



---

Basic Research in Computer Science

BRICS RS-02-44 Ésik & Németh: Higher Dimensional Automata

## Higher Dimensional Automata

Zoltán Ésik  
Zoltán L. Németh

BRICS Report Series

RS-02-44

---

ISSN 0909-0878

November 2002

**Copyright © 2002, Zoltán Ésik & Zoltán L. Németh.  
BRICS, Department of Computer Science  
University of Aarhus. All rights reserved.**

**Reproduction of all or part of this work  
is permitted for educational or research use  
on condition that this copyright notice is  
included in any copy.**

**See back inner page for a list of recent BRICS Report Series publications.  
Copies may be obtained by contacting:**

**BRICS  
Department of Computer Science  
University of Aarhus  
Ny Munkegade, building 540  
DK-8000 Aarhus C  
Denmark  
Telephone: +45 8942 3360  
Telefax: +45 8942 3255  
Internet: BRICS@brics.dk**

**BRICS publications are in general accessible through the World Wide  
Web and anonymous FTP through these URLs:**

`http://www.brics.dk`  
`ftp://ftp.brics.dk`  
**This document in subdirectory RS/02/44/**

# Higher Dimensional Automata\*

Z. Ésik and Z. L. Németh

Dept. of Computer Science  
University of Szeged  
P.O.B. 652  
6701 Szeged, Hungary

November 19, 2002

## 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 Hoogetboom and ten Pas [17], 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. An extended abstract of this paper under the title *Automata on Series-Parallel Buposets* appeared in the proceedings of *DLT 2001*, Wien, (W. Kuich, G. Rozenberg, A. Salomaa, eds.) LNCS 2295, Springer, 2002, 217-227.

Our study owes much to the work of Hoogeboom and ten Pas [16, 17] on text languages, and to the recent work of Lodaya and Weil [20, 21] and Kuske [18, 19] 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. [15]. 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  $\Sigma$  denote a finite alphabet. A  $\Sigma$ -labelled  $n$ -poset, or  $n$ -poset, for short, is a finite nonempty set  $P$  of vertices equipped with  $n$  (irreflexive) partial orders  $<_i, i \in [n] = \{1, \dots, n\}$ , and a labelling function  $\lambda : P \rightarrow \Sigma$ . A  $\Sigma$ -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. We say that  $P$  is an *induced sub- $n$ -poset of  $Q$*  if  $P \subseteq Q$  and the partial orders  $<_i^P$  of  $P$  are the restrictions of the corresponding orders  $<_i^Q$  of  $Q$ , moreover,  $\lambda_P$  is the restriction of  $\lambda_Q$ . Suppose that  $P = (P, <_1^P, \dots, <_n^P, \lambda_P)$  and  $Q = (Q, <_1^Q, \dots, <_n^Q, \lambda_Q)$  are  $\Sigma$ -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

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

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  $\text{SP}_n(\Sigma)$  denote the collection of  $n$ -posets that can be generated from the singletons by the  $n$  product operations.

THEOREM 2.1 [10] *An  $n$ -poset  $P = (P, <_1, \dots, <_n, \lambda_P)$  is in  $\text{SP}_n(\Sigma)$  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, and when  $n = 1$  the first condition implies the second. Thus, when  $n = 1$ , an  $n$ -poset is in  $\text{SP}_n(\Sigma)$  iff it is a labelled linear order, i.e., a word. An immediate consequence of Theorem 2.1 is the following fact.

COROLLARY 2.2 *If  $P$  is an induced sub- $n$ -poset of  $Q$  and  $Q \in \text{SP}_n(\Sigma)$ , then  $P \in \text{SP}_n(\Sigma)$ .*

PROPOSITION 2.3 [10]  *$\text{SP}_n(\Sigma)$  is freely generated by  $\Sigma$  in the variety of algebras equipped with  $n$  associative operations.*

Call an  $n$ -poset  $P$  in  $\text{SP}_n(\Sigma)$   $\circ_i$ -irreducible, where  $i \in [n]$ , if  $P$  has no decomposition into the  $\circ_i$ -product of two or more  $n$ -posets (in  $\text{SP}_n(\Sigma)$ ). If this condition does not hold, call  $P$   $\circ_i$ -reducible. Proposition 2.3 relies on the fact that each  $n$ -poset  $P$  in  $\text{SP}_n(\Sigma)$  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  $\text{SP}_2(\Sigma)$  *series-parallel*. The  $\circ_1$ -reducible series-parallel biposets are also called *horizontal*, and the  $\circ_2$ -reducible biposets *vertical series-parallel biposets*. For later use we note the following fact.

PROPOSITION 2.4 *If  $P = (P, <_1, \dots, <_n, \lambda)$  is in  $\text{SP}_n(\Sigma)$ , the relation  $< = <_1 \cup \dots \cup <_n$  is a linear order.*

*Proof.* Either by induction on the structure of  $P$ , or by Theorem 2.1. ■

REMARK 2.5 Theorem 2.1 is a particular instance of a more general result proved in [10] which concerns  $\Sigma$ -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 [14] and Valdes et al. [27], 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]. The papers cited above form only a small fragment of the by now extensive literature on 2-structures.

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  $\text{SP}_n(\Sigma)$ . We will denote  $\text{SP}_2(\Sigma)$  by  $\text{SPB}(\Sigma)$  and write  $\bullet$  for the horizontal and  $\circ$  for the vertical product.

REMARK 2.6 Labelled biposets with the property that any two elements are related by exactly one of the two partial orders correspond to the *texts* of Ehrenfeucht and Rozenberg [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, by Proposition 2.4, 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  $\text{SPB}(\Sigma)$  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  $\text{SPB}(\Sigma)$ , where  $\Sigma$  is a finite alphabet, can be derived from standard general notions, cf. [11]. Recall that a *bisemigroup* is an algebra  $B = (B, \bullet, \circ)$  equipped with two associative binary operations  $\bullet$  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  $L \subseteq \text{SPB}(\Sigma)$  is recognizable if there is a finite bisemigroup  $B$  and a homomorphism  $h : \text{SPB}(\Sigma) \rightarrow B$  such that  $L = h^{-1}(h(L))$ .*

It is clear that  $L \subseteq \text{SPB}(\Sigma)$  is recognizable iff there is a finite index congruence  $\vartheta$  of  $\text{SPB}(\Sigma)$  which *saturates*  $L$ , i.e.,  $L$  is the union of some blocks of the partition induced by  $\vartheta$ . It follows by standard arguments that the class  $\text{Rec}$  of recognizable sp-biposet languages is (effectively) closed under the boolean operations and inverse homomorphisms, so that if  $h$  is a homomorphism  $\text{SPB}(\Sigma) \rightarrow \text{SPB}(\Sigma')$  and  $L \subseteq \text{SPB}(\Sigma')$  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  $\Omega$  of parentheses. *We assume that  $\Omega$  is partitioned into opening and closing parentheses that are in a bijective correspondence.* We usually denote the corresponding pairs by  $(_1, )_1$  and  $(_2, )_2$ , etc.

**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  $S$ ,  $\Sigma$  is the input alphabet,  $\Omega$  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 transition relation,
- $I, F \subseteq S$  are the sets of initial and final states, respectively.

**DEFINITION 3.3** *Suppose that  $P \in \text{SPB}(\Sigma)$  and  $p, q \in S$ . We say that  $\mathbf{S} = (S, H, V, \Sigma, \Omega, \delta, \gamma, I, F)$  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  $P = P_1 \bullet \dots \bullet P_n$ , where  $n \geq 2$ , and  $\exists r_1, \dots, r_{n-1} \in S$ ,  $r_0 = p$ ,  $r_n = q$  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  $P = P_1 \circ \dots \circ P_n$ , where  $n \geq 2$ , and  $\exists r_1, \dots, r_{n-1} \in S$ ,  $r_0 = p$ ,  $r_n = q$  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  $P = P_1 \circ \dots \circ P_n$ , 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  $[p', P, q']_{\mathbf{S}}$  holds.

(VH)  $p, q \in V$  and  $P$  has maximal horizontal decomposition  $P = P_1 \bullet \dots \bullet P_n$ , 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  $[p', P, q']_{\mathbf{S}}$  holds.

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

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

$$L(\mathbf{S}) = \{P \in \text{SPB}(\Sigma) \mid \exists i \in I, f \in F : [i, P, f]_{\mathbf{S}}\}.$$

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

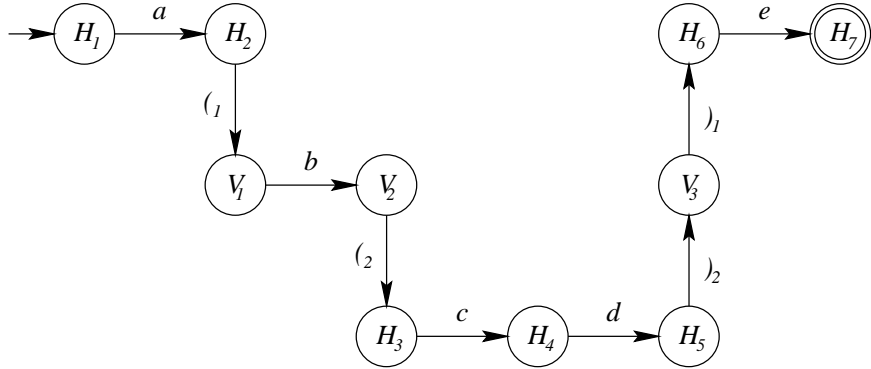


Figure 1: A parenthesizing automaton accepting  $\{a \bullet (b \circ (c \bullet d)) \bullet e\}$ .

EXAMPLE 3.5 The automaton given on Figure 1 accepts the single sp-biposet  $a \bullet (b \circ (c \bullet d)) \bullet e$ . The horizontal states are those labelled  $H_i$  and the vertical states those labelled  $V_j$ , for some  $i$  and  $j$ . There is a single initial state,  $H_1$ , and a single final state,  $H_7$ .



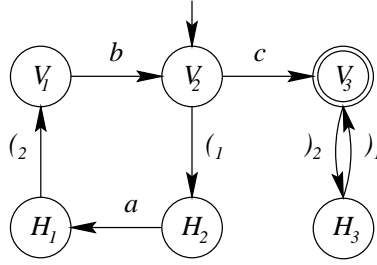


Figure 2: An automaton accepting  $\{c, a \bullet (b \circ c), a \bullet (b \circ (a \bullet (b \circ c))), \dots\}$ .

EXAMPLE 3.6 The automaton shown in Figure 2 accepts the infinite language

$$\{c, a \bullet (b \circ c), a \bullet (b \circ (a \bullet (b \circ c))), \dots\}.$$

This language is the least solution of the fixed point equation  $X = \{a\} \bullet (\{b\} \circ X) \cup \{c\}$ . (See Section 4 for the definition of the  $\bullet$  and  $\circ$  operations on languages.)

DEFINITION 3.7 An *sp-biposet language*  $L \subseteq \text{SPB}(\Sigma)$  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 3.8  $\text{Rec} = \text{Reg}$ , *i.e.*, an *sp-biposet language*  $L \subseteq \text{SPB}(\Sigma)$  is *recognizable* iff  $L$  is *regular*.

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

Let  $L \subseteq \text{SPB}(\Sigma)$  be an *sp-biposet language*, and let  $\mathbf{S}$  be a parenthesizing automaton accepting  $L$ . Define the relation  $\cong$  on  $\text{SPB}(\Sigma)$  as follows:

$$P \cong Q \Leftrightarrow (\forall p, q \in S : [p, P, q]_{\mathbf{S}} \Leftrightarrow [p, Q, q]_{\mathbf{S}}).$$

It is clear that  $\cong$  is an equivalence relation on  $\text{SPB}(\Sigma)$  and that it has index at most  $2^{m^2+n^2}$ , where  $m$  and  $n$  denote the number of horizontal and vertical states of  $\mathbf{S}$ , respectively. We verify that  $\cong$  is a congruence relation that saturates  $L$ . Suppose that  $P \cong Q$ , and let  $R$  be an arbitrary element of  $\text{SPB}(\Sigma)$ . We need to show that  $P \bullet R \cong Q \bullet R$ ,  $R \bullet P \cong R \bullet Q$ ,  $P \circ R \cong Q \circ R$  and  $R \circ P \cong R \circ Q$ . The argument is based on the following lemma which is a straightforward consequence of Definition 3.3.

LEMMA 3.9 *The following hold for any sp-biposets  $P$  and  $R$  in  $\text{SPB}(\Sigma)$ .*

(i) If  $p, q \in H$ , then  $[p, P \bullet R, q]_{\mathbf{S}} \Leftrightarrow \exists r \in H : [p, P, r]_{\mathbf{S}}$  and  $[r, R, q]_{\mathbf{S}}$ .

(ii) If  $p, q \in V$ , then  $[p, P \bullet R, q]_{\mathbf{S}} \Leftrightarrow \exists (k, )_k \in \Omega, p', r', q' \in H :$   
 $(p, (k, p')) \in \gamma, [p', P, r']_{\mathbf{S}}, [r', R, q']_{\mathbf{S}},$  and  $(q', )_k, q) \in \gamma.$

*Similar statements hold for the vertical product.*

We now show that  $P \bullet R \cong Q \bullet R$ . First, let  $p, q \in H$ , then by Lemma 3.9,

$$\begin{aligned} [p, P \bullet R, q]_{\mathbf{S}} &\Leftrightarrow \exists r : [p, P, r]_{\mathbf{S}} \text{ and } [r, R, q]_{\mathbf{S}} \\ &\Leftrightarrow \exists r : [p, Q, r]_{\mathbf{S}} \text{ and } [r, R, q]_{\mathbf{S}} \\ &\Leftrightarrow [p, Q \bullet R, q]_{\mathbf{S}} . \end{aligned}$$

In the other case  $p, q \in V$ , and

$$\begin{aligned} [p, P \bullet R, q]_{\mathbf{S}} &\Leftrightarrow \exists (k, )_k \in \Omega, p', r', q' \in H : (p, (k, p')) \in \gamma, \\ &\quad [p', P, r']_{\mathbf{S}}, [r', R, q']_{\mathbf{S}}, (q', )_k, q) \in \gamma \\ &\Leftrightarrow \exists (k, )_k \in \Omega, p', r', q' \in H : (p, (k, p')) \in \gamma, \\ &\quad [p', Q, r']_{\mathbf{S}}, [r', R, q']_{\mathbf{S}}, (q', )_k, q) \in \gamma \\ &\Leftrightarrow [p, Q \bullet R, q]_{\mathbf{S}} . \end{aligned}$$

So  $P \cong Q$  implies  $P \bullet R \cong Q \bullet R$ . One can verify similarly that  $P \cong Q$  also implies  $R \bullet P \cong R \bullet Q, P \circ R \cong Q \circ R$  and  $R \circ P \cong R \circ Q$ . Thus,  $\cong$  is a congruence relation on  $\text{SPB}(\Sigma)$ .

Finally,  $\cong$  saturates  $L$ , since  $P \in L$  and  $P \cong Q$  imply  $Q \in L$ :

$$\begin{aligned} P \in L &\Rightarrow \exists i \in I, \exists f \in F : [i, P, f]_{\mathbf{S}} \\ &\Rightarrow \exists i \in I, \exists f \in F : [i, Q, f]_{\mathbf{S}} \\ &\Rightarrow Q \in L. \end{aligned}$$

Thus  $\text{Reg} \subseteq \text{Rec}$ .

As for the inclusion  $\text{Rec} \subseteq \text{Reg}$ , let  $B$  be a finite bisemigroup,  $\varphi : \text{SPB}(\Sigma) \rightarrow B$  a homomorphism, and  $F \subseteq B$ , and let  $L = \varphi^{-1}(F)$ . Without loss of generality, we may assume that  $B$  has an element, denoted by 1, which is unit for both the horizontal and the vertical product. In order to prove that  $L$  is regular we construct a parenthesizing automaton from  $B, \varphi$  and  $F$ :

$$\mathbf{S}(B, \varphi, F) := (S, H, V, \Sigma, \Omega, \delta, \gamma, I, F), \text{ where}$$

- $S = H \cup V$ , where  $H = \{s^H \mid s \in B\}$ ,  $V = \{s^V \mid s \in B\}$ ,
- $\Omega = \{(s, \cdot)_s \mid s \in B\}$ ,
- $I = \{1^H\}$ ,
- $F = \{f^H \mid f \in F\}$ .

We define  $\delta$  and  $\gamma$  by using the operations of  $B$ . For all  $s, t \in B$  and  $a \in \Sigma$ , let

$$\begin{aligned} (s^H, a, t^H) &\in \delta \text{ if and only if } s \bullet \varphi(a) = t, \text{ and} \\ (s^V, a, t^V) &\in \delta \text{ if and only if } s \circ \varphi(a) = t. \end{aligned}$$

Moreover, let  $(s^H, (s, 1^V) \in \gamma$ , for all  $s \in B$ , and  $(u^V, \cdot)_s, t^H) \in \gamma$ , for all  $s, u, t \in B$  with  $s \bullet u = t$ . Similarly, for the vertical product, let  $(s^V, (s, 1^H) \in \gamma$ , for all  $s \in B$ , and  $(u^H, \cdot)_s, t^V) \in \gamma$ , for all  $s, u, t \in S$  such that  $s \circ u = t$ .

EXAMPLE 3.10 The construction of  $\mathbf{S}(B, \varphi, F)$  is illustrated in Figure 3. Suppose that  $s \bullet u = t$  and  $s \bullet r = w$  in  $B$ , then we have the following parenthesizing transitions in  $\mathbf{S}(B, \varphi, F)$ :

$$(s^H, (s, 1^V), (u^V, \cdot)_s, t^H), (r^V, \cdot)_s, w^H) \in \gamma.$$

If  $P, Q_1, Q_2, R, Q'$  and  $R'$  are sp-biposets such that  $\varphi(P) = s$ ,  $\varphi(Q_1) = v$ ,  $v \circ \varphi(Q_2) = u$  and  $t \bullet \varphi(R) = f$ , then the existence of a run  $[1^H, P \bullet (Q_1 \circ Q_2) \bullet R, f^H]$  can be inferred from the runs  $[1^H, P, s^H]$ ,  $[1^V, Q_1, v^V]$ ,  $[v^V, Q_2, u^V]$ ,  $[t^H, R, f^H]$ :

$$\begin{aligned} [1^V, Q_1, v^V]_{\mathbf{S}}, \quad [v^V, Q_2, u^V]_{\mathbf{S}}, & \quad \text{hence } [1^V, Q_1 \circ Q_2, u^V]_{\mathbf{S}} ; \\ (s^H, (s, 1^V) \in \gamma, \quad [1^V, Q_1 \circ Q_2, u^V]_{\mathbf{S}}, \quad (u^V, \cdot)_s, t^H) \in \gamma, & \quad \text{hence } [s^H, Q_1 \circ Q_2, t^H]_{\mathbf{S}} ; \\ [1^H, P, s^H]_{\mathbf{S}}, \quad [s^H, Q_1 \circ Q_2, t^H]_{\mathbf{S}}, \quad [t^H, R, f^H]_{\mathbf{S}}, & \quad \text{hence } [1^H, P \bullet (Q_1 \circ Q_2) \bullet \\ & \quad R, f^H]_{\mathbf{S}} . \end{aligned}$$

Similarly, if  $\varphi(Q') = r$  and  $w \bullet \varphi(R') = z$ , and if  $[1^H, P, s^H]$ ,  $[1^V, Q', r^V]$ , and  $[w^H, R', z^H]$ , then we have  $[1^H, P \bullet Q' \bullet R', z^H]_{\mathbf{S}}$ .

In order to prove our theorem it is enough to show the following:

LEMMA 3.11 *For all  $P \in \text{SPB}(\Sigma)$  and  $s, t \in B$ ,*

$$\begin{aligned} [s^H, P, t^H]_{\mathbf{S}} &\Leftrightarrow s \bullet \varphi(P) = t; \\ [s^V, P, t^V]_{\mathbf{S}} &\Leftrightarrow s \circ \varphi(P) = t. \end{aligned}$$

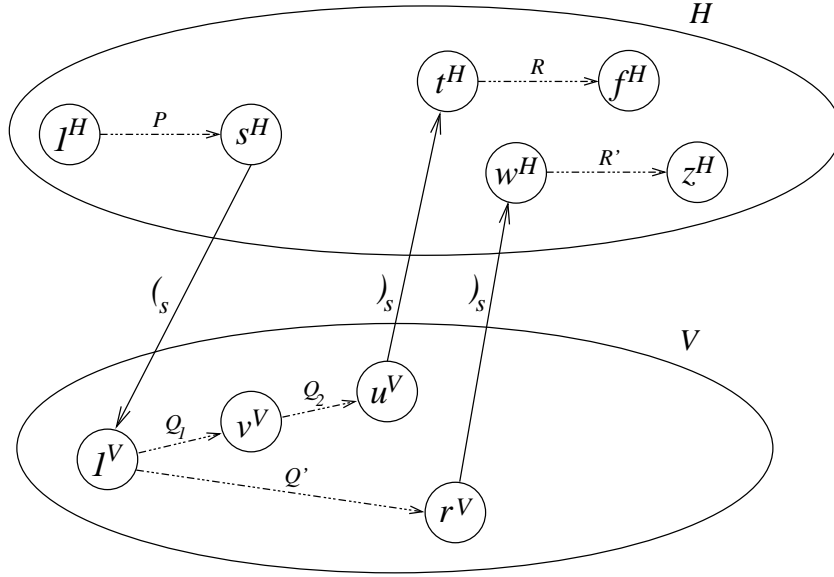


Figure 3: The construction of  $\mathbf{S}(B, \varphi, F)$ .

*Proof of Lemma 3.11.* According to Definition 3.3 there are five cases:

(Base) If  $P = a \in \Sigma$ , then our statements hold by the definition of  $\delta$ .

(HH) If  $s^H, t^H \in H$  and  $P$  has maximal horizontal decomposition  $P = P_1 \bullet P_2 \bullet \dots \bullet P_n$ ,  $n \geq 2$ , then

$$\begin{aligned}
[s^H, P, t^H]_{\mathbf{S}} &\Leftrightarrow \exists r_1^H, r_2^H, \dots, r_{n-1}^H, r_0^H = s^H, r_n^H = t^H : \\
&\quad [r_{i-1}^H, P_i, r_i^H]_{\mathbf{S}}, \quad i \in [n] \\
&\Leftrightarrow \exists r_1, r_2, \dots, r_{n-1}, r_0 = s, r_n = t : \\
&\quad r_i = r_{i-1} \bullet \varphi(P_i), \quad i \in [n] \\
&\Leftrightarrow s \bullet \varphi(P_1) \bullet \varphi(P_2) \bullet \dots \bullet \varphi(P_n) = t \\
&\Leftrightarrow s \bullet \varphi(P) = t.
\end{aligned}$$

(VV) If  $s^V, t^V \in V$  and  $P$  has maximal vertical decomposition  $P = P_1 \circ P_2 \circ \dots \circ P_n$ ,  $n \geq 2$ , then the proof is analogous to (HH).

(HV) If  $s^H, t^H \in H$  and  $P$  has maximal vertical decomposition  $P = P_1 \circ P_2 \circ \dots \circ P_n$ ,  $n \geq 2$ , then

$$\begin{aligned}
[s^H, P, t^H]_{\mathbf{S}} &\Leftrightarrow \exists u^V \in V : (s^H, (s, 1^V) \in \gamma, [1^V, P, u^V]_{\mathbf{S}}, \\
&\quad (u^V, )_s, t^H) \in \gamma \\
&\Leftrightarrow \exists u \in B : s \bullet u = t \text{ and } 1 \circ \varphi(P) = u \\
&\Leftrightarrow s \bullet \varphi(P) = t.
\end{aligned}$$

The second equivalence follows from the definition of  $\gamma$  and from the induction hypothesis.

(VH) If  $s^V, t^V \in V$  and  $P$  has maximal horizontal decomposition  $P = P_1 \bullet P_2 \bullet \dots \bullet P_n$ ,  $n \geq 2$ , then the proof is analogous to (HV). This concludes the proof of Lemma 3.11.  $\blacksquare$

*Proof of Theorem 3.8, completed.* For all  $P \in \text{SPB}(\Sigma)$ ,

$$\begin{aligned} \varphi(P) \in F &\Leftrightarrow 1 \bullet \varphi(P) = f \in F \\ &\Leftrightarrow [1^H, P, f^H]_{\mathbf{S}} \text{ for some } f \in F \\ &\Leftrightarrow \mathbf{S}(B, \varphi, F) \text{ accepts } P. \end{aligned}$$

Thus  $\varphi^{-1}(F) = L(\mathbf{S}(B, \varphi, F))$ .  $\blacksquare$

REMARK 3.12 Each language  $L \subseteq \text{SPB}(\Sigma)$  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* [23] of a word language, and to the *syntactic algebra* of a tree language, cf. [25]. For our present purpose it is sufficient to define  $B_L$  as the quotient of  $\text{SPB}(\Sigma)$  with respect to the largest congruence  $\sim_L$  that saturates  $L$ . We clearly have that  $L$  is recognizable iff  $B_L$  is finite. The natural homomorphism  $\varphi_L : \text{SPB}(\Sigma) \rightarrow B_L$  is called the *syntactic morphism* of  $L$ .

## 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 birational sets. It will be clear from the definitions that every series rational or parallel rational set is birational, and every birational set is generalized birational. We will show that the reverse inclusions fail.

DEFINITION 4.1 *Let  $L_1, L_2 \subseteq \text{SPB}(\Sigma)$ , where  $\Sigma$  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).*

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

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

**DEFINITION 4.2** *The class of birational languages is the least class  $\text{BRat}$  of sp-biposet languages containing the finite sp-biposet languages in  $\text{SPB}(\Sigma)$ , for all  $\Sigma$ , and closed under union, horizontal and vertical product and horizontal and vertical iteration. The class of generalized birational languages is the least class  $\text{GRat}$  of sp-biposet languages containing the finite languages and closed under the above operations and complementation. We denote the class of finite languages by  $\text{Fin}$  and the class of cofinite languages by  $\text{CoFin}$ .*

Clearly,  $\text{BRat} \subseteq \text{GRat}$ . In order to show that  $\text{GRat} \subseteq \text{Rec}$  we need some preparations.

**PROPOSITION 4.3** *The class of regular (i.e., recognizable) languages is closed under  $\xi$ -substitution.*

*Proof.* Suppose that  $\mathbf{S}_1 = (S_1, H_1, V_1, \Sigma, \Omega, \delta_1, \gamma_1, I_1, F_1)$  and  $\mathbf{S}_2 = (S_2, H_2, V_2, \Sigma \cup \{\xi\}, \Omega, \delta_2, \gamma_2, I_2, F_2)$  are parenthesizing automata accepting  $L_1 \subseteq \text{SPB}(\Sigma)$  and  $L_2 \subseteq \text{SPB}(\Sigma \cup \{\xi\})$ , respectively. Without loss of generality, we may assume that  $S_1 \cap S_2 = \emptyset$  and that  $\mathbf{S}_1$  and  $\mathbf{S}_2$  use the same set of parentheses  $\Omega$ . We construct an automaton  $\mathbf{R}$  accepting  $L_2[L_1/\xi]$  by “replacing” each transition  $(p, \xi, q)$  of  $\mathbf{S}_2$  by several disjoint copies of  $\mathbf{S}_1$ . These copies will serve to provide a run from  $p$  to  $q$  under each sp-biposet belonging to  $L_1$ . For the sake of simplicity, we only describe how to replace a single  $\xi$ -transition  $(p, \xi, q)$ , the other  $\xi$ -transitions are handled in the same way.

First, we construct an intermediate automaton  $\mathbf{S}'$  that consists of two disjoint copies of  $\mathbf{S}_1$  and  $\mathbf{S}_2$ . We only describe here the case when  $p$  and  $q$  are in  $H$ .

If a singleton biposet  $a$  is in  $L_1$ , that is  $(i, a, f) \in \delta_1$  for some  $i \in I_1$  and  $f \in F_1$ , then add a labelled transition  $(p, a, q)$  to  $\mathbf{S}'$ . We regard this transition *new* even if it already appeared in  $\mathbf{S}_2$ .

In order to have a run from  $p$  to  $q$  under the sp-biposets in  $L_1$  accepted with a horizontal initial state and a horizontal vertical state of  $\mathbf{S}_1$ , simulate the first and the last transitions of  $\mathbf{S}_1$  as follows. Duplicate all transitions starting from any horizontal initial state  $i$  in  $I_1$  using  $p$  as

origin instead of  $i$ , and duplicate all transitions ending in any horizontal vertical state  $f \in F_1$  using  $q$  as target instead of  $f$ .

As for the vertical sp-biposets in  $L_1$  accepted with vertical initial and final states, suppose that  $(*_$  and  $)_*$  is a new pair of parentheses not used in  $\mathbf{S}_1$  or  $\mathbf{S}_2$ . Add new transitions  $(p, (*, i)$  and  $(f, )_*, q)$  for each vertical initial state  $i$  and each vertical final state  $f$ . Formally,

$$\begin{aligned}
\mathbf{S}' &= (S', H', V', \Sigma, \Omega', \delta', \gamma', I_2, F_2), \text{ where} \\
S' &= S_1 \cup S_2; \quad H' = H_1 \cup H_2; \quad V' = V_1 \cup V_2; \\
\Omega' &= \Omega \cup \{(*, )_*\}; \\
\delta' &= \delta_1 \cup (\delta_2 \setminus \{(p, \xi, q)\}) \cup \\
&\quad \{(p, a, q) \mid a \in \Sigma, \exists i \in I_1, \exists f \in F_1 : (i, a, f) \in \delta_1\} \cup \quad (1) \\
&\quad \{(p, a, x) \mid a \in \Sigma, x \in H_1, \exists i \in I_1 \cap H_1 : (i, a, x) \in \delta_1\} \cup \quad (2) \\
&\quad \{(x, a, q) \mid a \in \Sigma, x \in H_1, \exists f \in F_1 \cap H_1 : (x, a, f) \in \delta_1\}; \quad (3) \\
\gamma' &= \gamma_1 \cup \gamma_2 \cup \\
&\quad \{(p, (, x) \mid ( \in \Omega, x \in V_1, \exists i \in I_1 \cap H_1 : (i, (, x) \in \gamma_1\} \cup \quad (4) \\
&\quad \{(x, ), q) \mid ) \in \Omega, x \in V_1, \exists f \in F_1 \cap H_1 : (x, ), f) \in \gamma_1\} \cup \quad (5) \\
&\quad \{(p, (*, i) \mid i \in I_1 \cap V_1\} \cup \quad (6) \\
&\quad \{(f, )_*, q) \mid f \in F_1 \cap V_1\}. \quad (7)
\end{aligned}$$

At this point the proof is almost finished, but some technical difficulties arise. We say that a pair of states  $(q_2, r_2)$  of an automaton is *linked to* another pair of states  $(q_1, r_1)$  if there is a pair of parentheses  $(, )$  in  $\Omega$  such that  $(q_1, (, q_2)$  and  $(r_2, ), r_1)$  are parenthesizing transitions. Suppose that  $(s, t)$  is a pair of horizontal states in  $S_1$  that is linked to a pair of vertical states  $(i, f)$ , where  $i$  is an initial and  $f$  is a final state of  $\mathbf{S}_1$ . Moreover, suppose that  $P$  is a horizontal sp-biposet such that  $[s, P, t]_{\mathbf{S}_1}$ . Then we have  $[i, P, f]_{\mathbf{S}_1}$ , so that  $P \in L_1$ . However,  $[p, P, q]_{\mathbf{S}'}$  is not necessarily true, since

$$p \xrightarrow{(*} i \xrightarrow{(} s \xrightarrow{P} t \xrightarrow{)} f \xrightarrow{)_*} q$$

is not allowed as a run from  $p$  to  $q$ , because “double parenthesizing” is not allowed.

We resolve this difficulty by regarding  $s$  as an initial state and  $t$  as a final state of  $\mathbf{S}_1$ , and simulating the transitions starting from  $s$  and the transitions ending in  $t$  by transitions starting from  $p$  and ending in  $q$ , respectively.

So, let

$$T = \{(s, t) \in H_1^2 \mid \exists i \in I_1 \cap V_1, \exists f \in F_1 \cap V_1 : (s, t) \text{ is linked to } (i, f)\}.$$

However, we need a more sophisticated method than a simply copying of the transitions as before. If another pair of states  $(s', t')$  is linked to  $(s, t) \in T$ , and if  $[s', P, t']_{\mathbf{S}_1}$ , for some vertical sp-biposet  $P$ , then  $[i, P, f]_{\mathbf{S}_1}$  is not necessarily true (double parenthesizing), so we have to avoid  $[p, P, q]_{\mathbf{R}}$ , too, unless there exist runs  $[i, P_1, r]_{\mathbf{S}_1}$  and  $[r, P_2, f]_{\mathbf{S}_1}$  such that  $P_1 \circ P_2 = P$ , where  $r$  is in  $V_1$ , or  $[p, P, q]_{\mathbf{S}_2}$  holds. Thus  $(s, (, s')$  and  $(t', ), t) \in \delta_1$  cannot be simply copied as  $(p, (, s')$  and  $(t', ), q)$ , since  $(p, (, s'), [s', P, t']$  and  $(t', ), q)$  would make a valid run  $[p, P, q]_{\mathbf{R}}$ . Therefore, we use new parentheses, say  $(first, )_{first}$  and  $(last, )_{last}$  to duplicate transitions involving ( and ), so that we do not link any  $(s', t') \in V_1 \times V_1$  to  $(p, q)$ . For each  $(s, t) \in T$  we use a disjoint copy of  $\mathbf{S}_1$  denoted by  $\mathbf{S}_1^{s,t} = (S_1^{s,t}, H_1^{s,t}, V_1^{s,t}, \Sigma, \Omega, \delta_1^{s,t}, \gamma_1^{s,t}, I_1^{s,t}, F_1^{s,t})$ . We refer to the copy of a state  $x \in S$  in  $S_1^{s,t}$  as  $x^{s,t}$ .

We are now ready to define the automaton  $\mathbf{R}$  accepting  $L_2[L_1/\xi]$ . Let

$$\begin{aligned}
\mathbf{R} &= (R, H, V, \Sigma, \Omega'', \delta'', \gamma'', I_2, F_2), \text{ where} \\
H &= H' \cup \bigcup_{(s,t) \in T} H_1^{s,t}, \quad V = V' \cup \bigcup_{(s,t) \in T} V_1^{s,t}, \quad R = H \cup V, \\
\Omega'' &= \Omega' \cup \{(first, )_{first}, (last, )_{last} \mid (, ) \in \Omega\}, \\
\delta'' &= \delta' \cup \bigcup_{(s,t) \in T} (\delta_1^{s,t} \cup \{(p, a, x^{s,t}) \mid (s, a, x) \in \delta_1\} \cup \\
&\quad \{(x^{s,t}, a, q) \mid (x, a, t) \in \delta_1\}), \\
\gamma'' &= \gamma' \cup \bigcup_{(s,t) \in T} (\gamma_1^{s,t} \cup \\
&\quad \{(p, (first, x^{s,t}), (y^{s,t}, )_{first}, z^{s,t}) \mid (s, (, x), (y, ), z) \in \gamma_1\} \cup \\
&\quad \{(x^{s,t}, (last, y^{s,t}), (z^{s,t}, )_{last}, q) \mid (x, (, y), (z, ), t) \in \gamma_1\}). \tag{8}
\end{aligned}$$

Let us write, for states  $x, y \in R$  and  $P \in \text{SPB}(\Sigma)$ ,  $[x, P, y]_{\mathbf{R}}^+$  if and only if there is a run of  $\mathbf{R}$  from  $x$  to  $y$  under  $P$ , whose labelling and parenthesizing transitions are all new, i.e., do not appear in  $\mathbf{S}_2$ .

We claim that  $\mathbf{R}$  accepts  $L_2[L_1/\xi]$ . To show this, it is enough to verify that an sp-biposet  $P$  is in  $L_1$  if and only if  $[p, P, q]_{\mathbf{R}}^+$ . We only prove that  $P \in L_1$  implies  $[p, P, q]_{\mathbf{R}}^+$ , the converse implication can be seen similarly. We have that  $P$  is in  $L_1$  if and only if there exist states  $i$  in  $I_1$  and  $f$  in  $F_1$  such that  $[i, P, f]_{\mathbf{S}_1}$ . We will distinguish between the cases  $(Base)$ ,  $(HH)$ ,  $(HV)$ ,  $(VH)$  and  $(VV)$  according to the cases of the Definition 3.3.

$(Base)$  When  $P$  is a singleton sp-biposet our claim is trivial by construction.



(*HH*), (*HV*) The cases when  $i$  and  $f$  are horizontal states are also easy. We only need to change the first and the last transition according to (2)-(5) in a run  $[i, P, f]_{\mathbf{S}_1}$  to get a corresponding run  $[p, P, q]_{\mathbf{R}}^+$ .

(*VV*) In the case when  $i$  and  $f$  are vertical states and  $P$  is a vertical sp-biposet,  $[i, P, f]_{\mathbf{S}_1}$  corresponds to a run  $(p, (*, i), [i, P, f]_{\mathbf{S}_1}, (f, *)_*, q) = [p, P, q]_{\mathbf{R}}^+$  by (6) and (7). Since  $P$  is vertical, “double parenthesizing” cannot occur in  $[p, P, q]_{\mathbf{R}}^+$ . Thus it is a valid run.

In the most difficult case (*VH*), when  $i, f \in V_1$  and  $P$  is a horizontal sp-biposet, we use the following lemma.

LEMMA 4.4 *For all  $P_1, P_2$  in  $\text{SPB}(\Sigma)$ ,  $(s, t) \in T$  and  $r \in H_1$ ,*

$$\begin{aligned} [s, P_1, r]_{\mathbf{S}_1} &\Leftrightarrow [p, P_1, r^{s,t}]_{\mathbf{R}}^+; \\ [r, P_2, t]_{\mathbf{S}_1} &\Leftrightarrow [r^{s,t}, P_2, q]_{\mathbf{R}}^+. \end{aligned}$$

The proof that  $[p, P_1, r^{s,t}]_{\mathbf{R}}^+$  implies  $[s, P_1, r]_{\mathbf{S}_1}$  and  $[r^{s,t}, P_2, q]_{\mathbf{R}}^+$  implies  $[r, P_2, t]_{\mathbf{S}_1}$  is based on the fact that the runs  $[p, P_1, r^{s,t}]_{\mathbf{R}}^+$  and  $[r^{s,t}, P_2, q]_{\mathbf{R}}^+$  have to use the transitions introduced in (8) and (9). We omit the details. Now, since  $P$  is horizontal we can write it as  $P = P_1 \bullet P_2$ . Thus, using Lemma 4.4,

$$\begin{aligned} [i, P_1 \bullet P_2, f]_{\mathbf{S}_1} &\Rightarrow \exists(, ) \in \Omega, \exists s, t \in H_1 : (i, (, s) \in \gamma_1, [s, P_1 \bullet P_2, t]_{\mathbf{S}_1}, \\ &\quad (t, ), f) \in \gamma_1 \\ &\Rightarrow (s, t) \in T, \exists r \in H_1 : [s, P_1, r]_{\mathbf{S}_1}, \text{ and } [r, P_2, t]_{\mathbf{S}_1} \\ &\Rightarrow [p, P_1, r^{s,t}]_{\mathbf{R}}^+, \text{ and } [r^{s,t}, P_2, q]_{\mathbf{R}}^+ \\ &\Rightarrow [p, P_1 \bullet P_2, q]_{\mathbf{R}}^+ \quad \blacksquare \end{aligned}$$

We have already noted that  $\text{Rec}$  is closed under the Boolean operations and inverse homomorphisms. Using Proposition 4.3, we can immediately derive some further closure properties of  $\text{Rec}$ .

COROLLARY 4.5 *The class  $\text{Rec}$  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  $\text{GRat} \subseteq \text{Rec}$ .*

*Proof.* The closure of recognizable sets under the Boolean operations and taking inverse homomorphic images comes from general observations on recognizable sets. E.g., if  $L_1$  and  $L_2$  can be recognized in bisemigroups

$B_1$  and  $B_2$ , respectively, then  $L_1 \cap L_2$  and  $L_1 \cup L_2$  can be recognized in the direct product  $B_1 \times B_2$ .

The rest of the results immediately follow from Proposition 4.3. Indeed, if  $\xi_1, \xi_2 \notin \Sigma$  then, since  $\{\xi_1 \bullet \xi_2\}$  and  $\xi_1^{+\bullet}$  are recognizable,

$$\begin{aligned} L_1 \bullet L_2 &= (\{\xi_1 \bullet \xi_2\}[L_1/\xi_1])[L_2/\xi_2], \text{ and} \\ L_1^{+\bullet} &= \xi_1^{+\bullet}[L_1/\xi_1] \end{aligned}$$

are also recognizable. Obviously, the same holds for the vertical product and the vertical iteration operation. Closure under homomorphism also follows from closure under substitution.  $\blacksquare$

However, the recognizable sets are not closed under  $\xi$ -iteration as the following classical example shows:

$$L = (a \bullet \xi \bullet b)^{\ast\xi} = \{a^n \bullet \xi \bullet b^n \mid n \geq 0\} \notin \text{Rec}.$$

Here the  $\xi$ -iteration of an sp-biposet language is defined as usual,  $L^\xi = \cup_{i \geq 0} L_i^\xi$ , where  $L_0^\xi = \{\xi\}$ , and  $L_{i+1}^\xi = L_i^\xi[L/\xi] \cup L_i^\xi$ .

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

- if  $P$  is a singleton sp-biposet (i.e. a letter in  $\Sigma$ ), then  $\text{ad}(P) = 0$ ,
- if  $P = P_1 \bullet \dots \bullet P_n$ , then  $\text{ad}(P) = \max\{\text{ad}(P_1), \dots, \text{ad}(P_n)\} + 1$ ,
- if  $P = P_1 \circ \dots \circ P_n$ , then  $\text{ad}(P) = \max\{\text{ad}(P_1), \dots, \text{ad}(P_n)\} + 1$ ,

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:  $\text{ad}(L) := \sup\{\text{ad}(P) \mid P \in L\}$ .

Note that  $\text{ad}(L)$  may be infinite. A recognizable language which has unbounded alternation depth was given in Example 3.6. We denote by  $\text{BD}^{\leq n}$  the class of sp-biposet languages  $L$  with  $\text{ad}(L) \leq n$ , and by  $\text{BD}$  the class of *bounded alternation depth sp-biposet languages*  $\bigcup_{n < \infty} \text{BD}^{\leq n}$ .

**THEOREM 4.7**  $\text{BRat} = \text{Rec} \cap \text{BD}$ .

*Proof.* The inclusion  $\text{BRat} \subseteq \text{Rec}$  follows from Corollary 4.5. It is easy to prove that  $\text{BD}$  is closed under union and the product and iteration operations, hence  $\text{BRat} \subseteq \text{BD}$ . Thus  $\text{BRat} \subseteq \text{Rec} \cap \text{BD}$ .

As for the inclusion  $\text{Rec} \cap \text{BD} \subseteq \text{BRat}$ , we show by induction on  $n$  that for any regular sp-biposet language  $L$  we have  $L^{\leq n} \in \text{BRat}$ , for all  $n \geq 0$ , where  $L^{\leq n} \subseteq L$  is the set of all biposets in  $L$  of alternation depth  $\leq n$ . Thus, if  $L \in \text{Reg} \cap \text{BD}$ , then there exists an  $n \geq 0$  such that  $L \in \text{BD}^{\leq n}$ , so  $L = L^{\leq n} \in \text{BRat}$ .

Assume that  $L = L(\mathbf{S})$ , for some parenthesizing automaton  $\mathbf{S} = (S, H, V, \Sigma, \Omega, \delta, \gamma, I, F)$ . For any states  $q_1, q_2 \in H$  or  $q_1, q_2 \in V$  and  $n \geq 0$ , denote by  $L_{q_1, q_2}^{\leq n}$  the language consisting of all sp-biposets in  $\text{SPB}(\Sigma)$  of alternation depth at most  $n$  having a run  $[q_1, P, q_2]_{\mathbf{S}}$ . As

$$L^{\leq n} = \bigcup_{i \in I, f \in F} L_{i, f}^{\leq n},$$

we conclude our proof by verifying the following lemma.

**LEMMA 4.8** *For any states  $q_1, q_2 \in H$  or  $q_1, q_2 \in V$  and  $n \geq 0$ , we have  $L_{q_1, q_2}^{\leq n} \in \text{BRat}$ .*

*Proof.* In case  $n = 0$ ,

$$L_{q_1, q_2}^{\leq 0} = \{a \mid (q_1, a, q_2) \in \delta\} \in \text{BRat}.$$

Note that the case  $n = 1$  follows from the classical Kleene theorem for finite words, since only labelling transitions are involved.

Assume that  $L_{s_1, s_2}^{\leq n} \in \text{BRat}$ , for any  $s_1$  and  $s_2$ , horizontal or vertical states of  $\mathbf{S}$ . In order to obtain a birational expression for  $L_{q_1, q_2}^{\leq n+1}$ , we introduce some new notations.

By induction, there exists a birational expression  $E_{s_1, s_2}^{\leq n}$  for each language  $L_{s_1, s_2}^{\leq n}$ . For any two horizontal states  $s_1, s_2$ , let  $HL_{s_1, s_2}$  denote the ordinary regular language of nonempty words accepted by the ordinary automaton with state set  $H$ , initial state  $s_1$  and final state  $s_2$ , whose set of input letters is the set  $H^2$  with transitions  $p_1 \xrightarrow{(p_1, p_2)} p_2$ , for all  $(p_1, p_2) \in H^2$ . Let  $HE_{s_1, s_2}$  denote an ordinary rational expression for  $HL_{s_1, s_2}$ .

When  $s_1, s_2$  are vertical states, define  $VL_{s_1, s_2}$  in the same way, and let  $VE_{s_1, s_2}$  denote a rational expression for  $VL_{s_1, s_2}$ . Let  $VE_{s_1, s_2}^{++}$  be rational expressions representing the words of  $VL_{s_1, s_2}$  having length at least 2. Now,  $L_{q_1, q_2}^{\leq n+1}$  is given by the expression

$$E_{q_1, q_2}^{\leq n+1} = HE_{q_1, q_2}[E_{s_1, s_2}^{\leq n}/(s_1, s_2)] \cup \bigcup_{\substack{q_1 \xrightarrow{\leftarrow} p_1, \\ p_2 \xrightarrow{\rightarrow} q_2}} VE_{p_1, p_2}^{++}[E_{s_1, s_2}^{\leq n}/(s_1, s_2)].$$

It is understood that when substituting in  $HE_{q_1, q_2}$ , all product operations in  $HE_{q_1, q_2}$  are replaced by horizontal product, and all stars by horizontal iteration. Similarly, each product and star in  $VE_{p_1, p_2}^{++}$  is replaced by the vertical version of the operation. ■

Two subclasses of **BRat** are also of interest.

**DEFINITION 4.9** *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  $L \subseteq \text{SPB}(\Sigma)$  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.*

Note that series rational languages are parallel bounded since **PB** is closed under the operations used in the definition of **SRat**. Moreover, if  $L$  is a parallel bounded birational language, then parallel iteration cannot be used in the construction of  $L$ , so  $L$  is series rational. Thus

$$\text{SRat} = \text{BRat} \cap \text{PB} \text{ and } \text{PRat} = \text{BRat} \cap \text{SB}.$$

**PROPOSITION 4.10**  $\text{SRat} = \text{Rec} \cap \text{PB}$  and  $\text{PRat} = \text{Rec} \cap \text{SB}$ .

*Proof.* By Theorem 4.7,  $\text{BRat} = \text{Rec} \cap \text{BD}$ , and it is not hard to see that  $\text{PB} \subseteq \text{BD}$ . Thus,

$$\text{SRat} = \text{BRat} \cap \text{PB} = \text{Rec} \cap \text{BD} \cap \text{PB} = \text{Rec} \cap \text{PB}. \quad \blacksquare$$

Our next aim is to use the previous statements to show that it is decidable for a recognizable sp-biposet language whether it is birational, series rational or parallel rational. We do not now the answer for generalized birational languages.

**DEFINITION 4.11** *Let  $B$  be an arbitrary bisemigroup and suppose that  $p, q \in B$ . We denote the fact that  $q$  is a horizontal (vertical) factor of  $p$  by*

$$\begin{aligned} p \succ_h q &\Leftrightarrow \exists r, s \in B : p = r \bullet q \bullet s \text{ or } p = r \bullet q \text{ or } p = q \bullet s; \\ p \succ_v q &\Leftrightarrow \exists r, s \in B : p = r \circ q \circ s \text{ or } p = r \circ q \text{ or } p = q \circ s. \end{aligned}$$

Since the operations  $\bullet$  and  $\circ$  are associative, the relations  $\succ_h$  and  $\succ_v$  are transitive. The following facts are clear

LEMMA 4.12 (i) If  $\varphi : B_1 \rightarrow B_2$  is a homomorphism of bisemigroups and  $p, q \in B_1$ , then  $p \succ_h q \Rightarrow \varphi(p) \succ_h \varphi(q)$ ;

(ii) If  $\varphi : B_1 \rightarrow B_2$  is a surjective homomorphism of bisemigroups and  $p', q' \in B_2$ , then  $p' \succ_h q' \Rightarrow \exists p \in \varphi^{-1}(p'), \exists q \in \varphi^{-1}(q') : p \succ_h q$ .  
Similar statements hold for  $\succ_v$ .

DEFINITION 4.13 Suppose that  $x_1, x_2, \dots$  is a finite or infinite sequence of elements of a bisemigroup. For all  $i \geq 1$ , let  $\succ_i$  be the relation  $\succ_h$  or the relation  $\succ_v$ . When  $x_i \succ_i x_{i+1}$  holds for each  $i \geq 1$ , we call the sequence  $x_1 \succ_1 x_2 \succ_2 x_3 \succ_3 \dots$  a composition chain. Moreover, if  $\succ_i = \succ_h$  for an infinite number of indices  $i$ , then we call the chain horizontally infinite. Vertically infinite composition chains are defined symmetrically. Finally, a chain is an alternating composition chain if for all  $n$ ,  $\succ_n = \succ_h$  and  $\succ_{n+1} = \succ_v$ , or  $\succ_n = \succ_v$  and  $\succ_{n+1} = \succ_h$ . (E.g.  $x_1 \succ_h x_2 \succ_v x_3 \succ_h x_4 \succ_v x_5$  is an alternating composition chain.)

PROPOSITION 4.14 For any  $L \in \text{SPB}(\Sigma)$ , let  $B_L$  denote the syntactic bisemigroup of  $L$  and let  $\varphi_L : \text{SPB}(\Sigma) \rightarrow B_L$  denote the corresponding syntactic morphism. (See Remark 3.12.) Then,

- $L \in \text{Fin} \Leftrightarrow B_L$  is finite and there is no infinite composition chain in  $B_L$  starting from some element of  $\varphi_L(L)$ ;
- $L \in \text{CoFin} \Leftrightarrow B_L$  is finite and there is no infinite composition chain in  $B_L$  starting from some element of  $B_L \setminus \varphi_L(L)$ ;
- $L \in \text{SRat} \Leftrightarrow B_L$  is finite and there is no vertically infinite composition chain in  $B_L$  starting from some element of  $\varphi_L(L)$ ;
- $L \in \text{PRat} \Leftrightarrow B_L$  is finite and there is no horizontally infinite composition chain in  $B_L$  starting from some element of  $\varphi_L(L)$ ;
- $L \in \text{BRat} \Leftrightarrow B_L$  is finite and there is no infinite alternating composition chain in  $B_L$  starting from some element of  $\varphi_L(L)$ .

*Proof.* All statements are clear. For example, for the first two claims, one observes that for every finite bisemigroup  $B$  and  $b \in B$ , there is no infinite composition chain starting from  $b$  if and only if there is a constant  $K$  such that all composition chains are of length at most  $K$ . For the last three claims one also uses Theorem 4.7 and Proposition 4.10. ■

COROLLARY 4.15 It is decidable for a recognizable sp-biposet language whether it is finite, cofinite, birational, series rational or parallel rational.

We now set out to prove that GRat is properly included in Rec. Suppose that  $B$  is a finite bisemigroup. An *elementary horizontal translation* of  $B$  is a function  $B \rightarrow B$ ,

$$\begin{aligned} x &\mapsto b \bullet x, & \text{or} \\ x &\mapsto x \bullet b, \end{aligned}$$

where  $b$  is any element of  $B$ . *Elementary vertical translations* are defined in the same way. Note that the same function can be both an elementary vertical translation and an elementary horizontal translation. A *translation*  $p$  is any composition of elementary translations. An *alternating translation* is any composition of elementary translations  $p_1, \dots, p_k$  such that some  $p_i$  is an elementary horizontal translations and some  $p_j$  (possibly  $j = i$ ) is an elementary vertical translation. We call  $B$  a *generalized aperiodic bisemigroup* if for all alternating translations  $p : B \rightarrow B$  and all  $b \in B$  and  $n \geq 1$ , if  $p^n(b) = b$  then  $p(b) = b$ .

REMARK 4.16 If  $B$  is a (finite) bisemigroup, the alternating translations of  $B$  form a semigroup with respect to function composition. It is clear from the definition that  $B$  is a generalized aperiodic bisemigroup iff the semigroup of alternating translations is aperiodic. (See [9] for the definition of aperiodic semigroups.) One might wish to call  $B$  aperiodic if the semigroup of *all* translations is aperiodic.

The statements of the following two lemmas are analogous to the corresponding facts for aperiodic semigroups, cf. [9]. We omit the proofs.

LEMMA 4.17 *A finite bisemigroup  $B$  is an alternating aperiodic semigroup iff there exists an integer  $n$  such that  $p^n = p^{n+1}$  holds for all alternating translations  $p$ .*

LEMMA 4.18 *The (finite) generalized aperiodic bisemigroups form a pseudo-variety, i.e., they are closed with respect to taking finite direct products, subalgebras and homomorphic images.*

We will prove:

PROPOSITION 4.19 *If  $L \subseteq \text{SPB}(A)$  is a generalized birational language, then the syntactic bisemigroup of  $L$  is a generalized aperiodic bisemigroup.*

Suppose that  $L_i \subseteq \text{SPB}(A)$ ,  $i = 1, 2$  are regular sp-biposet languages recognizable by the finite bisemigroups  $B_1$  and  $B_2$ , respectively. We construct a finite bisemigroup  $B = B_1 \bullet B_2$  recognizing  $L_1 \bullet L_2$ . Our construction is an appropriate modification of the Schützenberger product [24, 9]. We set

$$B := B_1 \times P(B_1 \times B_2) \times B_2,$$

where  $P(B_1 \times B_2)$  denotes the set of all subsets of  $B_1 \times B_2$ , and define the operations of horizontal and vertical product by

$$\begin{aligned} (b_1, U, b_2) \bullet (c_1, V, c_2) &:= (b_1 \bullet c_1, W, b_2 \bullet c_2), \\ (b_1, U, b_2) \circ (c_1, V, c_2) &:= (b_1 \circ c_1, \emptyset, b_2 \circ c_2), \end{aligned}$$

where

$$W := \{(u_1, u_2 \bullet c_2) : (u_1, u_2) \in U\} \cup \{(b_1 \bullet v_1, v_2) : (v_1, v_2) \in V\} \cup \{(b_1, c_2)\}.$$

It is a straightforward matter to check that both operations are associative.

Suppose now that  $\varphi_i : \text{SPB}(A) \rightarrow B_i$  is a homomorphism recognizing  $L_i$ ,  $i = 1, 2$ . Define  $\varphi : \text{SPB}(A) \rightarrow B$  by

$$P \mapsto (P\varphi_1, \{(Q_1\varphi_1, Q_2\varphi_2) : Q_1 \bullet Q_2 = P\}, P\varphi_2).$$

Thus, if  $P$  is a singleton or a vertical sp-biposet, then the middle component of  $P\varphi$  is empty. Again, it is a straightforward matter to check that  $\varphi$  is a homomorphism  $\text{SPB}(A) \rightarrow B_1 \bullet B_2$ .

Assume that  $L_i = \varphi^{-1}(B'_i)$ , where  $B'_i \subseteq B_i$ ,  $i = 1, 2$ . Let

$$B' = \{(b_1, U, b_2) : U \cap (B'_1 \times B'_2) \neq \emptyset\}.$$

Then for all  $P \in \text{SPB}(A)$ ,

$$\begin{aligned} P\varphi \in B' &\Leftrightarrow \exists P_1, P_2 : P = P_1 \bullet P_2, P_1\varphi_1 \in B'_1, P_2\varphi_2 \in B'_2 \\ &\Leftrightarrow P \in L_1 \bullet L_2. \end{aligned}$$

Thus,  $L_1 \bullet L_2$  is recognizable by  $B_1 \bullet B_2$ .

Symmetrically, given finite bisemigroups  $B_1$  and  $B_2$  accepting regular languages  $L_1$  and  $L_2$ , one can define a finite bisemigroup  $B_1 \circ B_2$ . Thus, we have:

PROPOSITION 4.20 *For any finite bisemigroups  $B_1, B_2$  recognizing  $L_1$  and  $L_2$ , the languages  $L_1 \bullet L_2$  and  $L_1 \circ L_2$  are recognizable by  $B_1 \bullet B_2$  and  $B_1 \circ B_2$ , respectively.*

PROPOSITION 4.21 *If  $B_1$  and  $B_2$  are generalized aperiodic bisemigroups, then so are  $B_1 \bullet B_2$  and  $B_1 \circ B_2$ .*

*Proof.* Suppose that  $(b_1, U, b_2) \in B_1 \bullet B_2$  and

$$p^n((b_1, U, b_2)) = (b_1, U, b_2) \quad (10)$$

holds for an alternating translation  $p$  of  $B_1 \bullet B_2$  and an integer  $n > 1$ . Since the operations in  $B_1 \bullet B_2$  are defined pointwise in the first and third components, there exist alternating translations  $p_i$  of  $B_i$ ,  $i = 1, 2$  such that for all  $(c_1, V, c_2) \in B_1 \bullet B_2$ , the first component of  $p(c_1, V, c_2)$  is  $p_1(c_1)$ , and the third component is  $p_2(c_2)$ . Thus, by (10),

$$\begin{aligned} p^n((b_1, U, b_2)) &= (p_1^n(b_1), U, p_2^n(b_2)) \\ &= (b_1, U, b_2). \end{aligned}$$

Since  $B_1$  and  $B_2$  are generalized aperiodic bisemigroups, it follows that  $p_1(b_1) = b_1$  and  $p_2(b_2) = b_2$ , so that

$$p((b_1, U, b_2)) = (b_1, W, b_2), \quad (11)$$

for some  $W$ . To complete the proof, we need to show that  $W = U$ . Write  $p$  as the composition  $p'' \circ q \circ p'$ , where  $p', p''$  are translations and  $q$  is an elementary vertical translation of  $B_1 \bullet B_2$ . Functions  $p_1$  and  $p_2$  can in turn be written as compositions  $p_1 = p_1'' \circ q_1 \circ p_1'$  and  $p_2 = p_2'' \circ q_2 \circ p_2'$ , where  $p_i', p_i''$  are translations of  $B_i$  and  $q_i$  is an elementary vertical translation of  $B_i$ ,  $i = 1, 2$ . Thus, by the definition of the vertical product operation in  $B_1 \bullet B_2$ , the middle component of  $q((c_1, V, c_2))$  is  $\emptyset$ , for all  $(c_1, V, c_2) \in B_1 \bullet B_2$ . It follows that the middle component of  $p((c_1, V, c_2))$  does not depend on  $V$ , since it equals the middle component of  $p''((d_1, \emptyset, d_2))$ , where  $d_i = q_i(p_i'(c_i))$ ,  $i = 1, 2$ . It follows now from (11) by induction on  $k$  that  $p^k((b_1, U, b_2)) = (b_1, W, b_2)$ , for all  $k \geq 1$ . When  $k = n$ , by (10), we have that  $W = U$ . We have thus proved that  $B_1 \bullet B_2$  is a generalized aperiodic bisemigroup. The same argument proves that  $B_1 \circ B_2$  is also generalized aperiodic. ■

We now turn to the iteration operations. Suppose that  $L \subseteq \text{SPB}(A)$  is a regular language recognized by a finite bisemigroup  $B$ , homomorphism



$\varphi : \text{SPB}(A) \rightarrow B$  and set  $B' \subseteq B$ . Our next task is to construct a finite bisemigroup  $B^{+\bullet} = C$  accepting the horizontal iterate of  $L$ . We assume that  $I$  is not in  $B$  and denote by  $B^I$  the bisemigroup that results by adding  $I$  to  $B$  so that  $I$  is a unit element with respect to both the horizontal and the vertical product. We set

$$C = B \times P(B^I)^{B^I}.$$

We also regard any function in  $P(B^I)^{B^I}$  as a relation from  $B^I$  to  $B^I$  and write the composite  $V \circ U = \{(x, y) : \exists z (x, z) \in U, (z, y) \in V\}$  of two relations  $U, V$  as  $UV$ . The horizontal and vertical product operations on  $C$  are defined as follows. For all  $(b, U), (c, V) \in C$ ,

$$\begin{aligned} (b, U) \bullet (c, V) &= (b \bullet c, UV), \\ (b, U) \circ (c, V) &= (b \circ c, W), \end{aligned}$$

where

$$xW = \begin{cases} \{x \bullet (b \circ c)\} & \text{if } x \bullet (b \circ c) \notin B' \\ \{x \bullet (b \circ c), I\} & \text{otherwise.} \end{cases}$$

Note that  $W$  does not depend on  $U$  and  $V$ . Since the operations are associative in  $B$ , it follows immediately that they are also associative in  $C$ , so that  $C$  is a bisemigroup.

Let  $P \in \text{SPB}(A)$  with maximal horizontal decomposition  $P = P_1 \bullet \dots \bullet P_k$ ,  $k \geq 1$ . We define  $F_P \in P(B^I)^{B^I}$  by  $y \in xF_P$  iff  $y = x \bullet P\varphi$  or there exist  $1 \leq i_1 < i_2 < \dots < i_m \leq k + 1$  with

$$\begin{aligned} x \bullet (P_1 \bullet \dots \bullet P_{i_1-1})\varphi &\in B' \\ I \bullet (P_{i_j} \bullet \dots \bullet P_{i_{j+1}-1})\varphi &\in B', \quad j = 1, \dots, m-1, \\ I \bullet (P_{i_m} \bullet \dots \bullet P_k)\varphi &= y. \end{aligned}$$

In particular, when  $i_m = k + 1$ , we have  $I = y$ . We omit the proof of the following fact.

*Claim* For all  $P, Q$ , it holds that  $F_{P \bullet Q} = F_P F_Q$ .

We now define, for each  $P \in \text{SPB}(A)$ ,

$$P\psi := (P\varphi, F_P).$$

By the above claim, we have

$$\begin{aligned} P\psi \bullet Q\psi &= (P\varphi, F_P) \bullet (Q\varphi, F_Q) \\ &= (P\varphi \bullet Q\varphi, F_P F_Q) \\ &= ((P \bullet Q)\varphi, F_{P \bullet Q}) \\ &= (P \bullet Q)\psi. \end{aligned}$$

Also, by letting  $G$  denote the function in  $P(B^I)^{B^I}$ ,

$$xG := \begin{cases} \{x \bullet (P \circ Q)\varphi\} & \text{if } x \bullet (P \circ Q)\varphi \notin B^I \\ \{x \bullet (P \circ Q)\varphi, I\} & \text{otherwise,} \end{cases}$$

we have

$$\begin{aligned} P\psi \circ Q\psi &= (P\varphi, F_P) \circ (Q\varphi, F_Q) \\ &= (P\varphi \circ Q\varphi, G) \\ &= ((P \circ Q)\varphi, G) \\ &= (P \circ Q)\psi. \end{aligned}$$

Thus,  $\psi$  is a homomorphism. It follows from the definition of  $\psi$  that  $P \in L^{+\bullet}$  if and only if  $IU$  contains  $I$ , where  $U$  is the second component of  $P\psi$ , so that  $L^{+\bullet}$  is recognizable by  $C = B^{+\bullet}$ . Symmetrically, by interchanging the operations of horizontal and vertical product (as well as the corresponding orders), we can construct in the same way a bisemigroup  $B^{+\circ}$ . We have proved:

**PROPOSITION 4.22** *For each regular language  $L$  recognized by a finite bisemigroup  $B$ , the horizontal and vertical iterates of  $L$  are recognizable by  $B^{+\bullet}$  and  $B^{+\circ}$ , respectively.*

**PROPOSITION 4.23** *If  $B$  is a generalized aperiodic bisemigroup, then so are  $B^{+\bullet}$  and  $B^{+\circ}$ .*

*Proof.* By symmetry, it suffices to prove the claim for  $B^{+\bullet}$  that we denote by  $C$  for short. Let  $p$  denote an alternating translation of  $C$ . There exist an alternating translation  $p_1$  of  $B$  such that the first component of  $p((c, V))$  is  $p_1(c)$ , for each  $(c, V)$ .

Suppose that  $p^n((b, U)) = (b, U)$ , for some  $(b, U) \in C$  and  $n > 1$ . Then  $p_1^n(b) = b$ , so that  $p_1(b) = b$ , since  $B$  is a generalized aperiodic bisemigroup. Thus,  $p((b, U)) = (b, W)$ , for some  $W$ . Using the fact that the vertical product is independent of the second components of the elements of  $C$ , we can argue as in the proof of Proposition 4.21 that  $W$  is in fact  $U$ . ■

*Proof of Proposition 4.19.* Suppose that  $L \subseteq \text{SPB}(A)$  is a generalized rational language. It is a routine matter to check that  $B_L$ , the syntactic bisemigroup of  $L$  is a generalized aperiodic bisemigroup whenever  $L$  has 0 or 1 element. To complete the proof, we need to show that if  $B_{L_1}$

and  $B_{L_2}$  are generalized aperiodic bisemigroups, then so are  $B_{L_1 \cup L_2}$ ,  $B_{\overline{L_1}}$ ,  $B_{L_1 \bullet L_2}$ ,  $B_{L_1 \circ L_2}$ ,  $B_{L_1^+ \bullet}$  and  $B_{L_1^+ \circ}$ . All these claims follow from Lemma 4.18, Propositions 4.20, 4.21, 4.22, 4.23, together with the facts that  $L_1 \cup L_2$  is recognizable by the direct product  $B_{L_1} \times B_{L_2}$  and  $\overline{L_1}$  is recognizable by  $B_{L_1}$ . ■

**PROPOSITION 4.24**  $\text{SRat} \cup \text{PRat} \subset \text{BRat} \subset \text{GRat} \subset \text{Rec}$ , where each inclusion is proper. Moreover,  $\text{SRat}$  and  $\text{PRat}$  are incomparable with respect to set inclusion.

Indeed, it is clear that  $\text{SRat} \cup \text{PRat} \subseteq \text{BRat} \subseteq \text{GRat} \subseteq \text{Rec}$  and that the first two inclusions are proper. As for the last inclusion, let  $\Sigma = \{a\}$ , and define  $P_0 = a$ ,  $P_{n+1} = a \bullet (P_n \circ a)$ ,  $n \geq 0$ . Then let  $L = \{P_{2i} \mid i \geq 0\}$ , the set consisting of every second of the  $P_n$ . The syntactic bisemigroup  $B_L$  is finite but not generalized aperiodic. Indeed,  $B_L$  contains  $L$  and the language  $L' = \{P_{2i+1} \mid i \geq 0\}$ , and we have  $p(L) = L'$  and  $p(L') = L$  for the alternating translation  $p : x \mapsto L \bullet (x \circ L)$ . Thus  $L \in \text{Rec} - \text{GRat}$  by Proposition 4.19.

**OPEN PROBLEM** Does the converse of Proposition 4.19 hold?

If it does, then it is decidable for a recognizable language whether or not it is generalized birational.

## 5 Logical definability

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

Suppose that  $\Sigma$  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  $\Sigma$ ,  $x, y$  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)*  $\varphi$  holds in, or is satisfied by an sp-biposet  $P$ , denoted  $P \models \varphi$ . The language  $L_\varphi$  defined by  $\varphi$  is  $\{P \in \text{SPB}(\Sigma) \mid P \models \varphi\}$ .

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

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

It is not hard to show that  $\text{MSO} \subseteq \text{Rec}$ . We can argue by formula induction. In order to do that, we first associate a language  $L_\varphi \subseteq \text{SPB}(\Sigma \times \mathcal{V} \times \mathcal{P}(\mathcal{W}))$  to any formula  $\varphi$  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 [26], and makes use of the closure properties of recognizable languages given in Corollary 4.5. (An alternative way of proving  $\text{MSO} \subseteq \text{Rec}$  would be through a compositionality property of the monadic theories of sp-biposets. See Kuske [19] 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 Hooġeboom and ten Pas in [17]. (See Remark 2.6.) 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 [17], we immediately have:

THEOREM 5.2  $\text{Rec} = \text{MSO}$ .

The inclusion  $\text{Rec} \subseteq \text{MSO}$  is shown for texts in [17] 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  $L \subseteq \text{SPB}(\Sigma)$  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.*

As an illustration of Corollary 5.3, consider the sp-biposet language  $L$  over the alphabet  $\Sigma = \{a\}$  containing the sp-biposets that are a parallel product of an even number of blocks each of which is a series product of  $a$  with itself by an even number of times. Formally,

$$L = \{a^{\bullet 2n_1} \circ a^{\bullet 2n_2} \circ \dots \circ a^{\bullet 2n_{2k}} \mid k \geq 1, n_i \geq 1, i \in [2k]\}.$$

Here, for each  $m \geq 1$ ,  $a^{\bullet m}$  stands for the  $m$ -fold series product of  $a$  with itself.

It is easy to see that  $L$  is birational and has alternation depth 2. Thus  $L$  is recognizable, regular and MSO-definable. Indeed,

$$L = ((a \bullet a)^{+\bullet} \circ (a \bullet a)^{+\bullet})^{+\circ},$$

showing that  $L$  is birational.

To show that  $L$  is recognizable, define the 5-element bisemigroup  $B = \{0, 1, \mathbf{0}, \mathbf{1}, \perp\}$  so that

$$\begin{aligned} i \bullet j &= i + j \bmod 2, & i, j \in \{0, 1\}, \\ \mathbf{i} \circ \mathbf{j} &= \mathbf{i} + \mathbf{j} \bmod 2, & \mathbf{i}, \mathbf{j} \in \{\mathbf{0}, \mathbf{1}\}, \\ 0 \circ 0 &= \mathbf{0}, \\ \mathbf{0} \circ 0 = 0 \circ \mathbf{0} &= \mathbf{1}, \\ \mathbf{1} \circ 0 = 0 \circ \mathbf{1} &= \mathbf{0}. \end{aligned}$$

In the missing cases, both operations give  $\perp$ . Let  $\varphi$  denote the homomorphism  $\text{SPB}(\Sigma) \rightarrow B$  that maps  $a$  to 1. It follows that  $L = \varphi^{-1}(\{\mathbf{0}\})$ , proving that  $L$  is recognizable.  $B$  is in fact isomorphic to the syntactic bisemigroup  $B_L$ .

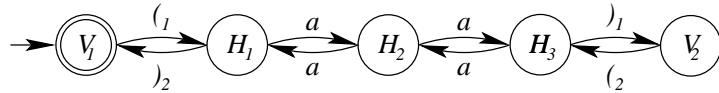


Figure 4: An automaton accepting  $L = ((a \bullet a)^{+\bullet} \circ (a \bullet a)^{+\bullet})^{+\circ}$ .

$L$  is also regular since  $L$  is accepted by the automaton shown in Figure 4. Finally,  $L$  is MSO-definable. Indeed, we can express in our logic that an sp-biposet is the disjoint sum of an even number of maximal blocks linearly ordered by  $<_h$ , each having an even number of vertices such that any two different blocks are related by  $<_v$ .

## 6 Conclusions

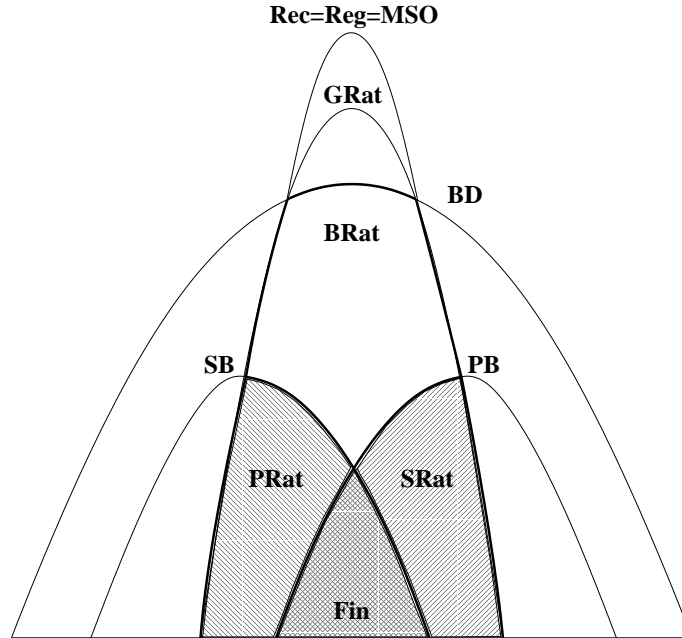


Figure 5: Comparison of classes of sp-biposets

Our main results can be summarized on Figure 5. The classes of recognizable, regular and MSO-definable languages coincide. The class **GRat** of generalized birational sp-biposet languages is strictly included in this class, which in turn strictly includes the class **BRat** of birational languages. The class **BRat** can be characterized as those recognizable sp-biposet languages that have bounded alternation depth, and the class **SRat** (**PRat**) as those that are parallel (series) bounded. It is obvious that the intersection of **SB** and **PB** is the class of finite languages. All inclusions suggested by the figure are strict.

## 7 Comparison with other work

Our investigations have been influenced to a great extent by the work of Hoogetboom and ten Pas [16, 17] 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. [14].<sup>1</sup> In [20, 21], 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 [16] (actually this notion is well established in a very general setting, just as 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 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 sp-biposet languages and showed in [21] that for such languages, the concepts of recognizability, regularity and series rationality are all equivalent. Moreover, Kuske proved in [19] 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 [18, 19], 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  $\text{Rec}$  is included in  $\text{MSO}$ . We have since learned that the equality  $\text{Rec} = \text{MSO}$  has been established by Hoogetboom and ten Pas in [17] for text languages, from which Theorem 5.2 follows immediately.

By a generalized rational sp-poset language Lodaya and Weil understood

---

<sup>1</sup>Grabowski called a bisemigroup with a neutral element a double monoid.

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 sp-biposet languages.

The main object of study in [22] 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. Proposition 4.10 also appears in [22].

Automata and languages over free bisemigroups (more precisely, free bisemigroups with identity) have also been studied in Hashiguchi et al. [15]. 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 these 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. [12]. 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 [13, 28].

**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 Ingólfssdóttir, 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] S. Eilenberg. *Automata, Languages, and Machines. Vol. B.* Academic Press, New York, 1976.
- [10] Z. Ésik. Free algebras for generalized automata and language theory. *RIMS Kokyuroku 1166*, Kyoto University, Kyoto, 2000, 52–58.
- [11] F. Gécseg and M. Steinby. *Tree Automata.* Akadémiai Kiadó, Budapest, 1984.

- [12] D. Giammarresi and A. Restivo. Two-dimensional finite state recognizability. *Fund. Inform.*, 25(1996), 399-422.
- [13] 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.
- [14] J. Grabowski. On partial languages. *Fund. Inform.*, 4(1981), 427–498.
- [15] K. Hashiguchi, S. Ichihara and S. Jimbo. Formal languages over free binoids. *J. Autom. Lang. Comb.*, 5(2000), 219–234.
- [16] H. J. Hoogeboom and P. ten Pas. Text languages in an algebraic framework. *Fund. Inform.*, 25(1996), 353–380.
- [17] H. J. Hoogeboom and P. ten Pas. Monadic second-order definable text languages. *Theory Comput. Syst.*, 30(1997), 335–354.
- [18] D. Kuske. Infinite series-parallel posets: logic and languages. In: proc. *ICALP 2000*, LNCS 1853, Springer, 2001, 648–662. for N-free pomsets.
- [19] D. Kuske. Towards a language theory for infinite N-free pomsets. *Theoret. Comput. Sci.* To appear.
- [20] K. Lodaya and P. Weil. Kleene iteration for parallelism. In: proc. *FST & TCS 98*, LNCS 1530, Springer-Verlag, 1998, 355–366.
- [21] K. Lodaya and P. Weil. Series-parallel languages and the bounded-width property. *Theoret. Comput. Sci.*, 237(2000), 347–380.
- [22] K. Lodaya and P. Weil. Rationality in algebras with series operation. *Inform. and Comput.*, 171(2001), 269-293.
- [23] J.-E. Pin. *Varieties of Formal Languages*. Plenum Publishing Corp., New York, 1986.
- [24] M.-P. Schützenberger. On finite monoids having only trivial subgroups. *Inform. and Control*, 8(1965), 190-194.
- [25] M. Steinby. General varieties of tree languages. *Theoret. Comput. Sci.*, 205(1998), 1–43.
- [26] H. Straubing. *Automata, Formal Logic and Circuit Complexity*. Birkhauser, Boston, 1994.
- [27] J. Valdes, R. E. Tarjan, and E. L. Lawler. The recognition of series-parallel digraphs. *SIAM J. Comput.*, 11(1982), 298–313.
- [28] 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-02-44 Zoltán Ésik and Zoltán L. Németh. *Higher Dimensional Automata*. November 2002. 32 pp. A preliminary version appears under the title *Automata on Series-Parallel Biposets* in Kuich, Rozenberg and Salomaa, editors, *5th International Conference, Developments in Language Theory, DLT '01 Revised Papers*, LNCS 2295, 2001, pages 217–227. This report supersedes the earlier BRICS report RS-01-24.
- RS-02-43 Mikkel Christiansen and Emmanuel Fleury. *Using IDDs for Packet Filtering*. October 2002. 25 pp.
- RS-02-42 Luca Aceto, Jens A. Hansen, Anna Ingólfssdóttir, Jacob Johnsen, and John Knudsen. *Checking Consistency of Pedigree Information is NP-complete (Preliminary Report)*. October 2002. 16 pp.
- RS-02-41 Stephen L. Bloom and Zoltán Ésik. *Axiomatizing Omega and Omega-op Powers of Words*. October 2002. 16 pp.
- RS-02-40 Luca Aceto, Willem Jan Fokkink, and Anna Ingólfssdóttir. *A Note on an Expressiveness Hierarchy for Multi-exit Iteration*. September 2002. 8 pp.
- RS-02-39 Stephen L. Bloom and Zoltán Ésik. *Some Remarks on Regular Words*. September 2002. 27 pp.
- RS-02-38 Daniele Varacca. *The Powerdomain of Indexed Valuations*. September 2002. 54 pp. Short version appears in Plotkin, editor, *Seventeenth Annual IEEE Symposium on Logic in Computer Science, LICS '02 Proceedings, 2002*, pages 299–308.
- RS-02-37 Mads Sig Ager, Olivier Danvy, and Mayer Goldberg. *A Symmetric Approach to Compilation and Decompilation*. August 2002. To appear in Neil Jones's Festschrift.
- RS-02-36 Daniel Damian and Olivier Danvy. *CPS Transformation of Flow Information, Part II: Administrative Reductions*. August 2002. 9 pp. To appear in the *Journal of Functional Programming*. This report supersedes the earlier BRICS report RS-01-40.
- RS-02-35 Patricia Bouyer. *Timed Automata May Cause Some Troubles*. August 2002. 44 pp.