BRICS

Basic Research in Computer Science

Automata on Series-Parallel Biposets

Zoltán Ésik
Zoltán L. Németh

BRICS Report Series RS-01-24
ISSN 0909-0878 June 2001

Automata on Series-Parallel Biposets*

Z. Ésik and Z. L. Németh
Dept. of Computer Science
University of Szeged
P.O.B. 652
6701 Szeged, Hungary
esik/zlnemeth@inf.u-szeged.hu

Abstract

We provide the basics of a 2-dimensional theory of automata on series-parallel biposets. We define recognizable, regular and rational sets of series-parallel biposets and study their relationship. Moreover, we relate these classes to languages of series-parallel biposets definable in monadic second-order logic.

1 Introduction

Finite automata process words, i.e., elements of a finitely generated free semigroup. In this paper, we define automata whose input structure is a finitely generated free bisemigroup equipped with two associative operations. The elements of the free bisemigroup may be represented by labelled series-parallel biposets. We introduce recognizable, regular and rational sets of series-parallel biposets and study their relationship. Moreover, by relying on the main result of Hoogeboom and ten Pas [16], we relate these classes to languages of series-parallel biposets definable in monadic second-order logic. All of our results can be generalized to higher dimensions, i.e., to any finite number of associative operations.

*Research supported by grant no. T30511 from the National Foundation of Hungary for Scientific Research.

Our study owes much to the work of Hoogeboom and ten Pas [15, 16] on text languages, and to the recent work of Lodaya and Weil [19, 20] and Kuske [17, 18] on languages of series-parallel posets that may be seen as a two-dimensional extension of the classical theory to a situation where one of the two associative operations is commutative. We believe that the case that none of the two operations is commutative is more fundamental. An independent study of automata and languages over free bisemigroups was also initiated by Hashiguchi et. al. [14]. However, the approach taken in *op. cit.* is very syntactic. See the last section for a comparison.

2 Biposets

Let n denote a positive integer and let Σ denote a finite alphabet. A *Σ-labelled n-poset*, or n-poset, for short, is a finite nonempty set P of vertices equipped with n (irreflexive) partial orders $<_i$, $i \in [n] = \{1, \ldots, n\}$, and a labelling function $\lambda : P \to \Sigma$. A *Σ-labelled biposet*, or biposet, is an n-poset for $n = 2$. The two partial orders of a biposet are called the *horizontal* and the *vertical order*. Accordingly, we write $<_h$ and $<_v$. A *morphism* between n-posets P and Q is a function on the vertices that preserves the partial orders and the labelling. An *isomorphism* is a bijective morphism whose inverse is also a morphism. Below we will identify isomorphic n-posets.

Suppose that $P = (P, <_1^P, \ldots, <_n^P, \lambda_P)$ and $Q = (Q, <_1^Q, \ldots, <_n^Q, \lambda_Q)$ are Σ-labelled n-posets. Without loss of generality assume that P and Q are disjoint. For each $i \in [n]$, we define the *$\circ_i$-product* $P \circ_i Q$ to be the n-poset with underlying set $P \cup Q$, partial orders

$$<_j^{P \circ_i Q} \quad = \quad \begin{cases} <_j^P \cup <_j^Q & \text{if } j \neq i \\ <_i^P \cup <_i^Q \cup (P \times Q) & \text{if } j = i \end{cases}$$

and labelling $\lambda_{P \circ_i Q} = \lambda_P \cup \lambda_Q$. When $n = 2$, the product operations $\circ_1$ and $\circ_2$ are called the *series product* or *horizontal product* and the *parallel product* or *vertical product*, respectively. It is clear that the product operations $\circ_i$ are associative.

Each letter $a \in \Sigma$ may be identified with the singleton n-poset labelled a. Let $\mathrm{SP}_n(\Sigma)$ denote the collection of n-posets that can be generated from the singletons by the n product operations.

THEOREM 2.1 [9] *An n-poset $P = (P, <_1, \ldots, <_n, \lambda_P)$ is in $\mathrm{SP}_n(\Sigma)$ iff the following conditions hold.*

1. *For every $u, v \in P$ with $u \neq v$ there is exactly one $i \in [n]$ such that $u <_i v$ or $v <_i u$ holds.*
2. *Each poset $(P, <_i)$, $i \in [n]$ is N-free, i.e., it does not have an induced subposet isomorphic to the poset $([4], <)$ with $1 < 3$, $2 < 3$, $2 < 4$.*
3. *P satisfies the following* triangle condition*: If u, v, w are different vertices of P, then u, v, w are related by at most 2 of the partial orders $<_i$ (i.e., there is no triangle whose sides have different "colours").*

Note that when $n = 1, 2$, the last condition holds automatically as does the second for $n = 1$. Thus, when $n = 1$, an n-poset is in $\mathrm{SP}_n(\Sigma)$ iff it is a labelled linear order, i.e., a word. An immediate consequence of Theorem 2.1 is the fact that any "induced sub-n-poset" of an n-poset in $\mathrm{SP}_n(\Sigma)$ is also in $\mathrm{SP}_n(\Sigma)$.

PROPOSITION 2.2 [9] *$\mathrm{SP}_n(\Sigma)$ is freely generated by Σ in the variety of algebras equipped with n associative operations.*

Call an n-poset P in $\mathrm{SP}_n(\Sigma)$ *$\circ_i$-irreducible*, where $i \in [n]$, if P has no decomposition into the $\circ_i$-product of two or more n-posets (in $\mathrm{SP}_n(\Sigma)$). If this condition does not hold, call P *$\circ_i$-reducible*. Proposition 2.2 relies on the fact that each n-poset P in $\mathrm{SP}_n(\Sigma)$ is either a singleton or there is a unique i such that P is $\circ_i$-reducible. Moreover, in that case, P has, up to associativity, a unique *maximal decomposition* into a $\circ_i$-product of $\circ_i$-irreducible n-posets. We call the biposets in $\mathrm{SP}_2(\Sigma)$ *series-parallel.*

REMARK 2.3 Theorem 2.1 is a particular instance of a more general result proved in [9] which concerns Σ-labelled sets equipped with n partial orders $<_i$ and m symmetric irreflexive relations $\sim_j$. The relations $<_i$ define n associative operations and the relations $\sim_j$ define m associative and commutative operations. The general result is a common extension of the geometric characterization of series-parallel partial orders by Grabowski [13] and Valdes et al. [25], and the characterization of cographs by Corneil et al. [2]. For the case that $n = 1$ and $m = 2$, see also Boudol and Castellani [1].

Labelled n-posets satisfying the first condition of Theorem 2.1 correspond to those (labelled) reversible antisymmetric *2-structures* of Ehrenfeucht and Rozenberg [6] which are transitive. The third condition is the angularity property [7] for these 2-structures. The n-posets satisfying both the first and the third condition correspond to the T-structures of [8], while the n-posets satisfying all three conditions correspond to the uniformly nonprimitive T-structures. Uniformly nonprimitive 2-structures are studied in detail in Engelfriet et al. [4].

In the subsequent sections we will only consider biposets, and in particular series-parallel biposets, or *sp-biposets*, for short. All of our results can be generalized, in a straightforward way, to n-posets in $\mathrm{SP}_n(\Sigma)$. *We will denote* $\mathrm{SP}_2(\Sigma)$ *by* $\mathrm{SPB}(\Sigma)$ *and write* $\cdot$ *for the horizontal and* $\circ$ *for the vertical product.*

REMARK 2.4 Labelled biposets with the property that any two elements are related by exactly one of the two partial orders are called *texts* by Ehrenfeucht and Rozenberg in [8]. The sp-biposets are the uniformly nonprimitive, or *alternating texts*. Suppose that $P = (P, <_h, <_v, \lambda_P)$ is a labelled biposet which is a text. Then the relations $\sqsubset_1 = <_h \cup <_v$ and $\sqsubset_2 = <_h \cup <_v^{-1}$ are strict linear orders on P, where $<_v^{-1}$ is the reverse of the relation $<_v$. Moreover, the relations $<_h$ and $<_v$ can be recovered from these linear orders. In fact, this correspondence defines a bijection between texts and finite nonempty labelled sets equipped with two not necessarily different strict linear orders, see [8]. The operations of horizontal and parallel product on texts correspond to natural operations on (isomorphism classes of) labelled biposets equipped with two strict linear orders. It follows that $\mathrm{SPB}(\Sigma)$ can be represented as an algebra of isomorphism classes of such biposets satisfying a condition ("primitive quartet-freeness") corresponding to N-freeness. See [7] and [5] for details.

3 Recognizable and regular languages

The concept of recognizable sp-biposet languages, i.e., recognizable subsets of $\mathrm{SPB}(\Sigma)$, where Σ is a finite alphabet, can be derived from standard general notions, cf. [10]. Recall that a *bisemigroup* is an algebra $B = (B, \cdot, \circ)$ equipped with two associative binary operations $\cdot$ and $\circ$. Homomorphisms and congruences of bisemigroups are defined as usual. A congruence, or equivalence relation of a bisemigroup is of *finite index* if the partition induced by the relation has a finite number of blocks.

DEFINITION 3.1 *A language* $L \subseteq \mathrm{SPB}(\Sigma)$ *is* recognizable *if there is a finite bisemigroup* B *and a homomorphism* $h : \mathrm{SPB}(\Sigma) \to B$ *such that* $L = h^{-1}(h(L))$.

It is clear that $L \subseteq \mathrm{SPB}(\Sigma)$ is recognizable iff there is a finite index congruence ϑ of $\mathrm{SPB}(\Sigma)$ which *saturates* L, i.e., L is the union of some blocks of the partition induced by ϑ. It follows by standard arguments that the class Rec of recognizable sp-biposet languages is (effectively) closed under the boolean operations and inverse homomorphisms, so that if h is a homomorphism $\mathrm{SPB}(\Sigma) \to \mathrm{SPB}(\Sigma')$ and $L \subseteq \mathrm{SPB}(\Sigma')$ is recognizable, then so is $h^{-1}(L)$. Other closure properties will be given later.

Regular sets of sp-biposets will be defined using parenthesising automata that process sp-biposets in a sequential manner. The definition below involves a finite set Ω of parentheses. *We assume that* Ω *is partitioned into opening and closing parentheses that are in a bijective correspondence.*

DEFINITION 3.2 *A (nondeterministic)* parenthesizing automaton *is a 9-tuple* $\mathbf{S} = (S, H, V, \Sigma, \Omega, \delta, \gamma, I, F)$, *where* S *is the nonempty, finite set of* states, H *and* V *are the sets of* horizontal *and* vertical states, *which give a disjoint decomposition of* S, Σ *is the* input alphabet, Ω *is a finite set of parentheses, moreover,*

- $\delta \subseteq (H \times \Sigma \times H) \cup (V \times \Sigma \times V)$ *is the* labelled transition relation,
- $\gamma \subseteq (H \times \Omega \times V) \cup (V \times \Omega \times H)$ *is the* parenthesizing relation,
- $I, F \subseteq S$ *are the sets of* initial *and* final states, *respectively.*

DEFINITION 3.3 *Suppose that* $P \in \mathrm{SPB}(\Sigma)$ *and* $p, q \in S$. *We say that* $\mathbf{S}$ has a run on P from p to q, *denoted* $(p, P, q)_{\mathbf{S}}$ *if one of the following conditions holds.*

(Base) $P = a \in \Sigma$ *and* $(p, a, q) \in \delta$.

(HH) $p, q \in H$ *and* P *has maximal horizontal decomposition* $P = P_1 \cdot \ldots \cdot P_n$, *where* $n \geq 2$, *and* $\exists r_1, \ldots, r_{n-1} \in S$, $r_0 = p$, $r_n = q$ *such that* $(r_{i-1}, P_i, r_i)_{\mathbf{S}}$, *for all* $i \in [n]$.

(VV) $p, q \in V$ *and* P *has maximal vertical decomposition* $P = P_1 \circ \ldots \circ P_n$, *where* $n \geq 2$, *and* $\exists r_1, \ldots, r_{n-1} \in S$, $r_0 = p$, $r_n = q$ *such that* $(r_{i-1}, P_i, r_i)_{\mathbf{S}}$ *for all* $i \in [n]$.

(*HV*) $p,q \in H$ *and* P *has maximal vertical decomposition* $P = P_1 \circ \ldots \circ P_n$, *where* $n \geq 2$, *and* $\exists ({}_{k},)_k \in \Omega$, $p', q' \in V$ *and* $(p, ({}_{k}, p'), (q',)_k, q) \in \gamma$ *such that* $(p', P, q')_{\mathsf{S}}$ *holds.*

(*VH*) $p,q \in V$ *and* P *has maximal horizontal decomposition* $P = P_1 \cdot \ldots \cdot P_n$, *where* $n \geq 2$, *and* $\exists ({}_{k},)_k \in \Omega$, $p', q' \in H$ *and* $(p, ({}_{k}, p'), (q',)_k, q) \in \gamma$ *such that* $(p', P, q')_{\mathsf{S}}$ *holds.*

REMARK 3.4 Note that $(p, P, q)_{\mathsf{S}}$ implies $p,q \in H$ or $p,q \in V$. So in (HH) we have $r_1, r_2, \ldots r_{n-1} \in H$, and similarly, in (VV) we have $r_1, r_2, \ldots r_{n-1} \in V$.

The sp-biposet *language* $L(\mathsf{S})$ *accepted by the automaton* S is defined as the set of all labels of a run from an initial state to a final state. Formally,

$$L(\mathsf{S}) \;=\; \{P \in \mathrm{SPB}(\Sigma) \mid \exists i \in I,\ f \in F\ :\ (i, P, f)_{\mathsf{S}}\}.$$

We say that two automata are *equivalent* if they accept the same sp-biposet language.

DEFINITION 3.5 *An sp-biposet language* $L \subseteq \mathrm{SPB}(\Sigma)$ *is said to be* regular *if it is accepted by a parenthesizing automaton. We denote the class of all regular sp-biposet languages by* Reg.

THEOREM 3.6 Reg = Rec, *i.e., an sp-biposet language* $L \subseteq \mathrm{SPB}(\Sigma)$ *is recognizable iff* L *is regular.*

In our proof, we show how to construct a parenthesizing automaton from a finite bisemigroup, and conversely, how to construct a finite bisemigroup from a parenthesizing automaton.

REMARK 3.7 Each language $L \subseteq \mathrm{SPB}(\Sigma)$ can be recognized by a smallest bisemigroup, called the *syntactic bisemigroup* of L. This bisemigroup B_L, unique up to isomorphism, corresponds to the syntactic semigroup [22] of a word language, and to the syntactic algebra of a tree language, cf. [23]. For our present purpose it is sufficient to define B_L as the quotient of $\mathrm{SPB}(\Sigma)$ with respect to the largest congruence $\sim_L$ that saturates L. We clearly have that L is recognizable iff B_L is finite.

4 Rationality

There are several meaningful definitions of rational sets of sp-biposets. Here we will only consider the simplest of them: series rational, parallel rational, birational and generalized rational sets.

DEFINITION 4.1 *Let* $L_1, L_2 \subseteq \mathrm{SPB}(\Sigma)$, *where* Σ *is an alphabet. We define the following operations, called* horizontal *(or* series product*),* vertical *(or* parallel product*),* horizontal iteration *(or* series iteration*) and* vertical iteration *(or* parallel iteration*).*

$$\begin{aligned} L_1 \cdot L_2 &:= \{P \cdot Q \mid P \in L_1,\ Q \in L_2\} \\ L_1 \circ L_2 &:= \{P \circ Q \mid P \in L_1,\ Q \in L_2\} \\ L_1^{\cdot +} &:= \{P_1 \cdot \ldots \cdot P_n \mid P_i \in L_1,\ n \geq 1\} \\ L_1^{\circ +} &:= \{P_1 \circ \ldots \circ P_n \mid P_i \in L_1,\ n \geq 1\}. \end{aligned}$$

Moreover, if $L_1 \in \mathrm{SPB}(\Sigma)$, $\xi \notin \Sigma$, $L_2 \in \mathrm{SPB}(\Sigma \cup \{\xi\})$, *then the sp-biposet language in* $\mathrm{SPB}(\Sigma)$ *obtained by substituting (non uniformly) biposets in* L_1 *for* ξ *in the members of* L_2 *is denoted by* $L_2[L_1/\xi]$. *We refer to this operation as* ξ-substitution.

DEFINITION 4.2 *The class of* birational languages *is the least class* BRat *of sp-biposet languages containing the finite sp-biposet languages in* $\mathrm{SPB}(\Sigma)$, *for all* Σ, *and closed under union, sequential and parallel product and sequential and parallel iteration. The class of* generalized birational languages *is the least class* GRat *of sp-biposet languages containing the finite languages and closed under the above operations and complementation.*

Clearly, BRat $\subseteq$ GRat. Our first result is that GRat $\subseteq$ Rec.

LEMMA 4.3 *Every parenthesizing automaton* S *is equivalent to a parenthesizing automaton* S^h *(*S^v*, resp.) with a single initial and final state, both horizontal (vertical, resp.), such that the initial state is not the target and the final state is not the origin of any (labelled or parenthesising) transition.*

We use the above Lemma to prove:

PROPOSITION 4.4 *The class of regular (i.e., recognizable) languages is closed under ξ-substitution.*

We have already noted that Rec is closed under the Boolean operations and inverse homomorphisms. Using Proposition 4.4, we can immediately derive some further closure conditions of Rec.

COROLLARY 4.5 *The class* Rec *of recognizable (i.e., regular) sp-biposet languages is (effectively) closed under the Boolean operations, horizontal and vertical product, horizontal and vertical iteration, homomorphism and inverse homomorphism. Thus, since every finite language is recognizable, we have that* GRat $\subseteq$ Rec.

DEFINITION 4.6 *Define the* alternation depth $\mathrm{ad}(P)$ *of an sp-biposet* $P \in \mathrm{SPB}(\Sigma)$ *inductively as follows:*

- *if P is a letter in Σ, then* $\mathrm{ad}(P) = 0$,
- *if* $P = P_1 \cdot \ldots \cdot P_n$, *then* $\mathrm{ad}(P) = \max\{\mathrm{ad}(P_1), \ldots, \mathrm{ad}(P_n)\} + 1$,
- *if* $P = P_1 \circ \ldots \circ P_n$, *then* $\mathrm{ad}(P) = \max\{\mathrm{ad}(P_1), \ldots, \mathrm{ad}(P_n)\} + 1$,

where the decompositions are maximal and $n \geq 2$. The alternation depth *of an sp-biposet language L is defined as the supremum of the alternation depths of its elements:* $\mathrm{ad}(L) := \sup\{\mathrm{ad}(P) \mid P \in L\}$.

Note that $\mathrm{ad}(L)$ may be ∞. We denote by $\mathrm{BD}^{\leq n}$ the class of sp-biposet languages L with $\mathrm{ad}(L) \leq n$, and by BD the class of *bounded alternation depth sp-biposet languages* $\bigcup_{n<\infty} \mathrm{BD}^{\leq n}$.

THEOREM 4.7 BRat $=$ Rec $\cap$ BD.

Two subclasses of BRat are also of interest. The class SRat of *series rational languages* is the least class containing the finite languages closed under union, series and parallel product and series iteration. Call a language $L \subseteq \mathrm{SPB}(\Sigma)$ *series bounded* (SB) if there is a constant K such that for all $P \in L$, the length of each $<_h$-chain in P is bounded by K. *Parallel bounded languages* (PB) and *parallel rational languages* (PRat) are defined symmetrically. We clearly have that SB $\cup$ PB $\subseteq$ BD.

COROLLARY 4.8 SRat $=$ Rec $\cap$ PB *and* PRat $=$ Rec $\cap$ SB.

In the full version of the paper we will prove that it is decidable for a recognizable sp-biposet language whether it is birational, series rational or parallel rational. We do not know the answer for generalized rational languages.

Suppose that B is a bisemigroup. An *elementary $\cdot$-translation* on B is a function $f : B \to B$ of the form $f(x) = b \cdot x$ or $f(x) = x \cdot b$, where b is a fixed element of B. *Elementary $\circ$-translation* are defined in the same way. An *elementary translation* is an elementary $\cdot$-translation or an elementary $\circ$-translation. (Note that the same function can be both an elementary $\cdot$-translation and an elementary $\circ$-translation.) An *alternating translation* is any composition of elementary translations $f_1, \ldots, f_k$ such that for some i, f_i is an elementary $\cdot$-translation, and for some j, f_j is an elementary $\circ$-translation.

The following proposition provides a necessary condition for a language to be generalized rational. This result is related to Schützenberger's well-known characterization of star-free word languages, see [22].

PROPOSITION 4.9 *If a language* $L \subseteq \mathrm{SPB}(\Sigma)$ *is in* GRat *then its syntactic bisemigroup B_L satisfies the following condition: For all alternating translations $p(x)$ and $b \in B_L$, if $p^n(b) = b$ for some $n \geq 1$, then $p(b) = b$.*

PROPOSITION 4.10 SRat $\cup$ PRat $\subset$ BRat $\subset$ GRat $\subset$ Rec, *where each inclusion is proper. Moreover,* SRat *and* PRat *are incomparable with respect to set inclusion.*

Indeed, it is clear that SRat $\cup$ PRat $\subseteq$ BRat $\subseteq$ GRat $\subseteq$ Rec and that the first two inclusions are proper. As for the last inclusion, let $\Sigma = \{a\}$, and define $P_0 = a$, $P_{n+1} = a \cdot (P_n \circ a)$, $n \geq 0$. Then let $L = \{P_{2i} : i \geq 0\}$, the set consisting of every second of the P_n. The syntactic bisemigroup of L is finite but does not satisfy the condition given in Proposition 4.9, so that $L \in$ Rec $-$ GRat.

OPEN PROBLEM Does the converse of Proposition 4.9 hold?

If so, then it is decidable for a recognizable language whether or not it is generalized rational.

5 Logical definability

In this section we relate monadic second-order definable (MSO-definable) sp-biposet languages to recognizable languages.

Suppose that Σ is an alphabet. An *atomic formula* is of the form $P_a(x)$, $X(x)$, $x <_h y$ or $x <_v y$, where a is any letter in Σ, x, y are first-order variables ranging over vertices in an sp-biposet, and X is a (monadic) second-order variable ranging over subsets of the vertex set of an sp-biposet. Here, $P_a(x)$ means that vertex x is labelled by a and $X(x)$ means that x belongs to X. The atomic formulas $x <_h y$ and $x <_v y$ have their expected meanings. (We assume a fixed countable set of first-order, and a fixed countable set of second-order variables.) *Formulas* are composed from atomic formulas by the boolean connectives $\vee$ and $\neg$ and first- and second-order existential quantifiers $\exists x$ and $\exists X$. We define in the usual way when a *closed formula (sentence)* φ *holds in, or is satisfied by an sp-biposet* P, denoted $P \models \varphi$. The *language* L_φ *defined by* φ is $\{P \in \mathrm{SPB}(\Sigma) \mid P \models \varphi\}$.

DEFINITION 5.1 *We say that a language* $L \subseteq \mathrm{SPB}(\Sigma)$ *is* MSO-definable *if there is sentence* φ *with* $L = L_\varphi$.

We let MSO denote the class of MSO-definable languages in $\mathrm{SPB}(\Sigma)$, for all alphabets Σ.

It is not hard to show that MSO $\subseteq$ Rec. We can argue by formula induction. In order to do that, we first associate a language $L_\varphi \subseteq \mathrm{SPB}(\Sigma \times \mathcal{V} \times \mathcal{P}(\mathcal{W}))$ to any formula φ whose free variables are contained in the finite sets $\mathcal{V}$ of first-order and $\mathcal{W}$ of second-order variables, where $\mathcal{P}(\mathcal{W})$ denotes the powerset of $\mathcal{W}$. Our definition parallels that in [24], and makes use of the closure properties of recognizable languages given in Corollary 4.5. (An alternative way of proving MSO $\subseteq$ Rec would be through a compositionality property of the monadic theories of sp-biposets. See Kuske [18] for a general outline of this method.)

Recognizable and MSO-definable text languages, with texts defined as isomorphism classes of nonempty finite labelled sets equipped with two strict linear orders, were studied by Hoogeboom and ten Pas in [16]. (See Remark 2.4.) The notion of recognizability clearly does not depend on the concrete representation of the free bisemigroups. On the other hand, the equivalence of the representations of free bisemigroups by texts and as labelled

sp-biposets can be established within the language of first-order logic. Thus, from the (more general) equivalence results proved in [16], we immediately have:

THEOREM 5.2 Rec = MSO.

This inclusion Rec $\subseteq$ MSO is shown for texts in [16] by interpreting the "structure" of a text within the text. This method originates in [3].

COROLLARY 5.3 *The following conditions are equivalent for a language* $L \subseteq \mathrm{SPB}(\Sigma)$ *of bounded alternation depth: 1. L is recognizable. 2. L is regular. 3. L is birational. 4. L is generalized birational. 5. L is MSO-definable. When L is parallel-bounded, the above conditions are further equivalent to the condition that L is series rational.*

6 Comparison with other work

Our investigations have been influenced to a great extent by the work of Hoogeboom and ten Pas [15, 16] on text languages, in particular on logical definability, and the recent work of K. Lodaya and P. Weil, and subsequently by D. Kuske, on automata (and logic) on series-parallel posets (sp-posets), i.e., finite nonempty labelled sets equipped with a single partial order subject to the N-free condition. These posets, equipped with the series product and the parallel product, where the parallel product is now just disjoint union (hence commutative), form the free "semi-commutative" bisemigroups, cf. [13].[1] In [19, 20], Lodaya and Weil defined recognizable languages of sp-posets as well as regular languages accepted by "branching automata", and rational languages. They showed that a language of sp-posets is regular iff it is rational, and that the recognizable languages form a proper subclass of the regular languages. Aside from semi-commutativity, their notion of recognizability corresponds to ours, and the one in [15] (actually this notion is well established in a very general setting, just as the the notion of equational set, see below). On the other hand, their notion of rationality is much more general than our birationality, and although our parenthesizing automata owe much to their branching automata, they are not a non-commutative

[1] Grabowski called a bisemigroup with a neutral element a double monoid.

version of branching automata. The above differences, together with the well-known fact that rationality and recognizability do not coincide for free commutative semigroups explain why the above mentioned results of Lodaya and Weil are so different from ours.

Nevertheless Lodaya and Weil also obtained several results similar to ours. They studied bounded width poset languages that correspond to our parallel bounded biposet languages and showed in [20] that for such languages, the concepts of recognizability, regularity and series rationality are all equivalent. Moreover, Kuske proved in [18] that for bounded width poset languages, these conditions are equivalent to MSO-definability. These equivalences correspond to our Corollary 5.3, the parallel bounded case.

What we called a birational sp-biposet language corresponds to the series-parallel rational sp-poset languages of Lodaya and Weil. In [17, 18], Kuske showed that any series rational poset language is MSO-definable and that every MSO-definable poset language is recognizable. On the other hand, there easily exist recognizable but not MSO-definable sp-poset languages. In an earlier version of this paper we proposed as an open problem whether Rec is included in MSO. We have since learned that the equality Rec = MSO has been established by Hoogeboom and ten Pas in [16] for text languages, from which Theorem 5.2 follows immediately.

By a generalized rational sp-poset language Lodaya and Weil understood a language that would elsewhere be called equational. In the realm of associative operations, they correspond to the much broader class of context-free languages. We have not studied context-free biposet languages.

The main object of study in [21] is the extension of the classical framework to automata over free algebras with a single associative operation and a collection of operations not satisfying any nontrivial equations. It is shown that a suitably adapted version of branching automata captures recognizable languages, and that there exists a corresponding notion of rationality. Lodaya and Weil also discuss, in a rather indirect way, the situation when at least one of the additional operations is associative. In this case they find that the recognizable languages form a proper subclass of the regular languages which coincide with the rational languages. Their "asymmetric" notion of regularity is different from ours (which is "symmetric"), and their notion of rationality they show to correspond to regularity is much more general than ours. Corollary 4.8 also appears in [21].

Automata and languages over free bisemigroups (more precisely, free bisemigroups with identity) have also been studied in Hashiguchi et al. [14]. The elements of the free bisemigroup are represented by ordinary words (involving parentheses) in "standard form". Accordingly, ordinary finite automata are used to accept sp-biposet languages. More precisely, they define two kinds of acceptance modes: the free binoid mode and the free monoid mode. The free monoid mode is rather restricted, since the language accepted in the ordinary sense by the finite automaton should consist of only such words that are standard forms of sp-biposets. The free binoid mode is closer to our approach. We suspect that it corresponds to those parenthesising automata having a single pair of parentheses. No notion related to our recognizability, rationality, or logical definability is considered. On the other hand, they define phrase structure grammars (B-grammars) generating sp-biposet languages in standard form. (The definition takes a full page and consists of 31 items!) In particular, they define left and right linear B-grammars and show that they determine different language classes that lie somewhere between finite automata in the free monoid, and the free binoid mode.

A different two-dimensional generalization of the classical framework is provided by the *picture languages*. Pictures themselves are labeled biposets with a very regular structure. They come with two operations, corresponding to horizontal and vertical product, but these are only partially defined, cf. [11]. The notion of recognizability is based on tilings and behaves differently, since recognizable picture languages are not closed under complement and their emptiness problem is undecidable. For the description of picture languages using formal logic we refer to [12, 26].

Acknowledgment Parts of the results were obtained during the first author's visit at BRICS and the Department of Computer Science of the University of Aalborg. This author is indebted to Anna Ingólfsdóttir, Luca Aceto and Kim G. Larsen for their kind hospitality.

References

[1] G. Boudol and I. Castellani. Concurrency and atomicity. *Theoret. Comput. Sci.*, 59(1988), 25–34.

[2] D. G. Corneil, H. Lerchs and L. S. Burlinham. Complement reducible graphs. *Discr. Appl. Math.*, 3(1981), 163–174.

[3] B. Courcelle. The monadic second-order logic on graphs V: on closing the gap between definability and recognizability. *Theoret. Comput. Sci.*, 80(1991), 153–202.

[4] J. Engelfriet, T. Harju, A. Proskurowski, G. Rozenberg. Characterization and complexity of uniformly nonprimitive labeled 2-structures. *Theoret. Comput. Sci.*, 154(1996), 247–282.

[5] A. Ehrenfeucht, P. ten Pas and G. Rozenberg. Combinatorial properties of texts. *Theor. Inf. Appl.*, 27(1993), 433–464.

[6] A. Ehrenfeucht and G. Rozenberg. Theory of 2-structures, Part 1: clans, basic subclasses, and morphisms. Part 2: representation through labeled tree families. *Theoret. Comput. Sci.*, 70(1990), 277–303, 305–342.

[7] A. Ehrenfeucht and G. Rozenberg. Angular 2-structures. *Theoret. Comput. Sci.*, 92(1992), 227–248.

[8] A. Ehrenfeucht and G. Rozenberg. T-structures, T-functions and texts. *Theoret. Comput. Sci.*, 116(1993), 227–290.

[9] Z. Ésik. Free algebras for generalized automata and language theory. *RIMS Kokyuroku 1166*, Kyoto University, Kyoto, 2000, 52–58.

[10] F. Gécseg and M. Steinby. *Tree Automata*. Akadémiai Kiadó, Budapest, 1984.

[11] D. Giammarresi and A. Restivo. Two-dimensional finite state recognizability. *Fund. Inform.*, 25(1996), 399-422.

[12] D. Giammarresi, A. Restivo, S. Seibert and W. Thomas. Monadic second order logic over pictures and recognizability by tiling systems. In: proc. *STACS 94, Caen*, LNCS 775, Springer, 1994, 365–375.

[13] J. Grabowski. On partial languages. *Fund. Inform.*, 4(1981), 427–498.

[14] K. Hashiguchi, S. Ichihara and S. Jimbo. Formal languages over free binoids. *J. Autom. Lang. Comb.*, 5(2000), 219–234.

[15] H. J. Hoogeboom and P. ten Pas. Text languages in an algebraic framework. *Fund. Inform.*, 25(1996), 353–380.

[16] H. J. Hoogeboom and P. ten Pas. Monadic second-order definable text languages. *Theory Comput. Syst.*, 30(1997), 335–354.

[17] D. Kuske. Infinite series-parallel posets: logic and languages. In: proc. *ICALP 2000*, LNCS 1853, Springer, 2001, 648–662.

[18] D. Kuske. Towards a language theory for infinite N-free pomsets. To appear.

[19] K. Lodaya and P. Weil. Kleene iteration for parallelism. In: proc. *FST & TCS 98*, LNCS 1530, Springer-Verlag, 1998, 355–366.

[20] K. Lodaya and P. Weil. Series-parallel languages and the bounded-width property. *Theoret. Comput. Sci.*, 237(2000), 347–380.

[21] K. Lodaya and P. Weil. Rationality in algebras with series operation. *Inform. and Comput.*, to appear.

[22] J.-E. Pin. *Varieties of Formal Languages*. Plenum Publishing Corp., New York, 1986.

[23] M. Steinby. General varieties of tree languages. *Theoret. Comput. Sci.*, 205(1998), 1–43.

[24] H. Straubing. *Automata, Formal Logic and Circuit Complexity*. Birkhauser, Boston, 1994.

[25] J. Valdes, R. E. Tarjan, and E. L. Lawler. The recognition of series-parallel digraphs. *SIAM J. Comput.*, 11(1982), 298–313.

[26] Th. Wilke. Star-free picture expressions are strictly weaker than first-order logic. In: proc. *ICALP 97*, LNCS 1256, Springer-Verlag, 1997, 347–357.

Recent BRICS Report Series Publications

RS-01-24 Zoltán Ésik and Zoltán L. Németh. *Automata on Series-Parallel Biposets*. June 2001. 15 pp. To appear in Kuich, editor, *5th International Conference*, Developments in Language Theory DLT '01 Proceedings, LNCS, 2001.

RS-01-23 Olivier Danvy and Lasse R. Nielsen. *Defunctionalization at Work*. June 2001. Extended version of an article to appear in De Nicola, editor, *3rd International Conference on Principles and Practice of Declarative Programming*, PPDP '01 Proceedings, 2001.

RS-01-22 Zoltán Ésik. *The Equational Theory of Fixed Points with Applications to Generalized Language Theory*. June 2001. 21 pp. To appear in Kuich, editor, *5th International Conference*, Developments in Language Theory DLT '01 Proceedings, LNCS, 2001.

RS-01-21 Luca Aceto, Zoltán Ésik, and Anna Ingólfsdóttir. *Equational Theories of Tropical Semirings*. June 2001. 52 pp. Extended abstracts of parts of this paper have appeared in Honsell and Miculan, editors, *Foundations of Software Science and Computation Structures*, FoSSaCS '01 Proceedings, LNCS 2030, 2000, pages 42–56 and in Gaubert and Loiseau, editors, *Workshop on Max-plus Algebras and their Applications to Discrete-event Systems, Theoretical Computer Science, and Optimization*, MAX-PLUS '01 Proceedings, IFAC (International Federation of Automatic Control) IFAC Publications, 2001.

RS-01-20 Catuscia Palamidessi and Frank D. Valencia. *A Temporal Concurrent Constraint Programming Calculus*. June 2001. 31 pp.

RS-01-19 Jiří Srba. *On the Power of Labels in Transition Systems*. June 2001. 23 pp. Full and extended version of Larsen and Nielsen, editors, *Concurrency Theory: 12th International Conference*, CONCUR '01 Proceedings, LNCS, 2001.

RS-01-18 Katalin M. Hangos, Zsolt Tuza, and Anders Yeo. *Some Complexity Problems on Single Input Double Output Controllers*. 2001. 27 pp.