics for Traces

Ingmar Meinecke

Institut für Informatik, Universität Leipzig,
D-04109 Leipzig, Germany
meinecke@informatik.uni-leipzig.de

Abstract. We study a quantitative model of traces, *i.e.* trace series which assign to every trace an element from a semiring. We show the coincidence of recognizable trace series with those which are definable by restricted formulas from a weighted logics over traces. We use a translation technique from formulas over words to those over traces, and vice versa. This way, we show also the equivalence of aperiodic and first-order definable trace series.

1 Introduction

Traces as introduced by Mazurkiewicz [19] model concurrency by a global independence relation on a finite alphabet, *i.e.* traces are congruence classes of words modulo the independence relation. A fruitful theory of combinatorics on traces and of trace languages has developed the last twenty years, see [6, 5] for an overview. Droste and Gastin [7] started to explore quantitative aspects of traces a few years ago. They enriched the model with weights from a semiring as it was done for words already in the 1960s by Schützenberger [23]. Droste and Gastin obtained a result in the style of Kleene and Schützenberger, *i.e.* the coincidence of recognizability and a restricted form of rationality. Moreover, they defined and characterized in [8] a weighted concept of aperiodicity for traces. Kuske [16] showed recently the coincidence of recognizable trace series with those recognized by weighted asynchronous cellular automata, both in the non-deterministic and deterministic case. However, a characterization by weighted logics in the lines of Büchi [4] and Elgot [12] was missing even for words. This gap was closed recently by an introduction of weighted logics over words by Droste and Gastin [9]. The semantics of this weighted MSO-logics is a formal power series over words, *i.e.* a function from the free monoid into a semiring. Weighted logics was already extended to trees by Droste and Vogler [10] and to pictures by Mäurer [18].

Naturally, the question arises whether this concept carries over to traces and, therewith, generalizes the results of Droste and Gastin for weighted logics over words [9] on the one hand and the logical characterization of trace languages as done by Ebinger and Muscholl [11] and Thomas [24] on the other hand. Moreover, one could be interested in the execution time of a trace or in the multiplicity of a certain property satisfied by a trace. Such problems can be formulated often better by a logical formula than by a direct construction of a weighted automaton for traces. Therefore, we are interested in weighted logics over traces and in a result that states the coincidence of logically definable and recognizable trace series. Moreover, such a coincidence should be effective in order to open the way to something like quantitative model checking over traces.

D. Grigoriev, J. Harrison, and E.A. Hirsch (Eds.): CSR 2006, LNCS 3967, pp. 235–246, 2006.

Here, we can avoid to repeat the proof of [9] for traces. Instead of this we adapt a technique introduced by Ebinger and Muscholl [11] for their result about the coincidence of definable and recognizable trace languages. There, a formula over traces is translated into an appropriate one over words and vice versa using the lexicographic normal form. This way one is able to transfer the coincidence of definable and recognizable word languages to trace languages. For the weighted case the main problem is to keep the right weighted semantics within the translation of the formulas. Indeed, disjunction and existential quantification result in an addition, whereas conjunction and universal quantification result in multiplication within the underlying semiring. Certainly, these operations are not idempotent in general. Therefore, we are in need of certain "unambiguity" results that will guarantee the right semantics. We obtain such a result for first-order formulas over more general relational structures with a well-order on their elements which is definable by a propositional formula. We apply this result to traces, prove the "translation lemma", and succeed in proving the coincidence of recognizable trace series with trace series defined by restricted monadic second-order formulas. Moreover, for the underlying semiring being either a computable field or being locally finite we will show that decidability results carry over from words to traces. Finally, the coincidence of aperiodic and first-order definable trace series is shown.

For further research the consequences of these results should be explored more in detail. Moreover, application of weighted logics to other models of concurrency like sp-posets [17, 21, 20], MSCs [3], and Σ-DAGs [2] is in work.

2 Basic Concepts

Let Σ be a finite alphabet, Σ^* the free monoid, and $I \subseteq \Sigma^2$ an irreflexive and symmetric relation, called the *independence relation*. Then $D = \Sigma^2 \setminus I$ is reflexive and symmetric and called the *dependence relation*. We define $\sim \subseteq \Sigma^* \times \Sigma^*$ by

$$u \sim v \iff u = w_1abw_2 \wedge v = w_1baw_2 \text{ for } (a,b) \in I \text{ and } w_1, w_2 \in \Sigma^*.$$

By abuse of notation we denote the reflexive and transitive closure of $\sim$ also by $\sim$. Now $\sim$ is a congruence relation on Σ^* and the resulting quotient is called the *trace monoid* $\mathbb{M} = \mathbb{M}(\Sigma, D)$. Its elements are called *traces*. Let $\varphi : \Sigma^* \to \mathbb{M}$ be the canonical epimorphism with $\varphi(w) = [w]$ where $[w]$ is the congruence class of w. For $t \in \mathbb{M}$ there is a prominent representative among $\varphi^{-1}(t)$, the *lexicographic normal form* $\mathrm{LNF}(t)$ of t, *i.e.* the least representative of t with regard to the lexicographic order. The set of all lexicographic normal forms is denoted by LNF. $L \subseteq \mathbb{M}$ is called a *trace language*.

A *semiring* $\mathbb{K} = (K, \oplus, \circ, \mathbb{0}, \mathbb{1})$ is a set K equipped with two binary operations, called *addition* $\oplus$ and *multiplication* $\circ$, such that

1. $(K, \oplus, \mathbb{0})$ is a commutative monoid and $(K, \circ, \mathbb{1})$ a monoid,
2. multiplication distributes over addition: $k \circ (l \oplus m) = (k \circ l) \oplus (k \circ m)$ and $(l \oplus m) \circ k = (l \circ k) \oplus (m \circ k)$ for all $k, l, m \in K$, and
3. $\mathbb{0} \circ k = k \circ \mathbb{0} = \mathbb{0}$ for all $k \in K$.

If the multiplication is commutative we speak of a *commutative semiring*. Examples of semirings are the natural numbers $\mathbb{N} = (\mathbb{N}, +, \cdot, 0, 1)$, the *tropical semiring*

$\mathbb{T} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$, and the *Boolean semiring* $\mathbb{B} = (\{0,1\}, \vee, \wedge, 0, 1)$ which equals the two-element Boolean algebra. For an overview about semirings see [13, 14].

A *formal trace series* or just *trace series* over a trace monoid $\mathbb{M}(\Sigma, D)$ and a semiring $\mathbb{K}$ is a function $T : \mathbb{M}(\Sigma, D) \to \mathbb{K}$. It is often written as a formal sum

$$T = \sum_{t \in \mathbb{M}(\Sigma,D)} (T, t)\, t$$

where $(T, t) = T(t)$. Functions $S : \Sigma^* \to \mathbb{K}$ are called here *word series*. The collection of formal trace series over $\mathbb{M}$ and $\mathbb{K}$ is referred to as $\mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$, and, similarly, $\mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ is defined. For an overview about formal word series see [22, 15, 1].

For the Boolean semiring $\mathbb{B}$ there is a one-to-one correspondence between trace series $T = \sum_{t \in \mathbb{M}(\Sigma,D)} (T, t)\, t$ over $\mathbb{B}$ and their *support* $\mathrm{supp}(T) = \{t \in \mathbb{M}(\Sigma, D) \mid (T, t) \neq \mathbb{0}\}$. Vice versa, a trace language $L \subseteq \mathbb{M}(\Sigma, D)$ corresponds to its *characteristic series* $\mathbb{1}_L$ where

$$(\mathbb{1}_L, t) = \begin{cases} \mathbb{1} & \text{if } t \in L, \\ \mathbb{0} & \text{otherwise.} \end{cases}$$

Hence, formal power series extend formal language theory.

3 Recognizable Trace Series

Let $\mathbb{M}$ be a trace monoid and $\mathbb{K}$ a semiring. Let $\mathbb{K}^{n \times n}$ denote the monoid of $n \times n$-matrices over $\mathbb{K}$ equipped with multiplication. A *recognizable trace series* is a trace series $T \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ such that there are an $n \in \mathbb{N}$, a monoid homomorphism $\mu : \mathbb{M} \to \mathbb{K}^{n \times n}$, $\lambda \in K^{1 \times n}$, and $\gamma \in K^{n \times 1}$ with $(T, t) = \lambda \mu(t) \gamma$ for all $t \in \mathbb{M}$. The triple (λ, μ, γ) is called a *linear representation* of T. For $\varphi : \Sigma^* \to \mathbb{M}$ the canonical epimorphism and $S \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ we define $\varphi^{-1}(S) \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ by $(\varphi^{-1}(S), w) = (S, \varphi(w))$. Furthermore, for $S' \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ we denote by $S'_{|\,\mathrm{LNF}}$ the *restriction* of S' to LNF, *i.e.*

$$(S'_{|\,\mathrm{LNF}}, w) = \begin{cases} (S', w) & w \in \mathrm{LNF}, \\ \mathbb{0} & \text{otherwise.} \end{cases}$$

The following theorem is implicit in [7].

Theorem 3.1. *Let $\mathbb{K}$ be a commutative semiring. Then $S \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ is recognizable iff $S' = \varphi^{-1}(S)_{|\,\mathrm{LNF}} \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ is recognizable.*

The next lemma can be shown as for word series, cf. [1, L. III.1.3].

Lemma 3.2. *Let $\mathbb{K}$ be a semiring and $L \subseteq \mathbb{M}$ a recognizable trace language. Then $\mathbb{1}_L \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ is a recognizable trace series.*

Corollary 3.3. *Let $L_i \subseteq \mathbb{M}$ be recognizable trace languages and $k_i \in \mathbb{K}$ for $i = 1, \ldots, n$. Then $S = \sum_{i=1}^{n} k_i \mathbb{1}_{L_i}$ is a recognizable trace series.*

The last corollary justifies the name *recognizable step function* for a series of the form $S = \sum_{i=1}^{n} k_i \mathbb{1}_{L_i}$ with $L_i \subseteq \mathbb{M}$ recognizable for all $i = 1, \ldots, n$.

4 Definable Trace Series

We represent every trace $t \in \mathbb{M}(\Sigma, D)$ by its *dependence graph*. A dependence graph is (an isomorphism class of) a node-labeled acyclic graph (V, E, l) where V is an at most countable set of nodes[1], $E \subseteq V \times V$ is the edge relation such that (V, E) is acyclic and the induced partial order is well-founded, $l : V \to \Sigma$ is the node-labeling such that

$$(l(x), l(y)) \in D \iff (x, y) \in E \cup E^{-1} \cup id_V \ .$$

A concatenation of dependence graphs is defined by the disjoint union provided with additional edges between nodes with dependent labels, *i.e.*

$$\begin{aligned} &(V_1, E_1, l_1) \cdot (V_2, E_2, l_2) \\ =&(V_1 \dot\cup V_2, E_1 \dot\cup E_2 \dot\cup \{(x, y) \in V_1 \times V_2 \mid (l_1(x), l_2(y)) \in D\}, l_1 \dot\cup l_2) \,. \end{aligned}$$

The monoid $\mathbb{M}(\Sigma, D)$ of finite traces can be identified with the monoid of finite dependence graphs.

Let $t = (V, E, l) \in \mathbb{M}$ and $w = a_1 \ldots a_n \in \Sigma^*$ with $\varphi(w) = t$. Then we represent w as $(V, <, (R_a)_{a\in\Sigma})$ where $<$ is a strict total order on V (the order of positions) and $R_a = \{v \in V \mid l(v) = a\}$ for all $a \in \Sigma$.

Definition 4.1. *The syntax of formulas of* weighted MSO-logic over traces *from* $\mathbb{M}$ *and over a semiring* $\mathbb{K}$ *is given by*

$$\begin{aligned} \Phi ::= k \mid P_a(x) \mid \neg P_a(x) \mid E(x, y) \mid \neg E(x, y) \mid x \in X \mid \neg x \in X \mid \\ \Phi \vee \Psi \mid \Phi \wedge \Psi \mid \exists x.\Phi \mid \exists X.\Phi \mid \forall x.\Phi \mid \forall X.\Phi \end{aligned}$$

with $k \in \mathbb{K}$ *and* $a \in \Sigma$. *This class of formulas is denoted by* MSO$(\mathbb{K}, \mathbb{M})$.

Remark 4.2. The weighted MSO-logic is a generalization of the usual MSO-logic. Weighted MSO-logic differs in two aspects. Firstly, atomic formulas of type k for $k \in K$ are added. Secondly, negation is applied to "unweighted" atomic formulas only. This is due to the fact that a semantics of something like $\neg k$ cannot be defined properly for arbitrary semirings. Hence, we cannot negate neither k nor general MSO-formulas. Thus negation is pulled through the unweighted atomic formulas and conjunction and universal quantification have to be added.

Note 4.3. A weighted MSO-logic for words, denoted by MSO$(\mathbb{K}, \Sigma)$ was defined in [9]. It uses k, $x \leq y$, $P_a(x)$, and $x \in X$ as atomic formulas[2]. Here, we do not include the formula $x = y$ in our syntax because for traces this can be written as

$$\bigvee_{a\in\Sigma} (P_a(x) \wedge P_a(y)) \wedge \neg E(x, y) \wedge \neg E(y, x) \,.$$

[1] Here, we deal with finite objects, *i.e.* finite traces, only. But we stick to the more general case, keeping in mind the possibility to consider infinite objects.

[2] Later on, we will use for words $x < y$ instead of $x \leq y$ which is just a slight technical difference.

A variable is *free* in Φ if it is not within the scope of a quantifier. The collection of all free variables of Φ is denoted by $\mathrm{free}(\Phi)$. Let $\mathcal{V}$ be a finite set of first-order and second-order variables and $t = (V, E, l)$. A $(\mathcal{V}, t)$-assignment σ is a function mapping first-order variables of $\mathcal{V}$ to elements of V and second-order variables of $\mathcal{V}$ to subsets of V. An *update* $\sigma[x \to v]$ for $v \in V$ is defined as $\sigma[x \to v](x) = v$ and $\sigma[x \to v](y) = \sigma(y)$ for all $y \neq x$, and, similarly, for $\sigma[X \to W]$ where $W \subseteq V$. A pair (t, σ) where σ is a $(\mathcal{V}, t)$-assignment will be encoded as a trace over an extended dependence alphabet $\Sigma_{\mathcal{V}} = \Sigma \times \{0,1\}^{\mathcal{V}}$. The new dependence relation $D_{\mathcal{V}}$ is defined by $(a, \bar{x}) D_{\mathcal{V}} (b, \bar{y})$ iff aDb for $a, b \in \Sigma$ and $\bar{x}, \bar{y} \in \{0,1\}^{\mathcal{V}}$. A trace t' over $\Sigma_{\mathcal{V}}$ will be written as a pair (t, σ) where t is the projection of t' over Σ and σ is the projection over $\{0,1\}^{\mathcal{V}}$. Then σ represents a *valid $\mathcal{V}$-assignment* if for any first-order variable $x \in \mathcal{V}$ the x-row of σ contains exactly one 1. Similarly, valid $\mathcal{V}$-assignments are defined for words.

Proposition 4.4. *The trace language $A_{\mathcal{V}} = \{(t, \sigma) \mid \sigma$ is a valid $\mathcal{V}$-assignment$\}$ is recognizable.*

For any formula Φ of MSO we simply write $\Sigma_{\Phi} = \Sigma_{\mathrm{free}(\Phi)}$ and $A_{\Phi} = A_{\mathrm{free}(\Phi)}$. Now we turn to the semantics of our formulas.

Definition 4.5. *Let $\Phi \in \mathrm{MSO}(\mathbb{K}, \mathbb{M})$ and let $\mathcal{V}$ be a finite set of variables with* $\mathrm{free}(\Phi) \subseteq \mathcal{V}$. *The* semantics *of Φ is a formal trace series $[\![\Phi]\!]_{\mathcal{V}} \in \mathbb{K}\langle\langle \mathbb{M}(\Sigma_{\mathcal{V}}^*, D_{\mathcal{V}})\rangle\rangle$ defined as follows: Let $(t, \sigma) \in \mathbb{M}(\Sigma_{\mathcal{V}}, D_{\mathcal{V}})$. If σ is not a valid $\mathcal{V}$-assignment, then $[\![\Phi]\!]_{\mathcal{V}}(t, \sigma) = \mathbb{0}$. Otherwise, we define $[\![\Phi]\!]_{\mathcal{V}}(t, \sigma)$ for $t = (V, E, l)$ inductively as follows:*

- $[\![k]\!]_{\mathcal{V}}(t, \sigma) = k$,
- $[\![P_a(x)]\!]_{\mathcal{V}}(t, \sigma) = \begin{cases} \mathbb{1} & \text{if } l(\sigma(x)) = a, \\ \mathbb{0} & \text{otherwise}, \end{cases}$
- $[\![E(x,y)]\!]_{\mathcal{V}}(t, \sigma) = \begin{cases} \mathbb{1} & \text{if } (\sigma(x), \sigma(y)) \in E, \\ \mathbb{0} & \text{otherwise}, \end{cases}$
- $[\![x \in X]\!]_{\mathcal{V}}(t, \sigma) = \begin{cases} \mathbb{1} & \text{if } \sigma(x) \in \sigma(X), \\ \mathbb{0} & \text{otherwise}, \end{cases}$
- *if Φ is of the form $P_a(x)$, $E(x,y)$, or $x \in X$, then*

$$[\![\neg\Phi]\!]_{\mathcal{V}}(t, \sigma) = \begin{cases} \mathbb{1} & \text{if } [\![\Phi]\!]_{\mathcal{V}}(t, \sigma) = \mathbb{0}, \\ \mathbb{0} & \text{if } [\![\Phi]\!]_{\mathcal{V}}(t, \sigma) = \mathbb{1}, \end{cases}$$

- $[\![\Phi \vee \Psi]\!]_{\mathcal{V}}(t, \sigma) = [\![\Phi]\!]_{\mathcal{V}}(t, \sigma) \oplus [\![\Psi]\!]_{\mathcal{V}}(t, \sigma)$,
- $[\![\Phi \wedge \Psi]\!]_{\mathcal{V}}(t, \sigma) = [\![\Phi]\!]_{\mathcal{V}}(t, \sigma) \circ [\![\Psi]\!]_{\mathcal{V}}(t, \sigma)$,
- $[\![\exists x.\Phi]\!]_{\mathcal{V}}(t, \sigma) = \bigoplus_{v \in V} [\![\Phi]\!]_{\mathcal{V} \cup \{x\}}(t, \sigma[x \to v])$,
- $[\![\exists X.\Phi]\!]_{\mathcal{V}}(t, \sigma) = \bigoplus_{W \subseteq V} [\![\Phi]\!]_{\mathcal{V} \cup \{X\}}(t, \sigma[X \to W])$,
- $[\![\forall x.\Phi]\!]_{\mathcal{V}}(t, \sigma) = \prod_{v \in V} [\![\Phi]\!]_{\mathcal{V} \cup \{x\}}(t, \sigma[x \to v])$,
- $[\![\forall X.\Phi]\!]_{\mathcal{V}}(t, \sigma) = \prod_{W \subseteq V} [\![\Phi]\!]_{\mathcal{V} \cup \{X\}}(t, \sigma[X \to W])$.

where we fix some order both on V and on $\mathfrak{P}(V)$ so that the last two products are defined even if $\mathbb{K}$ is not commutative. We simply write $[\![\Phi]\!]$ for $[\![\Phi]\!]_{\mathrm{free}(\Phi)}$.

If Φ is a sentence, then $[\![\Phi]\!] \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$. As usual, the semantics of some formula Φ depends on the free variables only. We call $S = \sum_{i=1}^{n} k_i \mathbb{1}_{L_i}$ a *definable step function* if the languages L_i are definable trace languages, or word languages respectively, for all $i = 1, \ldots, n$. For words and traces the notions of recognizable and definable step functions coincide because of the results of Büchi & Elgot [4, 12] and Ebinger & Muscholl [11].

Definition 4.6. *A formula $\Phi \in \mathrm{MSO}(\mathbb{K}, \Sigma)$ or $\Phi \in \mathrm{MSO}(\mathbb{K}, \mathbb{M})$ is called* restricted*, if it contains no universal quantification of second-order $\forall X.\Psi$, and whenever Φ contains a universal first-order quantification $\forall x.\Psi$, then $[\![\Psi]\!]$ is a definable step function.*

Remark 4.7. Droste and Gastin [9] had to use restricted MSO-formulas over words to preserve recognizability of the defined series. For universal FO-quantification $\forall x.\Psi$ they required $[\![\Psi]\!] = \sum_{i=1}^{n} k_i \mathbb{1}_{L_i}$ being a recognizable step function. Since we define a class of formulas, we favor to speak of the logical counterpart, *i.e.* definable step functions.

$\mathrm{RMSO}(\mathbb{K}, \mathbb{M})$ is the class of all restricted formulas from $\mathrm{MSO}(\mathbb{K}, \mathbb{M})$. Moreover, let $\mathrm{REMSO}(\mathbb{K}, \mathbb{M})$ contain all restricted *existential* formulas $\Phi \in \mathrm{RMSO}(\mathbb{K}, \mathbb{M})$, *i.e.* Φ is of the form $\exists X_1.\exists X_2 \ldots \exists X_n.\Psi$ with $\Psi \in \mathrm{RMSO}(\mathbb{K}, \mathbb{M})$ containing no second-order quantification anymore. FO and RFO denote the classes of first-order formulas and restricted first-order formulas, respectively. Similar notations are used for formulas over words.

5 Characteristic Series of FO-Definable Languages

Let $\mathcal{C}$ be a class of finite relational structures. We define formulas of a *weighted* MSO-*logic over* $\mathcal{C}$ in the same manner as for traces, *i.e.* atomic formulas are beside k for $k \in K$, and $x \in X$ the relation symbols of $\mathcal{C}$ and possibly $x = y$, and negation is applied to atomic formulas only. The formulas are provided with the appropriate semantics $S : \mathcal{C} \to \mathbb{K}$ as for traces, *i.e.* atomic formulas are interpreted by the characteristic series of the defined language (a valid $\mathcal{V}$-assignment provided) and the semantics of composed formulas is given as above. Similarly, an unweighted MSO-logic for $\mathcal{C}$ is defined with a semantics of languages $L \subseteq \mathcal{C}$. Moreover, we suppose that there is a propositional formula $\Omega(x, y)$ (*i.e.* one without any quantifier) with free FO-variables x, y such that for any structure $t \in \mathcal{C}$ the binary relation defined by Ω is a linear order on the elements of t. We say that $\mathcal{C}$ has a *simply definable linear order*.

Let $L \subseteq \mathcal{C}$ be a language of $\mathcal{C}$ and $\overline{L} = \mathcal{C} \setminus L$ the complement of L.

Lemma 5.1. *Let $\mathcal{C}$ be a class of finite relational structures with a simply definable linear order. Let $L = L(\Phi)$ be defined by an FO-formula Φ. Then both $\mathbb{1}_L$ and $\mathbb{1}_{\overline{L}}$ are definable in RFO.*

Proof (sketch). Let $L \subseteq \mathcal{C}$ be defined by Φ. We proceed by induction giving for each FO-formula Φ RFO-formulas Φ^+ and Φ^- such that $[\![\Phi^+]\!] = \mathbb{1}_L$ and $[\![\Phi^-]\!] = \mathbb{1}_{\overline{L}}$. The interesting cases are $(\exists x.\Phi)^+$ and $(\forall x.\Phi)^-$. Therefore, we choose the "smallest" element that satisfies Φ^+, and Φ^- respectively, by using $\Omega(x, y)$ defining the linear order $\leq_\Omega$. Since $\Omega(x,y)$ is a propositional formula we can already define $\Omega^+(x,y)$. Now we put

$$\big(\exists x.\Phi(x)\big)^+ = \exists x.\Big(\Phi^+(x) \wedge \forall y.\big(\Phi^-(y) \vee \Omega^+(x,y)\big)^+\Big) .$$

This is an RFO-formula. Indeed, $\Phi^+(x)$ is an RFO-formula by induction hypothesis and so are $\Phi^-(y)$ and $\Omega^+(x,y)$. Moreover, $\big(\Phi^-(y) \vee \Omega^+(x,y)\big)^+$ defines a definable step function by induction hypothesis. Since we choose the "smallest" element x satisfying Φ we get for a valid $\mathcal{V}$-assignment

$$[\![\, (\exists x.\Phi(x))^+ \,]\!]_{\mathcal{V}}(t,\sigma) = \begin{cases} \mathbb{1} & \text{if there is an } v \text{ such that } (t,\sigma[x \to v]) \text{ satisfies } \Phi, \\ \mathbb{0} & \text{otherwise.} \end{cases}$$

Similarly we proceed for $(\forall x.\Phi)^-$. □

Corollary 5.2. *Let L be an FO-definable trace language. Then $\mathbb{1}_L$ is RFO-definable.*

Proof. For a fixed linear order $\preceq$ on the alphabet Σ put

$$\Omega(x,y) = \bigvee_{(a,b)\in\prec} \Big(P_a(x) \wedge P_b(y)\Big) \vee \bigvee_{a\in\Sigma} \Big(P_a(x) \wedge P_a(y) \wedge \neg E(y,x)\Big)$$

and apply Lemma 5.1. □

6 The Coincidence of Recognizable and Definable Trace Series

We will follow the ideas of the proof as given for trace languages, cf. [5, pp. 497–505] and use the result of the previous section.

Lemma 6.1. *Let $\mathbb{K}$ be a commutative semiring, $\varphi : \Sigma^* \to \mathbb{M}$ the canonical epimorphism, and $T \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ a trace series. The following are equivalent:*

(i) *T is definable in* RMSO*, and* REMSO *respectively.*
(ii) *$\varphi^{-1}(T) \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ is definable in* RMSO*, and* REMSO *respectively.*
(iii) *$S = \varphi^{-1}(T)_{|\,\mathrm{LNF}} \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ is definable in* RMSO*, and* REMSO *respectively.*

Proof. *(i) $\Longrightarrow$ (ii)* Let $T \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ be defined by some sentence Ψ. Let $t = (V,E,l)$ be any trace and $w \in \Sigma^*$ with $\varphi(w) = t$. We have for $v_1, v_2 \in V$ that $(v_1,v_2) \in E$ iff $v_1 < v_2$ in w and $(l(v_1), l(v_2)) \in D$. Thus, replacing every atomic formula $E(x,y)$ in Ψ by the propositional formula

$$x < y \wedge \bigvee_{(a,b)\in D} (P_a(x) \wedge P_b(y)) \tag{1}$$

yields an new sentence $\tilde{\Psi}$. One shows easily $([\![\, \Psi \,]\!], t) = ([\![\, \tilde{\Psi} \,]\!], w)$ for each $t \in \mathbb{M}, w \in \Sigma^*$ with $\varphi(w) = t$ by structural induction.

It still remains to show that for $\Psi \in \mathrm{RMSO}(\mathbb{K}, \mathbb{M})$ also $\tilde{\Psi} \in \mathrm{RMSO}(\mathbb{K}, \Sigma)$. Clearly, if Ψ contains no universal second-order quantification neither does $\tilde{\Psi}$. Consider $\Psi = \forall x.\Phi$ with $[\![\Phi]\!] = \sum_{i=1}^{n} k_i \mathbb{1}_{L_i}$ for definable and, hence, recognizable trace languages $L_i \subseteq \mathbb{M}$ for $i = 1, \ldots, n$. As we have shown, $[\![\tilde{\Psi}]\!](w, \sigma) = [\![\Psi]\!](t, \sigma)$ where $\varphi(w) = t$. Consider the word series $S = \sum_{i=1}^{n} k_i \mathbb{1}_{\varphi^{-1}(L_i)}$. It is a recognizable and, hence, definable step function over words since $\varphi^{-1}(L_i)$ is recognizable for $i = 1, \ldots, n$. Moreover, $(w, \sigma) \in \varphi^{-1}(L_i)$ for some i implies that σ is a valid $\mathcal{V}$-assignment. For (w, σ) with σ a valid $\mathcal{V}$-assignment we have

$$S(w,\sigma) = \bigoplus_{i=1}^{n} k_i \mathbb{1}_{\varphi^{-1}(L_i)}(w,\sigma) = \bigoplus_{\{i \mid w \in \varphi^{-1}(L_i)\}} k_i = \bigoplus_{\{i \mid t = \varphi(w) \in L_i\}} k_i = [\![\Phi]\!](t,\sigma).$$

Hence, $S = \tilde{\Phi}$ is a definable step function. Thus, if Ψ is reduced so is $\tilde{\Psi}$. Moreover, for $\Psi \in \mathrm{REMSO}(\mathbb{K}, \mathbb{M})$ also $\tilde{\Psi} \in \mathrm{REMSO}(\mathbb{K}, \Sigma)$ because (1) is a propositional formula.

(ii) $\Longrightarrow$ *(iii)* Let $\varphi^{-1}(T) \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ be defined by an RMSO-formula Φ, let $S = \varphi^{-1}(T)_{|\,\mathrm{LNF}}$, and let $\preceq$ be the fixed order on Σ. The language LNF of all lexicographic normal forms is defined by the FO-sentence

$$\forall i \forall k. \Big[(i \le k) \longrightarrow \Big(l(i) \preceq l(k) \vee \exists j. \big(i \le j < k \wedge (l(j), l(k)) \in D \big) \Big) \Big]$$

where implication $\longrightarrow$, $l(i) \preceq l(k)$, and $(l(j), l(k)) \in D$ are obvious abbreviations. By Corollary 5.2, there is an RFO-formula Λ with $[\![\Lambda]\!] = \mathbb{1}_{\mathrm{LNF}}$. Hence, $S = [\![\Phi \wedge \Lambda]\!]$. If $\Phi = \exists X_1 \ldots \exists X_n.\Psi$ is an REMSO-formula, then S is defined by the REMSO-formula $\exists X_1 \ldots \exists X_n.(\Psi \wedge \Lambda)$ because Λ is from RFO.

(iii) $\Longrightarrow$ *(i)* Let $S = \varphi^{-1}(T)_{|\,\mathrm{LNF}}$ be defined by $\Phi \in \mathrm{RMSO}$. Then $\mathrm{supp}(S) \subseteq \mathrm{LNF}$. We replace every atomic formula $x < y$ in Φ by a new formula $\mathrm{lex}(x, y)$ that models the order in the lexicographic normal form, *i.e.* for every t and a valid $\mathcal{V}$-assignment σ we get $(t, \sigma) \models \mathrm{lex}(x, y)$ iff $\sigma(x) < \sigma(y)$ in $\mathrm{LNF}(t)$. The formula $\mathrm{lex}(x, y)$ can be found in the literature (cf.[5, p. 502]) and is an FO-formula because transitive closure of E can be expressed for traces in FO. Hence, we apply Corollary 5.2 and obtain an RFO-formula $\mathrm{lex}^+(x, y)$ defining $\mathbb{1}_{L(\mathrm{lex}(x,y))}$, and similarly $\mathbb{1}_{L(\neg\,\mathrm{lex}(x,y))}$ can be defined by an RFO-formula $\mathrm{lex}^-(x, y)$. Let Ψ be any formula over words and $\tilde{\Psi}$ the formula over traces obtained from Ψ by replacing every occurence of the atomic formula $x < y$ by $\mathrm{lex}^+(x, y)$, and any occurence of $\neg(x < y)$ by $\mathrm{lex}^-(x, y)$. Then we get for every trace t and every valid $\mathcal{V}$-assignment σ

$$[\![\tilde{\Psi}]\!]_{\mathcal{V}}(t, \sigma) = [\![\Psi]\!]_{\mathcal{V}}(\mathrm{LNF}(t), \sigma). \tag{2}$$

We still have to show that $\tilde{\Psi}$ is restricted. For an RMSO-formula $\forall x.\Psi$ over words $[\![\Psi]\!]_{\mathcal{V}}$ is a definable and recognizable step function, *i.e.* $[\![\Psi]\!]_{\mathcal{V}}(t, \sigma) = \sum_{i=1}^{n} k_i \mathbb{1}_{L_i}$ with recognizable word languages L_i ($i = 1, \ldots, n$). Now we have by Equation (2) $[\![\tilde{\Psi}]\!]_{\mathcal{V}} = \sum_{i=1}^{n} k_i \mathbb{1}_{\varphi(L_i \cap \mathrm{LNF})}$. By [6, Thm. 6.3.12] the trace languages $\varphi(L_i \cap \mathrm{LNF})$ are recognizable languages. Thus $[\![\tilde{\Psi}]\!]_{\mathcal{V}}$ is a recognizable and, hence, definable step function. Hence, if $S \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ with $\mathrm{supp}(S) \subseteq \mathrm{LNF}$ is defined by some sentence Φ from RMSO then $T = \varphi(S) \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ is defined by the RMSO-sentence $\tilde{\Phi}$. Certainly, if Φ is in REMSO so is $\tilde{\Phi}$. □

Remark 6.2. The translations of formulas over traces to those over words and vice versa as given in the proof of Lemma 6.1 are effective.

By Lemma 6.1 and the result for word series [9, Thm. 3.7] we get:

Theorem 6.3. *Let $\mathbb{K}$ be a commutative semiring and $T \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$. The following are equivalent:*

(i) T is recognizable,
(ii) T is definable by some sentence of RMSO*,*
(iii) T is definable by some sentence of REMSO*.*

Proof. Let T be recognizable. Then $S = \varphi^{-1}(T)_{|\,\mathrm{LNF}}$ is a recognizable word series by Theorem 3.1. Hence, S is definable in RMSO and REMSO by the main result of [9]. By Lemma 6.1, T is definable in RMSO and REMSO, respectively.

Conversely, let T be definable in RMSO and REMSO, respectively. Then the series $\varphi^{-1}(T)_{|\,\mathrm{LNF}}$ is definable by Lemma 6.1, hence recognizable by [9]. Now, Theorem 3.1 shows the recognizability of T. □

Example 6.4. Let $\mathbb{K} = (\mathbb{N} \cup \{-\infty\}, \max, +, -\infty, 0)$. We show that $H \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ mapping every $t \in \mathbb{M}$ to $\mathrm{height}(t)$, *i.e.* the length of the longest chain in t, is recognizable. Let

$$\mathrm{chain}(X) = \forall x, y \in X. \big(x = y \vee (x, y) \in E^+ \vee (y, x) \in E^+\big)$$

be an unweighted formula stating that X is a chain. Since transitive closure of E can be expressed for traces by an FO-formula (cf. [5, p. 501]), $\mathrm{chain}(X)$ is an FO-formula. By Corollary 5.2, there is an RFO-formula $\mathrm{chain}(X)^+$ defining $\mathbb{1}_{L(\mathrm{chain}(X))}$. Moreover, the formula

$$\mathrm{card}(X) = \forall x. \big((x \in X \longrightarrow 1) \wedge (\neg x \in X \longrightarrow 0)\big)$$

has the semantics $|X|$ over $\mathbb{K}$. Hence, $H = \sum_{t \in \mathbb{M}} \mathrm{height}(t)\, t$ is defined by

$$\varPhi = \exists X.\, \mathrm{chain}(X)^+ \wedge \mathrm{card}(X)\,.$$

By Theorem 6.3, $H \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ is recognizable.

7 Some Notes About Decidability

Given a weighted MSO-formula $\varPhi$ over traces, there are two immediate questions:

- It is decidable whether $\varPhi$ is an RMSO-formula?
- If $\varPhi$ is in RMSO, can we effectively compute the semantics of $\varPhi$, *i.e.* compute $([\![\varPhi]\!], t)$ for every trace $t \in \mathbb{M}$?

Droste and Gastin [9] answer these questions for weighted logics over words where the underlying semiring is either a computable field or a locally finite semiring. We cannot expect to do any better. By the effective translation of formulas the results carry over.

Proposition 7.1. *Let $\mathbb{K}$ be a computable field, and let $\Phi \in \mathrm{MSO}(\mathbb{K}, \mathbb{M})$. It is decidable whether Φ is reduced. In this case we can compute effectively for every trace $t \in \mathbb{M}$ the coefficient $(\llbracket \Phi \rrbracket, t)$ in a uniform way.*

Corollary 7.2. *Let $\mathbb{K}$ be a computable field, and let $\Phi, \Psi \in \mathrm{RMSO}(\mathbb{K}, \mathbb{M})$. It is decidable whether $\llbracket \Phi \rrbracket$ has empty support, whether $\llbracket \Phi \rrbracket = \llbracket \Psi \rrbracket$, and whether $\llbracket \Phi \rrbracket$ and $\llbracket \Psi \rrbracket$ differ for finitely many traces only.*

Recall that a semiring $\mathbb{K}$ is *locally finite*, if each finitely generated subsemiring of $\mathbb{K}$ is finite. A monoid M is locally finite, if each finitely generated submonoid of M is finite. Clearly, a semiring $(K, \oplus, \circ, \mathbb{0}, \mathbb{1})$ is locally finite iff both monoids $(K, \oplus, 0)$ and $(K, \circ, \mathbb{1})$ are locally finite. Now even every MSO-definable trace series is recognizable as it is true for word series [9, Thm. 6.4].

Theorem 7.3. *Let $\mathbb{K}$ be a locally finite commutative semiring and $T \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$. Then the following are equivalent:*

(i) T is definable in MSO.
(ii) T is recognizable.

Proposition 7.4. *Let $\mathbb{K}$ be a locally finite commutative semiring and $\Phi \in \mathrm{MSO}(\mathbb{K}, \mathbb{M})$. Then the coefficient $(\llbracket \Phi \rrbracket, t)$ can be computed effectively for every $t \in \mathbb{M}$ in a uniform way. Moreover, it is decidable*

(a) whether two $\mathrm{MSO}(\mathbb{K}, \mathbb{M})$-formulas Φ and Ψ satisfy $\llbracket \Phi \rrbracket = \llbracket \Psi \rrbracket$, and
(b) whether an $\mathrm{MSO}(\mathbb{K}, \mathbb{M})$-formula Φ satisfies $\mathrm{supp}(\llbracket \Phi \rrbracket) = \mathbb{M}$.

8 FO-Definable Trace Series

By considering Lemma 6.1 and its proof we get:

Lemma 8.1. *Let $\mathbb{K}$ be a commutative semiring and $\varphi : \Sigma^* \to \mathbb{M}$ the canonical epimorphism. The following are equivalent:*

(i) $T \in \mathbb{K}\langle\langle \mathbb{M} \rangle\rangle$ is definable in RFO *(in* FO, *respectively).*
(ii) $\varphi^{-1}(T) \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ is definable in RFO *(in* FO, *respectively).*
(iii) $\varphi^{-1}(T)_{|\,\mathrm{LNF}} \in \mathbb{K}\langle\langle \Sigma^* \rangle\rangle$ is definable in RFO *(in* FO, *respectively).*

Droste and Gastin showed that the classes of aperiodic word series, RFO-definable and FO-definable word series coincide for commutative weakly bi-aperiodic semirings [9, Thm. 7.8]. A monoid M is *weakly aperiodic*, if for each $m \in M$ there is an $n \geq 0$ such that $m^n = m^{n+1}$. M is *aperiodic* if there is an $n \geq 0$ such that $m^n = m^{n+1}$ for all $m \in M$. A semiring $\mathbb{K}$ is *weakly bi-aperiodic*, if both $(K, \oplus)$ and $(K, \circ)$ are weakly aperiodic monoids. Note that every commutative weakly aperiodic semiring K is locally finite. Let $S \in \mathbb{K}\langle\langle M \rangle\rangle$ be a recognizable series over an arbitrary monoid M. Then S is called *aperiodic* if there exists a representation $S = (\lambda, \mu, \gamma)$ with $\mu(M)$ aperiodic, *i.e.* there is some integer $n \geq 0$ such that $\mu(u^n) = \mu(u^{n+1})$ for all $u \in M$. A recognizable series S is *weakly aperiodic* if there exists some integer $n \geq 0$ such that $(S, uv^n w) = (S, uv^{n+1} w)$ for all $u, v, w \in M$. Clearly, every aperiodic series is also weakly aperiodic. The converse is true for locally finite semirings as already Droste and Gastin noted [8, Sect. 3].

Lemma 8.2. *Let $\mathbb{K}$ be a locally finite semiring, M a finitely generated monoid, and $S \in \mathbb{K}\langle\langle M\rangle\rangle$ a recognizable series. Then S is aperiodic iff S is weakly aperiodic.*

Using Lemma 8.2, we can clarify the relation between aperiodic trace and aperiodic word series.

Proposition 8.3. *Let $\mathbb{K}$ be a locally finite semiring. Then $T \in \mathbb{K}\langle\langle \mathbb{M}\rangle\rangle$ is aperiodic iff $\varphi^{-1}(T) \in \mathbb{K}\langle\langle \Sigma^*\rangle\rangle$ is aperiodic.*

Proof. If T has the aperiodic representation $T = (\lambda, \mu, \gamma)$ such that there is an $r \in \mathbb{N}$ with $\mu(t^r) = \mu(t^{r+1})$ for all $t \in \mathbb{M}$. Then $\varphi^{-1}(T)$ has the aperiodic representation $(\lambda, \mu \circ \varphi, \gamma)$. Vice versa, let $\varphi^{-1}(T)$ be aperiodic and, hence, also weakly aperiodic, *i.e.* there is some $r \in \mathbb{N}$ with $(\varphi^{-1}(T), uv^r w) = (\varphi^{-1}(T), uv^{r+1}w)$ for all $u, v, w \in \Sigma^*$. By t' we denote some representative for the trace t. Now we have

$$\begin{aligned}(T, xt^r y) &= (\varphi^{-1}(T), (xt^r y)') = (\varphi^{-1}(T), x't'^r y') = (\varphi^{-1}(T), x't'^{r+1}y') \\ &= (\varphi^{-1}(T), (xt^{r+1}y)') = (T, xt^{r+1}y)\end{aligned}$$

and, thus, T is weakly aperiodic. By Lemma 8.2, T is aperiodic. □

Theorem 8.4. *Let $\mathbb{K}$ be a commutative, weakly bi-aperiodic semiring and $T \in \mathbb{K}\langle\langle \mathbb{M}\rangle\rangle$. Then the following are equivalent:*

(i) T *is aperiodic.*
(ii) T *is weakly aperiodic.*
(iii) T *is* RFO-*definable.*
(iv) T *is* FO-*definable.*

Proof. Recall that a commutative, weakly bi-aperiodic semiring $\mathbb{K}$ is locally finite. Then the equivalence of (i) and (ii) is clear by Lemma 8.2. Now, let T be aperiodic. Then $\varphi^{-1}(T) \in \mathbb{K}\langle\langle \Sigma^*\rangle\rangle$ is aperiodic by Proposition 8.3. Now, [9, Thm. 7.8] implies RFO- and FO-definability of $\varphi^{-1}(T)$. By Lemma 8.1, T is RFO- and FO-definable, respectively. The converse direction follows similarly. □

Acknowledgements. The author would like to thank Dietrich Kuske for a lot of fruitful discussions.

References

1. J. Berstel and C. Reutenauer. *Rational Series and Their Languages*, volume 12 of *EATCS Monographs on Theoret. Comp. Sc.* Springer, 1988.
2. B. Bollig. On the expressiveness of asynchronous cellular automata. In *Proceedings of the 15th International Symposium on Fundamentals of Computation Theory (FCT'05), Lübeck, Germany, August 2005*, volume 3623 of *Lecture Notes in Comp. Sc.*, pages 528–539. Springer, 2005.
3. B. Bollig and M. Leucker. Message-passing automata are expressively equivalent to EMSO logic. *Theoret. Comp. Sc.*, 2006. to appear.

4. J.R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, (6):66–92, 1960.
5. V. Diekert and Y. Métivier. Partial commutation and traces. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages, Beyond Words*, volume 3, chapter 8, pages 457–534. Springer, 1997.
6. V. Diekert and G. Rozenberg, editors. *The Book of Traces*. World Scientific, 1995.
7. M. Droste and P. Gastin. The Kleene-Schützenberger theorem for formal power series in partially commuting variables. *Information and Computation*, 153:47–80, 1999.
8. M. Droste and P. Gastin. On aperiodic and star-free formal power series in partially commuting variables. In *Formal Power Series and Algebraic Combinatorics (Moscow 2000)*, pages 158–169. Springer Berlin, 2000.
9. M. Droste and P. Gastin. Weighted automata and weighted logics. In *Automata, Languages and Programming (32nd ICALP, Lissabon)*, volume 3580 of *Lecture Notes in Comp. Sc.*, pages 513–525. Springer, 2005.
10. M. Droste and H. Vogler. Weighted tree automata and weighted logics. submitted, 2005.
11. W. Ebinger and A. Muscholl. Logical definability on infinite traces. *Theoret. Comp. Sc.*, (154):67–84, 1996.
12. C.C. Elgot. Decision problems of finite automata design and related arithmetics. *Trans. Amer. Math. Soc.*, (98):21–52, 1961.
13. J.S. Golan. *Semirings and their Applications*. Kluwer Academic Publishers, 1999.
14. U. Hebisch and H.J. Weinert. *Semirings: Algebraic Theory and Application*. World Scientific, 1999.
15. W. Kuich and A. Salomaa. *Semirings, Automata, Languages*, volume 5 of *EATCS Monographs on Theoret. Comp. Sc.* Springer, 1986.
16. D. Kuske. Weighted asynchronous cellular automata. In *STACS 2006*, volume 3884 of *Lecture Notes in Comp. Sc.*, pages 685–696. Springer, 2006.
17. D. Kuske and I. Meinecke. Branching automata with costs – a way of reflecting parallelism in costs. *Theoret. Comp. Sc.*, 328:53–75, 2004.
18. I. Mäurer. Weighted picture automata and weighted logics. In *STACS 2006*, Lecture Notes in Comp. Sc. Springer, 2006. to appear.
19. A. Mazurkiewicz. Trace theory. In *Advances in Petri Nets 1986, Part II on Petri Nets: Applications and Relationships to Other Models of Concurrency*, volume 255 of *Lecture Notes in Comp. Sc.*, pages 279–324. Springer, 1987.
20. I. Meinecke. *Weighted Branching Automata – Combining Concurrency and Weights*. Dissertation, Technische Universität Dresden, Germany, December 2004.
21. I. Meinecke. The Hadamard product of sequential-parallel series. *J. of Automata, Languages and Combinatorics*, 10(2), 2005. To appear.
22. A. Salomaa and M. Soittola. *Automata-Theoretic Aspects of Formal Power Series*. Texts and Monographs in Computer Science. Springer, 1978.
23. M.P. Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.
24. W. Thomas. On logical definability of trace languages. In V. Diekert, editor, *Proceedings of a workshop of the ESPRIT BRA No 3166: Algebraic and Syntactic Methods in Computer Science (ASMICS) 1989*, pages 172–182. Technical University of Munich, 1990.