Weighted Automata and Monadic Second Order Logic

Nadia Labai
Faculty of Computer Science
Technion–Israel Institute of Technology
nadia@cs.technion.ac.il

Johann A. Makowsky
Faculty of Computer Science
Technion–Israel Institute of Technology
janos@cs.technion.ac.il

Let $\mathscr{S}$ be a commutative semiring. M. Droste and P. Gastin have introduced in 2005 weighted monadic second order logic **WMSOL** with weights in $\mathscr{S}$. They use a syntactic fragment **RMSOL** of **WMSOL** to characterize word functions (power series) recognizable by weighted automata, where the semantics of quantifiers is used both as arithmetical operations and, in the boolean case, as quantification.

Already in 2001, B. Courcelle, J.Makowsky and U. Rotics have introduced a formalism for graph parameters definable in Monadic Second order Logic, here called **MSOLEVAL** with values in a ring $\mathscr{R}$. Their framework can be easily adapted to semirings $\mathscr{S}$. This formalism clearly separates the logical part from the arithmetical part and also applies to word functions.

In this paper we give two proofs that **RMSOL** and **MSOLEVAL** with values in $\mathscr{S}$ have the same expressive power over words. One proof shows directly that **MSOLEVAL** captures the functions recognizable by weighted automata. The other proof shows how to translate the formalisms from one into the other.

1 Introduction

Let f be a function from relational structures of a fixed relational vocabulary τ into some field, ring, or a commutative semiring $\mathscr{S}$ which is invariant under τ-isomorphisms. $\mathscr{S}$ is called a *weight structure*. In the case where the structures are graphs, such a function is called a graph parameter, or, if $\mathscr{S}$ is a polynomial ring, a graph polynomial. In the case where the structures are words, it is called a word function.

The study of definability of graph parameters and graph polynomials in Monadic Second Order Logic **MSOL** was initiated in [6] and further developed in [22, 20]. For a weight structure $\mathscr{S}$ we denote the set of functions of τ-structures definable in **MSOL** by $\mathbf{MSOLEVAL}(\tau)_{\mathscr{S}}$, or if the context is clear, just by $\mathbf{MSOLEVAL}_{\mathscr{S}}$. The original purpose for studying functions in $\mathbf{MSOLEVAL}_{\mathscr{S}}$ was to prove an analogue to Courcelle's celebrated theorem for polynomial rings as weight structures, which states that graph parameters $f \in \mathbf{MSOLEVAL}_{\mathscr{S}}$ are computable in linear time for graphs of fixed tree-width, [6], and various generalizations thereof. **MSOLEVAL** can be seen as an analogue of the *Skolem elementary functions* aka *lower elementary functions*, [25, 26], adapted to the framework of *meta-finite model theory* as defined in [15].

In [8] a different formalism to define $\mathscr{S}$-valued word functions was introduced, which the authors called *weighted monadic second order logic WMSOL*, and used a fragment, **RMSOL**, of it to prove that a word function is recognized by a weighted automaton iff it is definable in **RMSOL**. This can be seen as an analogue of the Büchi-Elgot-Trakhtenbrot Theorem characterizing regular languages for the case of weighted (aka multiplicity) automata.

Gabriele Puppis, Tiziano Villa (Eds.): Fourth International Symposium on Games, Automata, Logics and Formal Verification
EPTCS 119, 2013, pp. 122–135, doi:10.4204/EPTCS.119.12

© N. Labai & J.A. Makowsky
This work is licensed under the Creative Commons Attribution License.

Main results

Our main results explore various features of the two formalisms **MSOLEVAL** and **RMSOL** for word functions with values in a semiring $\mathcal{S}$. In the study of **MSOLEVAL** we show how *model theoretic tools* can be used to characterize the word functions in **MSOLEVAL** as the fuctions recognizable by weigthed automata. This complements the automata theoretic approach used in the study of weighted automata, [9, 11]. In particular, we give two proofs that **RMSOL** and **MSOLEVAL** with values in a semiring $\mathcal{S}$ have the same expressive power over words. To see this we show the following for a word function f with values in $\mathcal{S}$:

(i) If f is definable in **MSOLEVAL**, it is contained in a finitely generated stable semimodule of word functions, Theorem 11.

(ii) If f is recognizable by some weighted automaton, it is definable in **MSOLEVAL**, the "if" direction of Theorem 8.

(iii) If f is definable in **RMSOL**, we can translate it, using Lemma 15, into an expression in **MSOLEVAL**, Theorem 16.

(iv) If f is definable in **MSOLEVAL**, we can, again using Lemma 15, translate it into an expression in **RMSOL**, Theorem 17.

Items (i) and (ii) together with a classical characterization of recognizable word functions in terms of finitely generated stable semimodules, Theorem 10, cf. [1, 17, 13], give us a direct proof that **MSOLEVAL** captures the functions recognizable by weighted automata. To prove item (i) we rely on and extend results about **MSOLEVAL** from [22, 14, 18].

Items (iii) and (iv) together show how to translate the formalisms **RMSOL** and **MSOLEVAL** into each other. Lemma 15 also shows how the fragment **RMSOL** of the weighted logic **WMSOL** comes into play.

The point of separating (i) and (ii) from (ii) and (iv) and giving *two* proofs of Theorem 8 is to show that the model theoretic methods developed in the 1950ties and further developed in [22] suffice to characterize the functions recognized by weighted automata.

Background and outline of the paper

We assume the reader is familiar with Monadic Second Order Logic and Automata Theory as described in [12, 1] or similar references. In Section 2 we introduce **MSOLEVAL** by example, which suffices for our purposes. A full definition is given in Appendix 2.2. In Section 3 we show that the word functions which are recognizable by a weighted automaton are exactly the word functions definable in **MSOLEVAL**. In Section 4 we give the exact definitions of **WMSOL** and **RMSOL**, and present translations between **MSOLEVAL** and **RMSOL** in both directions. In Section 5 we draw our conclusions.

2 Definable word functions

Let $\mathcal{S}$ be a commutative semiring. We denote structures over a finite relational signature (aka vocabulary) τ by $\mathcal{A}$ and their underlying universe by A. The class of functions in **MSOLEVAL**$_{\mathcal{S}}$ consists of the functions which map relational structures into $\mathcal{S}$, and which are definable in Monadic Second Order Logic **MSOL**. The functions in **MSOLEVAL**$_{\mathcal{S}}$ are represented as terms associating with each τ-structure $\mathcal{A}$ a polynomial $p(\mathcal{A},\bar{X}) \in \mathcal{S}[\bar{X}]$. The class of such polynomials is defined inductively where monomials are products of constants in $\mathscr{S}$ and indeterminates in $\bar{X}$ and the product ranges over elements a of A which satisfy an **MSOL**-formula $\phi(a)$. The polynomials are then defined as sums of monomials where the sum ranges over *unary* relations $U \subseteq A$ satisfying an **MSOL**-formula $\psi(U)$. The word functions are obtained by substituting elements of $\mathscr{S}$ for the indeterminates. The details of the definition of $\mathbf{MSOLEVAL}_{\mathscr{S}}$ are given at the end of this section. We first explain the idea of $\mathbf{MSOLEVAL}_{\mathscr{S}}$ by examples for the case where structures represent words over a fixed alphabet Σ.

2.1 Guiding examples

Let $f : \Sigma^\star \to \mathscr{S}$ be an $\mathscr{S}$-valued function on words over the alphabet Σ and let w be a word in $\Sigma^\star$. We call such functions *word functions*, following [3, 4]. They are also called *formal power series* in [1], where the indeterminates are indexed by words and the coefficient of X_w is $f(w)$.

We denote by $w[i]$ the letter at position i in w, and by $w[U]$ the word induced by U, for U a set of positions in w. We denote the length of a word w by $\ell(w)$ and the concatenation of two words $u, v \in \Sigma^\star$ by $u \circ v$. We denote by $[n]$ the set $\{1, 2, \ldots, n\}$.

We will freely pass between words and structures representing words. For the sequel, let $\Sigma = \{0, 1\}$ and $w \in \{0, 1\}^\star$ be represented by the structure

$$\mathscr{A}_w = \langle \{0\} \cup [\ell(w)], <^w, P_0^w, P_1^w \rangle.$$

$P_0^w, P_1^w \subseteq [\ell(w)]$ and $P_0^w \cap P_1^w = \emptyset$ and $P_0^w \cup P_1^w = [\ell]$.

As structures are always non-empty, the universe of a word w is represented by a structure containing the zero position $[n] \cup \{0\} = \{0, 1, \ldots, n\}$. So strictly speaking the size of the structure of the empty word is one, and of a word of length n it is $n + 1$. The zero position, represented by 0, has no letter attached to it, and the elements of the structure different from 0 represent positions in the word which carry letters. The positions in P_0^w carry the letter 0 and the positions in P_1^w carry the letter 1.

Examples 1. *In the following examples the functions are word functions with values in the ring $\mathbb{Z}$ or the polynomial ring $\mathbb{Z}[X]$.*

(i) The function $\sharp_1(w)$ counts the number of occurrences of 1 in a word w and can be written as

$$\sharp_1(w) = \sum_{i \in [n] : P_1(i)} 1.$$

(ii) The polynomial $X^{\sharp_1(w)}$ can be written as

$$X^{\sharp_1(w)} = \prod_{i \in [n] : P_1(i)} X.$$

*(iii) Let L be a regular language defined by the **MSOL**-formula ϕ_L. The generating function of the number of (contiguous) occurrences of words $u \in L$ in a word w, can be written as*

$$\sharp_L(w) = \sum_{U \subseteq [n] : w[U] \models \psi_L} \prod_{i \in U} X,$$

where $\psi_L(U)$ says that U is an interval and ϕ_L^U, the relativization of ϕ_L to U, holds.

(iv) The functions $\mathrm{sq}(w) = 2^{\ell(w)^2}$ and $\mathrm{dexp}(w) = 2^{2^{\ell(w)}}$ are not representable in $\mathbf{MSOLEVAL}_{\mathscr{F}}$.

The *tropical semiring* $\mathcal{T}_{min}$ is the semiring with universe $\mathbb{R} \cup \{\infty\}$, consisting of the real numbers augmented by an additional element ∞, and *min* as addition with ∞ as neutral element and real addition $+$ as multiplication with 0 as neutral element. The tropical semiring $\mathcal{T}_{max}$, also sometimes called *arctic semiring*, is defined analogously, where ∞ is replaced by $-\infty$ and *min* by *max*. The choice of the commutative semiring $\mathcal{S}$ makes quite a difference as illustrated by the following:

Examples 2. *In the next examples the word functions take values in the ring $\mathbb{Z}$ with addition and multiplication, or in the subsemiring of $\mathcal{T}_{max}$ generated by $\mathbb{Z}$. A* block of 1's in a word $w \in \{0,1\}^\star$ *is a maximal set of consecutive positions $i \in [\ell(w)]$ in the word w with $P_1(i)$.*

(i) The function $b_1(w)$ counts the number of blocks of 1's in w. $b_1(w)$ can be written as

$$b_1(w) = \sum_{B \subseteq [\ell(w)]: B \text{ is a block of 1's}} 1$$

which is in $\mathbf{MSOLEVAL}_{\mathbb{Z}}$*. Alternatively, it can be written as*

$$b_1(w) = \sum_{v \in [\ell(w)]: First-in-Block(v)} 1, \tag{1}$$

where $First - in - Block(v)$ is the formula in **MSOL** *which says that v is a first position in a block of 1's. Equation (1) can be expressed in* $\mathbf{MSOLEVAL}_{\mathbb{Z}}$ *and also in both* $\mathbf{MSOLEVAL}_{\mathcal{T}_{min}}$ *and* $\mathbf{MSOLEVAL}_{\mathcal{T}_{max}}$*.*

(ii) Let $mb_1^{max}(w)$ be the function which assigns to the word w the maximum of the sizes of blocks of 1's, and $mb_1^{min}(w)$ be the function which assigns to the word w the minimum of the sizes of blocks of 1's. One can show, see Remark 3, that mb_1^{max} and mb_1^{min} are not definable over the ring $\mathbb{Z}$. However, they are definable over $\mathcal{T}_{max}$, respectively over $\mathcal{T}_{min}$, by writing

$$mb_1^{max} = \max_{B: B \text{ is a block of 1's}} \sum_{v: v \in B} 1$$

and

$$mb_1^{min} = \min_{B: B \text{ is a block of 1's}} \sum_{v: v \in B} 1$$

(iii) The function $b_1(w)^2$ is definable in $\mathbf{MSOLEVAL}_{\mathbb{Z}}$ *because* $\mathbf{MSOLEVAL}_{\mathbb{Z}}$ *is closed under the usual product, cf. Proposition 7. However, it is not definable over either of the two tropical semirings. To see this one notes that polynomials in a tropical semiring are piecewise linear.*

Remark 3. *Let f be a word function which takes values in a field $\mathcal{F}$. The Hankel matrix $\mathcal{H}(f)$ is the infinite matrix where rows and columns are labeled by words u, v and the entry $\mathcal{H}(f)_{u,v} = f(u \circ v)$. It is shown in [14] that for word functions f in* $\mathbf{MSOLEVAL}_{\mathcal{F}}$ *the Hankel matrix $\mathcal{H}(f)$ has finite rank. To show non-definability of f it suffices to show that $\mathcal{H}(f)$ has infinite rank over a field $\mathcal{F}$ extending $\mathbb{Z}$.*

2.2 Formal definition of MSOLEVAL

Let $\mathcal{S}$ be a commutative semiring, which contains the semiring of natural numbers $\mathbb{N}$. We first define **MSOL**-polynomials, which are multivariate polynomials. The functions in **MSOLEVAL** are obtained from **MSOL**-polynomials by substituting values from $\mathcal{S}$ for the indeterminates.

MSOL-polynomials have a fixed finite set of variables (indeterminates, if we distinguish them from the variables of **SOL**), **X**. We denote by $card_{M,v}(\varphi(v))$ the number of elements v in the universe that satisfy φ. We assume τ contains a relation symbol $\mathbf{R}_{\leq}$ which is always interpreted as a linear ordering of the universe.

Let $\mathfrak{M}$ be a τ-structure. We first define the $\mathbf{MSOL}(\tau)$-*monomials* inductively.

Definition 4 (**MSOL**-monomials)**.**

(i) Let $\phi(v)$ be a formula in $\mathbf{MSOL}(\tau)$, where v is a first order variable. Let $r \in \mathbf{X} \cup (\mathscr{S} - \{0\})$ be either an indeterminate or an integer. Then

$$r^{card_{M,v}(\phi(v))}$$

is a standard $\mathbf{MSOL}(\tau)$-monomial (whose value depends on $card_{M,v}(\phi(v))$.

(ii) Finite products of $\mathbf{MSOL}(\tau)$-monomials are $\mathbf{MSOL}(\tau)$-monomials.

Even if r is an integer, and $r^{card_{M,v}(\phi(v))}$ does not depend on $\mathfrak{M}$, the monomial stands as it is, and is not evaluated.

Note the degree of a monomial is polynomially bounded by the cardinality of $\mathfrak{M}$.

Definition 5 (**MSOL**-polynomials)**.** *The polynomials definable in $\mathbf{MSOL}(\tau)$ are defined inductively:*

(i) $\mathbf{MSOL}(\tau)$-monomials are $\mathbf{MSOL}(\tau)$-polynomials.

(ii) Let ϕ be a $\tau \cup \{\bar{\mathbf{R}}\}$-formula in **MSOL** *where $\bar{\mathbf{R}} = (\mathbf{R}_1, \ldots, \mathbf{R}_m)$ is a finite sequence of* unary *relation symbols not in τ. Let t be a $\mathbf{MSOL}(\tau \cup \{\bar{\mathbf{R}}\})$-polynomial. Then*

$$\sum_{\bar{R}:\langle \mathfrak{M}, \bar{R}\rangle \models \phi(\bar{R})} t$$

is a $\mathbf{MSOL}(\tau)$-polynomial.

For simplicity we refer to $\mathbf{MSOL}(\tau)$-polynomials as **MSOL**-polynomials when τ is clear from the context.

We shall use the following properties of **MSOL**-polynomials. The proofs can be found in [21].

Lemma 6.

(i) Every indeterminate $x \in \mathbf{X}$ can be written as an **MSOL**-*monomial.*

(ii) Every integer c can be written as an **MSOL**-*monomial.*

Proposition 7. *The pointwise product of two* **MSOL**-*polynomials is again an* **MSOL**-*polynomial.*

3 MSOLEVAL$_{\mathscr{S}}$ and Weighted Automata

Let $\mathscr{S}$ be a commutative semiring and Σ a finite alphabet. A weighted automaton A of size r over $\mathscr{S}$ is given by:

(i) Two vectors $\alpha, \gamma \in \mathscr{S}^r$, and

(ii) for each $\sigma \in \Sigma$ a matrix $\mu_\sigma \in \mathscr{S}^{r \times r}$.

For a matrix or vector M we denote by M^T the transpose of M.

For a word $w = \sigma_1 \sigma_2 \ldots \sigma_{\ell(w)}$ the automaton A defines the function

$$f_A(w) = \alpha \cdot \mu_{\sigma_1} \cdot \ldots \cdot \mu_{\sigma_{\ell(w)}} \cdot \gamma^T.$$

A word function $f : \Sigma^\star \to \mathscr{S}$ is recognized by an automaton A if $f = f_A$. f is recognizable if there exists a weighted automaton A which recognizes it.

Theorem 8. *Let f be a word function with values in a commutative semiring $\mathscr{S}$. Then $f \in$ **MSOLEVAL**$_{\mathscr{S}}$ iff f is recognized by some weigthed automaton A over $\mathscr{S}$.*

In this section we prove Theorem 8 using model theoretic tools, without going through weighted logic. We need a few definitions.

The *quantifier rank* $qr(f)$ of a word function f in **MSOLEVAL**$_{\mathscr{S}}$ is defined as the maximal quantifier rank of the formulas which appear in the definition of f. It somehow measures the complexity of f, but we do not need the technical details in this paper. Quantifier ranks of formulas in **MSOL** are defines as usual, cf. [12].

We denote by $\mathscr{S}^{\Sigma^\star}$ the set of word functions $\Sigma^\star \to \mathscr{S}$. A *semimodule* $\mathscr{M}$ is a subset of $\mathscr{S}^{\Sigma^\star}$ closed under point-wise addition of word functions in $\mathscr{M}$, and point-wise multiplication with elements of $\mathscr{S}$. Note that $\mathscr{S}^{\Sigma^\star}$ itself is a semimodule.

$M \subseteq \mathscr{S}^{\Sigma^\star}$ *is finitely generated* if there is a finite set $F \subseteq \mathscr{S}^{\Sigma^\star}$ such that each $f \in M$ can be written as a (semiring) linear combination of elements in F. Let w be a word and f a word function. Then we denote by $w^{-1}f$ the word function g defined by

$$g(u) = (w^{-1}f)(u) = f(w \circ u)$$

M is *stable* if for all words $w \in \Sigma^\star$ and for all $f \in \mathscr{M}$ the word function $w^{-1}f$ is also in M.

3.1 Word functions in MSOLEVAL$_{\mathscr{S}}$ are recognizable

To prove the "only if" direction of Theorem 8 we use the following two theorems.

For a commutative semiring $\mathscr{S}$ and a sequence of indeterminates $\bar{X} = (X_1, \ldots, X_t)$ we denote by $\mathscr{S}[\bar{X}]$ the commutative semiring of polynomials with indeterminates $\bar{X}$ and coefficients in $\mathscr{S}$. The first theorem is from [22].

Theorem 9 (Bilinear Decomposition Theorem for Word Functions)**.**
*Let $\mathscr{S}$ be a commutative semiring. Let $f \in$ **MSOLEVAL**$_{\mathscr{S}}$ be a word function $\Sigma^+ \to \mathscr{S}$ of quantifier rank $qr(f)$. There are:*

(i) a function $\beta : \mathbb{N} \to \mathbb{N}$,

(ii) a finite vector $F = (g_1, \ldots, g_{\beta(qr(f))})$ of functions in **MSOLEVAL**$_{\mathscr{S}}$ *of length $\beta(qr(f))$, with $f = g_i$ for some $i \le \beta(qr(f))$,*

(iii) and for each $g_i \in F$, a matrix $M^{(i)} \in \mathscr{S}^{\beta(qr(f)) \times \beta(qr(f))}$

such that

$$g_i(u \circ v) = F(u) \cdot M^{(i)} F(v)^T.$$

The other theorem was first proved by G. Jacob, [17, 1].

Theorem 10 (G. Jacob 1975)**.** *Let f be a word function $f : \Sigma^\star \to \mathscr{S}$. Then f is recognizable by a weighted automaton over $\mathscr{S}$ iff there exists a finitely generated stable semimodule $\mathscr{M} \subseteq \mathscr{S}^{\Sigma^\star}$ which contains f.*

In order to prove the "only if" direction of Theorem 8 we reformulate it.

Theorem 11 (Stable Semimodule Theorem)**.** *Let $\mathscr{S}$ be a commutative semiring and let $f \in$ **MSOLEVAL**$_{\mathscr{S}}$ be a word function of quantifier rank $qr(f)$.*
There are:

(i) a function $\beta : \mathbb{N} \to \mathbb{N}$,

(ii) a finite vector $F = (g_1, \dots, g_{\beta(qr(f))})$ of functions in **MSOLEVAL**$_{\mathscr{S}}$ *of length $\beta(qr(f))$, with $f = g_i$ for some $i \leq \beta(qr(f))$,*

such that the semimodule $\mathscr{M}[F]$ generated by F is stable.

Proof. We take F and the matrices $M^{(i)}$ from Theorem 9 stated in the introduction.

We have to show that for every fixed word w and $f \in \mathscr{M}[F]$ the function $w^{-1}f \in \mathscr{M}[F]$. As $f \in \mathscr{M}[F]$ there is a vector $A = (a_1, \dots, a_{\beta(qr(f))}) \in \mathscr{S}^{\beta(qr(f))}$ such that

$$f(w) = A \cdot F^T(w)$$

for every fixed word w. Here $F(w)$ is shorthand for $(g_1(w), \dots, g_{\beta(qr(f))}(w))$.

Let u be a word. We compute $(w^{-1}f)(u)$.

$$(w^{-1}f)(u) = f(w \circ u) = A \cdot F^T(w \circ u) =$$
$$\sum_{i=1}^{\beta(qr(f))} a_i g_i(w \circ u) = \sum_{i=1}^{\beta(qr(f))} a_i F(w) M^{(i)} F^T(u)$$

We put $B_i = a_i F(w) M^{(i)}$ and observe that $B_i \in \mathscr{S}^{\beta(qr(f))}$. If we take $B = \sum_i^{\beta(qr(f))} B_i$ we get that $(w^{-1}f)(u) = B \cdot F^T(u)$, hence $w^{-1}f \in \mathscr{M}[F]$. □

3.2 Recognizable word functions are definable in MSOLEVAL$_{\mathscr{S}}$

For the "if" direction we proceed as follows:

Proof. Let A be a weighted automaton of size r over $\mathscr{S}$ for words in $\Sigma^\star$. For a word w with $\ell(w) = n$, given as a function $w : [n] \to \Sigma$, the automaton A defines the function

$$f_A(w) = \alpha \cdot \mu_{w(1)} \cdot \ldots \cdot \mu_{w(n)} \cdot \gamma^T. \tag{2}$$

We have to show that $f_A \in$ **MSOLEVAL**$_{\mathscr{S}}$.

To unify notation we define

$$M^a_{i,j} = (\mu_a)_{i,j}.$$

Equation (2) is a product of n matrices and two vectors.

Let P be the product of these matrices,

$$P = \prod_{k=1}^{n} \mu_{w(k)}.$$

Using matrix algebra we get for the entry $P_{a,b}$ of P:

$$P_{a,b} = \sum_{i_{n-1}=1}^{r} \left(\sum_{i_{n-2}=1}^{r} \left(\ldots \left(\sum_{i_1=1}^{r} M^{w(1)}_{a,i_1} \cdot M^{w(2)}_{i_1,i_2} \right) M^{w(3)}_{i_2,i_3} \right) \ldots \right) M^{w(n)}_{i_{n-1},b}$$
$$= \sum_{i_1,\ldots i_{n-1} \leq r} \left(M^{w(1)}_{a,i_1} \cdot M^{w(2)}_{i_1,i_2} \cdot \ldots \cdot M^{w(n)}_{i_{n-1},b} \right)$$

Let $\pi : [n-1] \to [r]$ be the function with $\pi(k) = i_k$. We rewrite $P_{a,b}$ as:

$$P_{a,b} = \sum_{\pi:[n-1]\to[r]} \left(M^{w(1)}_{a,\pi(1)} \cdot M^{w(2)}_{\pi(1),\pi(2)} \cdot \ldots M^{w(n)}_{\pi(n-1),b} \right) \tag{3}$$

Next we compute the b coordinate of the vector $\alpha \cdot P$:

$$(\alpha \cdot P)_b = \sum_{i=1}^{r} \alpha_i \cdot P_{i,b}$$

Therefore

$$f_A(w) = \alpha \cdot P \cdot \gamma = \sum_{b=1}^{r} (\alpha \cdot P)_b \cdot \gamma_b$$
$$= \sum_{b=1}^{r} \left(\sum_{a=1}^{r} \alpha_a \cdot P_{a,b} \right) \cdot \gamma_b = \sum_{a,b\le r} \alpha_a \cdot P_{a,b} \cdot \gamma_b$$

and by using Equation (3) for $P_{a,b}$ we get:

$$\sum_{a,b\le r} \alpha_a \cdot \left(\sum_{\pi:[n-1]\to[r]} \left(M^{w(1)}_{a,\pi(1)} \cdot M^{w(2)}_{\pi(1),\pi(2)} \cdot \ldots M^{w(n)}_{\pi(n-1),b} \right) \right) \cdot \gamma_b$$

Now let $\pi' : [n] \cup \{0\} \to [r]$ be the function for which $\pi'(0) = a, \pi'(n) = b$ and $\pi'(k) = \pi(k) = i_k$ for $1 \le k \le n-1$. Then we get

$$f_A(w) =$$
$$\sum_{\pi':[n]\cup\{0\}\to[r]} \alpha_{\pi'(0)} \cdot \left[M^{w(1)}_{\pi'(0),\pi'(1)} \cdot \ldots \cdot M^{w(n)}_{\pi'(n-1),\pi'(n)} \right] \cdot \gamma_{\pi'(n)} =$$
$$\sum_{\pi':[n]\cup\{0\}\to[r]} \alpha_{\pi'(0)} \cdot \left(\prod_{k\in[n]} M^{w(k)}_{\pi'(k-1),\pi'(k)} \right) \cdot \gamma_{\pi'(n)} \tag{4}$$

To convert Equation (4) into an expression in $\mathbf{MSOLEVAL}_{\mathscr{S}}$ we use a few lemmas:

First, let S be any set and $\pi : S \to [r]$ be any function. π induces a partition of S into sets $U^{\pi}_1, \ldots, U^{\pi}_r$ by $U^{\pi}_i = \{s \in S : \pi(s) = i\}$. Conversely, every partition $\mathscr{U} = (U_1, \ldots, U_r)$ of S induces a function $\pi_{\mathscr{U}}$ by setting $\pi_{\mathscr{U}}(s) = i$ for $s \in U_i$. To pass between functions π with finite range $[r]$ and partitions into r-sets we use the following lemma:

Lemma 12. *Let $E(\pi)$ be any expression depending on π.*

$$\sum_{\pi:S\to[r]} E(\pi) = \sum_{\mathscr{U}} E(\pi_{\mathscr{U}}) = \sum_{U_1,\ldots U_r:Partition(U_1,\ldots,U_r)} E(\pi_{\mathscr{U}})$$

where $\mathscr{U}$ ranges over all partitions of S into r sets $U_i : i \in [r]$. Clearly, $Partition(U_1, \ldots, U_r)$ can be written in **MSOL***.*

Second, to convert the factors $\alpha_{\pi'(0)}$ and $\gamma_{\pi'(n)}$ we proceed as follows:

Lemma 13. *Let α_i be the unique value of the coordinate of α such that $0 \in U_i$. Similarly, let γ_i be the unique value of the coordinate of γ such that $n \in U_i$.*

$$\alpha_{\pi'(0)} = \prod_{i=1}^{r} \prod_{0 \in U_i} \alpha_i$$
$$\gamma_{\pi'(n)} = \prod_{i=1}^{r} \prod_{n \in U_i} \gamma_i$$

Proof. First we note that, as $\mathscr{U}$ is the partition induced by π', the restriction of π' to U_i is constant for all $i \in [r]$. Next we note that the product ranging over the empty set gives the value 1. □

Similarly, to convert the factor $\prod_{k\in[n]} M^{w(k)}_{\pi'(k-1),\pi'(k)}$ use the following lemma:

Lemma 14. *Let $m_{i,j,w(v)}$ be the unique value of the (i, j)-entry of the matrix $\mu_{w(v)}$ such that $v \in U_i$ and $v+1 \in U_j$.*

$$\prod_{k\in[n]} M^{w(k)}_{\pi'(k-1),\pi'(k)} = \prod_{i,j=1}^{r} \left(\prod_{v-1\in U_i, v\in U_j} m_{i,j,w(v)} \right)$$

Using the fact that every element which is the interpretation of a term in $\mathscr{S}$ can be written as an expression in $\mathbf{MSOLEVAL}_{\mathscr{S}}$, Lemma 6 in Section 2.2, we can write $U_i(v)$ instead of $v \in U_i$, and see that the monomials of Lemmas 12, 13 and 14 are indeed in $\mathbf{MSOLEVAL}_{\mathscr{S}}$. Now we apply the fact that the pointwise product of two word functions in $\mathbf{MSOLEVAL}_{\mathscr{S}}$ is again a function in $\mathbf{MSOLEVAL}_{\mathscr{S}}$, Proposition 7 in Section 2.2,

to Lemmas 12, 13 and 14 and complete the proof of Theorem 8. □

4 Weighted MSOL and MSOLEVAL

In this section we compare the formalism of weighted **MSOL**, **WMSOL**, with our $\mathbf{MSOLEVAL}_{\mathscr{S}}$ for arbitrary commutative semirings. In [7, 8] and [2] two fragments of weighted **MSOL** are discussed. One is based on *unambiguous* formulas (a semantic concept), the other on *step formulas* based on the Boolean fragment of weighted **MSOL** (a syntactic definition). The two fragments have equal expressive power, as stated in [2], and characterize the functions recognizable by weighted automata. We denote both versions by **RMSOL**.

4.1 Syntax of WMSOL, the weighted version of MSOL

The definitions and properties of **WMSOL** and its fragments are taken literally from [2]. The syntax of formulas ϕ of weighted **MSOL**, denoted by **WMSOL**, is given inductively in Backus–Naur form by

$$\phi ::= k \mid P_a(x) \mid \neg P_a(x) \mid x \leq y \mid \neg x \leq y \mid x \in X \mid x \notin X$$
$$\mid \phi \vee \psi \mid \phi \wedge \psi \mid \exists x.\phi \mid \exists X.\phi \mid \forall x.\phi \mid \forall X.\phi$$

where $k \in \mathscr{S}$, $a \in \Sigma$. The set of weighted **MSOL**-formulas over the field $\mathscr{S}$ and the alphabet Σ is denoted by $\mathbf{MSOL}(\mathscr{S}, \Sigma)$. **bMSOL** *formulas* and **bMSOL**-*step formulas* are defined below. **bMSOL** is the Boolean fragment of **WMSOL**, and its name is justified by Lemma 15. **RMSOL** is the fragment of

WMSOL where universal second order quantification is restricted to **bMSOL** and first order universal quantification is restricted to **bMSOL**-step formulas.

The syntax of weighted **bMSOL** is given by

$$\phi ::= 0 \mid 1 \mid P_a(x) \mid x \leq y \mid x \in X \mid \neg\phi \mid \phi \wedge \psi \mid \forall x.\phi \mid \forall X.\phi$$

where $a \in \Sigma$.

The set of weighted **MSOL**-formulas over the commutative semiring $\mathscr{S}$ and the alphabet Σ is denoted by **WMSOL**$(\mathscr{S}, \Sigma)$.

Instead of defining step-formulas as in [2] we use Lemma 3 from [2] as our definition.

A **bMSOL**-step formula ψ is a formula of the form

$$\psi = \bigvee_{i \in I} (\phi_i \wedge k_i) \tag{5}$$

where I is a finite set, $\phi_i \in$ **bMSOL** and $k_i \in \mathscr{S}$.

4.2 Semantics of WMSOL, and translation of RMSOL into MSOLEVAL$_{\mathscr{S}}$

Next we define the semantics of **WMSOL** and, where it is straightforward, simultaneously also its translations into **MSOLEVAL**$_{\mathscr{S}}$.

The evaluations of weighted formulas $\phi \in$ **WMSOL**$(\mathscr{S}, \Sigma)$ on a word w are denoted by $WE(\phi, w, \sigma)$, where σ is an assignment of the variables of ϕ to positions, respectively sets of positions, in w.

We denote the evaluation of term t of **MSOLEVAL**$_{\mathscr{S}}$ for a word w and an assignment for the free variables σ by $E(t, w, \sigma)$. $\mathrm{tv}(\phi)$ stands for the truth value of ϕ (subject to an assignment for the free variables), i.e., $E(\mathrm{tv}(\phi), w, \sigma) = 0 \in \mathscr{S}$ for false and $E(\mathrm{tv}(\phi), w, \sigma) = 1 \in \mathscr{S}$ for true. The term $\mathrm{tv}(\phi)$ is used as an abbreviation for

$$\mathrm{tv}(\phi) = \sum_{U : U = A \wedge \phi} 1$$

where $U = A$ stands for $\forall x (U(x) \leftrightarrow x = x)$ and U does not occur freely in ϕ. Indeed, we have

$$E(\mathrm{tv}(\phi), w, \sigma) = \begin{cases} 1 & (w, \sigma) \models \phi \\ 0 & \text{else} \end{cases}$$

We denote by $TRUE(x)$ the formula $x = x$ with free first order variable x. Similarly, $TRUE(X)$ denotes the formula $\exists y \in X \vee \neg \exists y \in X$ with free set variable X.

The evaluations of formulas $\phi \in$ **WMSOL** and their translations are now defined inductively.

(i) For $k \in \mathscr{S}$ we have $tr(k) = k$ and $WE(k, w, \sigma)) = E(tr(k), w, \sigma)) = k$.

(ii) For atomic formulas θ we have $tr(\theta) = \mathrm{tv}(\theta)$ and

$$WE(\theta, w, \sigma) = E(tr(\theta), w, \sigma) = E(\mathrm{tv}(\theta), w, \sigma)$$

(iii) For negated atomic formulas we have

$$tr(\neg\theta) = 1 - tr(\theta) = 1 - \mathrm{tv}(\theta)$$

and

$$WE(\neg\theta, w, \sigma) = 1 - E(\mathrm{tv}(\theta), w, \sigma).$$

(iv) $tr(\phi_1 \vee \phi_2) = tr(\phi_1) + tr(\phi_2)$ and

$$WE(\phi_1 \vee \phi_2, w, \sigma) = E(tr(\phi_1) + tr(\phi_2), w, \sigma) = E(tr(\phi_1), w, \sigma) + E(tr(\phi_2), w, \sigma).$$

(v) $tr(\exists x.\phi) = \sum_{x:TRUE(x)} tr(\phi)$ and

$$WE(\exists x.\phi, w, \sigma) = E(\sum_{x:TRUE(x)} tr(\phi, w, \sigma)) = \sum_{x:TRUE(x)} E(tr(\phi, w, \sigma)).$$

(vi) $tr(\exists X.\phi) = \sum_{X:TRUE(X)} tr(\phi)$ and

$$WE(\exists X.\phi, w, \sigma) = E(\sum_{X:TRUE(X)} tr(\phi, w, \sigma)) = \sum_{X:TRUE(X)} E(tr(\phi, w, \sigma)).$$

(vii) $tr(\phi_1 \wedge \phi_2) = tr(\phi_1) \cdot tr(\phi_2)$ and

$$WE(\phi_1 \wedge \phi_2, w, \sigma) = E(tr(\phi_1) \cdot tr(\phi_2), w, \sigma) = E(tr(\phi_1), w, \sigma) \cdot E(tr(\phi_2), w\sigma).$$

So far the definition of WE was given using the evaluation function E and the translation was straightforward. Problems arise with the universal quantifiers.

The unrestricted definition of WE for **WMSOL** given below gives us functions which are not recognizable by weighted automata, and the straightforward translation defined below gives us expressions which are not in $\mathbf{MSOLEVAL}_{\mathscr{S}}$:

(viii) $tr(\forall x.\phi) = \prod_{x:TRUE(x)} tr(\phi)$ and

$$WE(\forall x.\phi, w, \sigma) = E(\prod_{x:TRUE(x)} tr(\phi, w, \sigma)) = \prod_{x:TRUE(x)} E(tr(\phi, w, \sigma)).$$

The formula $\phi_{sq} = \forall x.\forall y.2$ gives the function $2^{\ell(w)^2}$ and is not a **bMSOL**-step formula. The straightforward translation tr gives the term

$$\prod_{x:TRUE(x)} \left(\prod_{y:TRUE(y)} 2 \right) = \prod_{(x,y):TRUE(x,y)} 2,$$

which is a product over the tuples of a binary relation, hence not in $\mathbf{MSOLEVAL}_{\mathscr{S}}$.

(ix) $tr(\forall X.\phi) = \prod_{X:TRUE(X)} tr(\phi)$ and

$$WE(\forall X.\phi, w, \sigma) = E(\prod_{X:TRUE(X)} tr(\phi, w, \sigma)) = \prod_{X:TRUE(X)} E(tr(\phi, w, \sigma)).$$

Here the translation gives a product $\prod_{X:TRUE(X)}$ ranging over subsets, which is not an expression in $\mathbf{MSOLEVAL}_{\mathscr{S}}$.

In **RMSOL**, universal second order quantification is restricted to formulas of **bMSOL**, and first order universal quantification is restricted to **bMSOL**-step formulas.

In [2, page 590], after Figure 1, the following is stated:

Lemma 15. *The evaluation WE of a* **bMSOL***-formula ϕ assumes values in $\{0,1\}$ and coincides with the standard semantics of ϕ as an unweighted* **MSOL***-formula.*

Because the translation of universal quantifiers using tr leads outside of $\mathbf{MSOLEVAL}_{\mathscr{S}}$, we define a proper translation $tr' : \mathbf{RMSOL} \to \mathbf{MSOLEVAL}_{\mathscr{S}}$.

Using Lemma 15 we set $tr'(\phi) = \mathrm{tv}(\phi)$, for ϕ a **bMSOL**-formula.

For universal first order quantification of **bMSOL**-step formulas

$$\psi = \bigvee_{i\in I} (\phi_i \wedge k_i) \tag{6}$$

we compute $WE(\forall x.\psi, w, \sigma)$ and $E(tr(\forall x.\psi), w, \sigma)$ as follows, leaving the steps for the translation of $tr(\forall x.\psi)$ to the reader.

$$\begin{aligned} WE((\forall x.\psi), w, \sigma) = E(tr(\forall x.\psi), w, \sigma) = \\ E(tr(\forall x. \bigvee_{i\in I} (\phi_i \wedge k_i)), w, \sigma) = \\ \prod_{x:TRUE(x)} E(tr(\bigvee_{i\in I} (\phi_i \wedge k_i))), w, \sigma) = \\ \prod_{x:TRUE(x)} (\sum_{i\in I} (E(tr'(\phi_i)) \cdot k_i), w, \sigma)) = \\ \prod_{x:TRUE(x)} (\sum_{i\in I} (E(\mathrm{tv}(\phi_i), w, \sigma) \cdot k_i))) \end{aligned}$$

Clearly, the formula of the last line, $\prod_{x:TRUE(x)}(\sum_{i\in I}(\mathrm{tv}(\phi_i)) \cdot k_i))$ is an expression in $\mathbf{MSOLEVAL}_{\mathscr{S}}$.

For universal second order quantification of **bMSOL**-formulas ψ we use Lemma 15 and get

$$WE(\forall X.\psi, w, \sigma) = E(tr'(\forall X \psi), w, \sigma) = E(\mathrm{tv}(\forall X \psi), w, \sigma)$$

Clearly, the expression $\mathrm{tv}(\forall X \psi)$ is an expression in $\mathbf{MSOLEVAL}_{\mathscr{S}}$. Thus we have proved:

Theorem 16. *Let $\mathscr{S}$ be a commutative semiring. For every expression $\phi \in$ **RMSOL** there is an expression $tr'(\phi) \in \mathbf{MSOLEVAL}_{\mathscr{S}}$ such that $WE(\phi, w, \sigma) = E(tr'(\phi), w, \sigma)$, i.e., ϕ and $tr'(\phi)$ define the same word function.*

4.3 Translation from $\mathbf{MSOLEVAL}_{\mathscr{S}}$ to RMSOL

It follows from our Theorem 8 and the characterization in [8] of recognizable word functions as the functions definable in **RMSOL**, that the converse is also true. We now give a direct proof of the converse without using weighted automata.

Theorem 17. *Let $\mathscr{S}$ be a commutative semiring. For every expression $t \in \mathbf{MSOLEVAL}_{\mathscr{S}}$ there is a formula $\phi_t \in$ **RMSOL** such that $WE(\phi_t, w, \sigma) = E(t, w, \sigma)$, i.e., ϕ_t and t define the same word function.*

Proof. (i) Let $t = \prod_{x:\phi(x)} \alpha$ be a $\mathbf{MSOLEVAL}_{\mathscr{S}}$- monomial. We note that

$$\alpha \cdot \mathrm{tv}(\phi) + \mathrm{tv}(\neg\phi) = \begin{cases} \alpha & \text{if } \phi \text{ is true} \\ 1 & \text{else} \end{cases}$$

Furthermore, by Lemma 15 $\phi \in$ **bMSOL**. So we put

$$\phi_t = \forall x.((\phi(x) \wedge \alpha) \vee \neg\phi(x))$$

(ii) Let $t_1 = \sum_{U:\phi(U)} t$ and let ϕ_t be the translation of t. Then

$$\phi_{t_1} = \exists U.(\phi_t \wedge \phi(U))$$

□

5 Conclusions

We have given two proofs that **RMSOL** and **MSOLEVAL** with values in $\mathscr{S}$ have the same expressive power over words. One proof uses model theoretic tools to show directly that **MSOLEVAL** captures the functions recognizable by weighted automata. The other proof shows how to translate the formalisms from one into the other. Adapting the translation proof, it should be possible to extend the result to tree functions as well, cf. [10].

Although in this paper we dealt only with word functions, our formalism **MSOLEVAL**, introduced first fifteen years ago, was originally designed to deal with definability of graph parameters and graph polynomials, [6, 22, 24, 21]. It has been useful, since, in many applications in algorithmic and structural graph theory and descriptive complexity. Its use in characterizing word functions recognizable by weighted automata is new. **MSOLEVAL** can be seen as an analogue of the *Skolem elementary functions* aka *lower elementary functions*, [25, 26], adapted to the framework of *meta-finite model theory* as defined in [15].

The formalism **WMSOL** of weighted logic was first invented in 2005 in [7] and since then used to characterize word and tree functions recognizable by weighted automata, [10]. These characterizations need some syntactic restrictions which lead to the formalisms of **RMSOL**. No such syntactic restrictions are need for the characterization of recognizable word functions using **MSOLEVAL**. The weighted logic **WMSOL** can also be defined for general relational structures. However, it is not immediate which syntactic restrictions are needed, if at all, to obtain algorithmic applications similar to the ones obtained using **MSOLEVAL**, cf. [6, 5, 23].

Acknowledgements We are thankful to Jacques Sakarovitch, Géraud Sénizergues and Amir Shpilka for useful guidence on the subject of weighted automata. We are thankful to Manfred Droste, Tomer Kotek and Elena Ravve and several anonymous readers, for their valuable comments.

References

[1] J. Berstel & C. Reutenauer (1984): *Rational Series and their languages*. *EATCS Monographs on Theoretical Computer Science* 12, Springer.

[2] B. Bollig, P. Gastin, B. Monmege & M. Zeitoun (2010): *Pebble weighted automata and transitive closure logics*. In: *ICALP'10*, *Lecture Notes in Computer Science* 6199, Springer, pp. 587–598, doi:10.1007/978-3-642-11301-7.

[3] J.W. Carlyle & A. Paz (1971): *Realizations by Stochastic Finite Automata*. *J. Comp. Syst. Sc.* 5, pp. 26–40, doi:10.1016/S0022-0000(71)80005-3.

[4] A. Cobham (1978): *Representation of a Word Function as the Sum of Two Functions*. *Mathematical Systems Theory* 11, pp. 373–377, doi:10.1007/BF01768487.

[5] B. Courcelle, J.A. Makowsky & U. Rotics (2000): *Linear Time Solvable Optimization Problems on Graphs of Bounded Clique-Width*. *Theory of Computing Systems* 33.2, pp. 125–150, doi:10.1007/s002249910009.

[6] B. Courcelle, J.A. Makowsky & U. Rotics (2001): *On the Fixed Parameter Complexity of Graph Enumeration Problems Definable in Monadic Second Order Logic*. *Discrete Applied Mathematics* 108(1-2), pp. 23–52, doi:10.1016/S0166-218X(00)00221-3.

[7] M. Droste & P. Gastin (2005): *Weighted Automata and Weighted Logics*. In: *ICALP 2005*, pp. 513–525, doi:10.1007/11523468_42.

[8] M. Droste & P. Gastin (2007): *Weighted automata and weighted logics*. *Theor. Comput. Sci.* 380(1-2), pp. 69–86, doi:10.1016/j.tcs.2007.02.055.

[9] M. Droste, W. Kuich & H. Vogler, editors (2009): *Handbook of Weighted Automata*. EATCS Monographs on Theoretical Computer Science, Springer.

[10] M. Droste & H. Vogler (2006): *Weighted tree automata and weighted logics*. *Theor. Comput. Sci.* 366, pp. 228–247, doi:10.1016/j.tcs.2006.08.025.

[11] Manfred Droste & Werner Kuich (2013): *Weighted finite automata over semirings*. *Theor. Comput. Sci.* 485, pp. 38–48, doi:10.1016/j.tcs.2013.02.028.

[12] H.-D. Ebbinghaus & J. Flum (1995): *Finite Model Theory*. Perspectives in Mathematical Logic, Springer, doi:10.1007/978-3-662-03182-7.

[13] M. Fliess (1974): *Matrices de Hankel*. *J Maths Pures Appl* 53, pp. 197–222. Erratum in volume 54.

[14] B. Godlin, T. Kotek & J.A. Makowsky (2008): *Evaluation of graph polynomials*. In: *34th International Workshop on Graph-Theoretic Concepts in Computer Science, WG08*, *Lecture Notes in Computer Science* 5344, pp. 183–194, doi:10.1007/978-3-540-92248-3_17.

[15] E. Grädel & Y. Gurevich (1998): *Metafinite Model Theory*. *Information and Computation* 140, pp. 26–81, doi:10.1006/inco.1997.2675.

[16] J. E. Hopcroft & J. D. Ullman (1980): *Introduction to Automata Theory, Languages and Computation*. Addison-Wesley Series in Computer Science, Addison-Wesley.

[17] G. Jacob (1975): *Représentations et substitutions matricielles dans la théorie algébrique des transductions*. Ph.D. thesis, Université de Paris, VII.

[18] T. Kotek (March 2012): *Definability of combinatorial functions*. Ph.D. thesis, Technion - Israel Institute of Technology, Haifa, Israel. Submitted.

[19] T. Kotek & J.A. Makowsky (2012): *Connection Matrices and the Definability of Graph Parameters*. In: *CSL 2012*, pp. 411–425, doi:10.4230/LIPIcs.CSL.2012.411.

[20] T. Kotek, J.A. Makowsky & B. Zilber (2008): *On Counting Generalized Colorings*. In: *Computer Science Logic, CSL'08*, *Lecture Notes in Computer Science* 5213, pp. 339–353, doi:10.1007/978-3-540-87531-4_25.

[21] T. Kotek, J.A. Makowsky & B. Zilber (2011): *On Counting Generalized Colorings*. In M. Grohe & J.A. Makowsky, editors: *Model Theoretic Methods in Finite Combinatorics*, *Contemporary Mathematics* 558, American Mathematical Society, pp. 207–242, doi:10.1090/conm/558/11052.

[22] J.A. Makowsky (2004): *Algorithmic uses of the Feferman-Vaught theorem*. *Annals of Pure and Applied Logic* 126.1-3, pp. 159–213, doi:10.1016/j.apal.2003.11.002.

[23] J.A. Makowsky (2005): *Coloured Tutte polynomials and Kauffman brackets for graphs of bounded tree width*. *Discrete Applied Mathematics* 145(2), pp. 276–290, doi:10.1016/j.dam.2004.01.016.

[24] J.A. Makowsky (2008): *From a Zoo to a Zoology: Towards a general theory of graph polynomials*. *Theory of Computing Systems* 43, pp. 542–562, doi:10.1007/s00224-007-9022-9.

[25] Th. Skolem (1962): *Proof of some theorems on recursively enumerable sets*. *Notre Dame Journal of Formal Logic* 3.2, pp. 65–74, doi:10.1305/ndjfl/1093957149.

[26] S.A. Volkov (2010): *On a class of Skolem elementary functions*. *Journal of Applied and Industrial Mathematics* 4.4, pp. 588–599, doi:10.1134/S1990478910040149.