Math. Struct. in Comp. Science (1997), *vol.* 7, *pp.* 701–713. Printed in the United Kingdom

Varieties generated by languages with poset operations

STEPHEN L. BLOOM[†] and ZOLTÁN ÉSIK[‡]

[†]*Stevens Institute of Technology, Department of Computer Science, Hoboken, NJ 07030*
Email: bloom@cs.stevens-tech.edu

[‡]*A. József University, Department of Computer Science, Szeged, Hungary*
Email: esik@inf.u-szeged.hu

Received 10 September 1995; revised 4 December 1996

A V-labelled poset P can induce an operation on the languages on any fixed alphabet, as well as an operation on labelled posets (as noticed by Pratt and Gischer (Pratt 1986; Gischer 1988)). For any collection X of V-labelled posets and any alphabet Σ we obtain an X-algebra Σ_X of languages on Σ. We consider the variety **Lang**(X) generated by these algebras when X is a collection of nonempty 'traceable posets'. The current paper contains several observations about this variety. First, we use one of the basic results in Bloom and Ésik (1996) to show that a concrete description of the A-generated free algebra in **Lang**(X) is the X-subalgebra generated by the singletons (labelled $a \in A$) in the X-algebra of all A-labelled posets. Equipped with an appropriate ordering, these same algebras are the free *ordered* algebras in the variety **Lang**$(X)_\leqslant$ of ordered language X-algebras. Further, if one enriches the language algebras by adding either a binary or infinitary union operation, the free algebras in the resulting variety are described by certain 'closed' subsets of the original free algebras. Second, we show that for 'reasonable sets' X, the variety **Lang**(X) has the property that for each $n \geqslant 2$, the n-generated free algebra is a subalgebra of the 1-generated free algebra. Third, knowing the free algebras enables us to show that these varieties are generated by the finite languages on a two-letter alphabet.

1. Preliminaries

In this paper, all posets are assumed finite and we identify isomorphic (labelled) posets. We let $[n]$ denote the set consisting of the first n positive integers. The set of all subsets of the free monoid Σ^* is denoted $\mathscr{P}(\Sigma^*)$. In order to avoid requiring familiarity with our paper 'Free Shuffle Algebras in Language Varieties' (Bloom and Ésik 1996), this section contains the definitions of several frequently used notions.

A **Σ^*-labelled poset**, or Σ^*-pomset $P = (|P|, \leqslant_P, \ell_P)$ consists of a poset $(|P|, \leqslant_P)$, sometimes written just $(|P|, \leqslant)$, and an assignment of a nonempty word $v\ell_P$ in Σ^* to each

[†] Supported in part by the U.S.-Hungarian Science and Technology Joint Fund under J.F. No. 351.
[‡] Supported in part by the National Foundation for Scientific Research of Hungary under grant No. T7383 and by the U.S.-Hungarian Science and Technology Joint Fund under J.F. No. 351.

vertex v in P. (Here $|P|$ denotes the underlying set of elements or 'vertices' of P, and we will sometimes just write P for this set. Thus the expression '$v \in P$' is meaningful.) A **morphism** $f : P \longrightarrow Q$ of Σ^*-labelled posets is a function $|P| \longrightarrow |Q|$ that preserves the ordering and the labelling.

Suppose that P is a Σ^*-pomset in which the label of each vertex is a word consisting of a single letter. We call such a labelled poset a **Σ-pomset**, and usually use the letter A rather than Σ for the alphabet. From now on, we use the term 'A-pomset' for 'A-labelled poset'.

We denote the empty poset by 1. Two particular operations on (labelled) posets are the sequential, or serial product $P \cdot Q$ and shuffle (or parallel) product $P \otimes Q$. Given (labelled) posets P, Q, with $|P| \cap |Q| = \varnothing$,

$$P \cdot Q := (|P| \cup |Q|, \leqslant_{P \cdot Q})$$
$$P \otimes Q := (|P| \cup |Q|, \leqslant_{P \otimes Q}),$$

where for $v, v' \in |P| \cup |Q|$,

$$v \leqslant_{P \cdot Q} v' \Leftrightarrow v \leqslant_P v' \text{ or } v \leqslant_Q v' \text{ or}$$
$$v \in |P| \text{ and } v' \in |Q|.$$
$$v \leqslant_{P \otimes Q} v' \Leftrightarrow v \leqslant_P v' \text{ or } v \leqslant_Q v'.$$

The labelling is extended to $P \otimes Q$ and $P \cdot Q$ in the obvious way. Note that the ordering $\leqslant_{P \otimes Q}$ is the disjoint union of the orderings on P and Q.

Definition 1.1. We let $\mathbf{SP}(\Sigma^*)$ denote the least class of posets containing the empty poset 1, the singleton posets σ, labelled σ, for each $\sigma \in \Sigma^*$, closed under the operations $P \cdot Q$, $P \otimes Q$. The posets in $\mathbf{SP}(\Sigma^*)$ will be called 'series-parallel' posets.

Recall that a **topological sort**, or **topological run** of a poset P is a bijection $s : [n] \longrightarrow |P|$ such that

$$s_i \leqslant_P s_j \Rightarrow i \leqslant j,$$

where s_i is the value of s on $i \in [n]$. Suppose that $(P, \leqslant_P, \ell)$ is a Σ^*-pomset. Suppose that each vertex v of P that is labelled by a word $\sigma_1 \ldots \sigma_k$, $k = k_v \geqslant 1$, is replaced by the linearly ordered poset $v = v(1) < v(2) < \ldots < v(k_v)$, in which the label of the i-th vertex $v(i)$ is σ_i. (For example, if P is a 2 element poset $\{v_1, v_2\}$ in which the two elements v_1, v_2 are unrelated, and if $v_1\ell = abb$ and $v_2\ell = ba$, the resulting poset has 5 elements: two disjoint chains, one of length 3 and one of length 2, labelled in the indicated way.) Call the resulting Σ-labelled poset $(P', \leqslant_{P'}, \ell')$ the **expansion** of $(P, \leqslant_P, \ell)$.

Definition 1.2. The expansion $(P', \leqslant_{P'}, \ell')$ of $(P, \leqslant_P, \ell)$ is denoted $P\mathbf{Exp}$. The ordering in the expansion of P is $v(i) \leqslant_{P'} v'(j)$ if either $v = v'$ and $1 \leqslant i \leqslant j \leqslant k_v$, or $v \neq v'$ and $v \leqslant_P v'$.

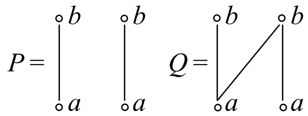
Definition 1.3. A **topological run** of a finite Σ^*-pomset P is a topological run of $P\mathbf{Exp}$. A **trace of** P is a word

$$v_1\ell' v_2\ell' \ldots$$

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

formed by concatenating the letters labelling the vertices of a topological run of P**Exp**. The set of **all traces** of P is denoted $\mathbf{Tr}(P)$.

Example 1.4. We give two $A = \{a, b\}$-pomsets with the same traces. Let P have vertices $\{1, 2, 3, 4\}$ with: $1, 2$ minimal and both labelled a; $3, 4$ maximal and labelled b; and $\mathit{pred}(3) = \{1\}$, $\mathit{pred}(4) = 2$. Let Q have vertices $\{1, 2, 3, 4\}$, with: $1, 2$ minimal, and both labelled a; $3, 4$ maximal and labelled b; and $\mathit{pred}(3) = \{1, 2\}$, and $\mathit{pred}(4) = 2$.



Then $\mathbf{Tr}(P) = \mathbf{Tr}(Q) = \{aabb, abab\}$.

Definition 1.5. A **bimonoid** $\mathrm{M} = (M, \cdot, \otimes, 1)$ consists of a monoid $(M, \cdot, 1)$ and a commutative monoid $(M, \otimes, 1)$. A **bimonoid morphism** $\mathrm{M} \longrightarrow \mathrm{M}'$ is a function $M \longrightarrow M'$ that preserves the unit and the two binary operations.

Note: The collection $\mathbf{SP}(A)$ of all A-labelled series-parallel posets is a bimonoid, as is the collection $\mathscr{P}(\Sigma^*)$ of all subsets of Σ^* under concatenation and shuffle (see Example 2.3).

Proposition 1.6. (Gischer 1988) For each set A, the bimonoid $(\mathbf{SP}(A), \cdot, \otimes, 1)$ is freely generated by A in the variety of all bimonoids.

For a fixed set A, we define the set $\Sigma(A)$ as $A \times \mathbf{N} \times [2]$[†], and write a_i instead of $(a, i, 1)$, and $\overline{a}_i$ instead of $(a, i, 2)$.

Now, since $\mathbf{SP}(A)$ is the free bimonoid, let $h_0 : \mathbf{SP}(A) \longrightarrow \mathscr{P}(\Sigma(A)^*)$ be the unique bimonoid morphism taking each letter $a \in A$ to the infinite set of two-letter words

$$ah_0 = \{a_0\overline{a}_0, a_1\overline{a}_1, \ldots\}, \quad a \in A.$$

For $P \in \mathbf{SP}(A)$, the value Ph_0 is the set of all traces of the $\Sigma(A)^*$-pomsets obtained by relabelling each vertex of P labelled $a \in A$ by a word $a_i\overline{a}_i$, for some $i \geqslant 0$. Note that this definition of Ph_0 is meaningful for any poset P, not just the series-parallel ones. We use this fact later.

Suppose that P is a $\Sigma(A)^*$-pomset in which each vertex is labelled by a two-letter word of the form $a_i\overline{a}_i$. Assume further that distinct vertices have distinct labels. Then a letter a_i occurs in a word in $\mathbf{Tr}(P)$ at most once. Each word in $u \in \mathbf{Tr}(P)$ may be written as a product

$$u = s_0 p_1 s_1 \ldots p_{n-1} s_{n-1} p_n, \tag{1}$$

where each s_i is a word on the 'open letters' a_i, $i \geqslant 0, a \in A$, and each word p_j is a word on the 'closed letters' $\overline{a}_i$, $i \geqslant 0$, $a \in A$. If the vertex v is labelled $a_i\overline{a}_i$ and the letter a_i appears in the word s_j, we say 'v occurs open in s_j'; similarly, if $\overline{a}_i$ appears in the word p_j, we say 'v occurs closed in p_j'.

A word u in $\mathbf{Tr}(P)$ is a **distinguishing trace** if a vertex v occurs open in s_0 iff v is minimal; a vertex occurs closed in p_n iff it is maximal, and if v occurs closed in p_i, then

[†] $\mathbf{N}$ denotes the set of nonnegative integers and for $n \in \mathbf{N}$, $[n] = \{1, 2, \ldots, n\}$.

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

v' occurs open in s_i iff v is an immediate predecessor of v'. Clearly, the structure of the poset can be recovered from any distinguishing trace.

Definition 1.7. An A-pomset P is **traceable** if one of the $\Sigma(A)^*$-pomsets obtained by relabelling the vertices of P labelled $a \in A$, for each $a \in A$, by distinct words $a_i\bar{a}_i$, $i \geqslant 0$, has a distinguishing trace.

The function h_0 separates any two distinct traceable posets.

Proposition 1.8. (Bloom and Ésik 1996) If P, Q are A-labelled traceable posets, then $Ph_0 = Qh_0 \Rightarrow P = Q$.

The traceable posets are characterized by a structural property.

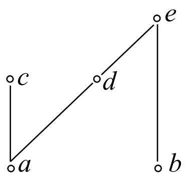
Definition 1.9. A poset P has the **zig-zag property** if for any two vertices u, v,

$$pred(u) \cap pred(v) \neq \emptyset \Rightarrow pred(u) = pred(v),$$

where $pred(u)$ is the set of all immediate predecessors of the vertex u.

Proposition 1.10. (Bloom and Ésik 1996) An A-pomset P is traceable iff P has the zig-zag property.

It follows that each series-parallel poset is traceable, but the converse is not true. Let $|P| = \{a, b, c, d, e\}$ be the poset determined by the relations $succ(a) = \{c, d\}$, $succ(d) = succ(b) = e$.



Then P has the zig-zag property, and is thus traceable, but P is not series-parallel.

Of course, the poset Q in Example 1.4 is not traceable.

Suppose that P is a poset. For each $v, v' \in P$ such that v is an immediate predecessor of v', add a new vertex u immediately above v and immediately below v'. The resulting poset is always traceable.

Remark 1.11. For series parallel posets, Proposition 1.8 was proved in Tschantz (1994) using special distinguishing traces.

2. Poset operations on languages

For any set A, the set of all A-pomsets is denoted $\mathbf{Pos}(A)$. Let V denote the countably infinite set

$$V := \{x_1, x_2, \ldots\}$$

of 'variables'. The posets in $\mathbf{Pos}(V)$ play the role of terms.

In this section, we show how a V-pomset P determines an operation on $\mathscr{P}(\Sigma^*)$, for any alphabet Σ. Suppose that

$$h : V \longrightarrow \mathscr{P}(\Sigma^*)$$

is any function. Write the value $x_i h$ as L_i. A P**-section** w of h is a choice of a word $w_v \in L_i$,

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

for each vertex $v \in P$, where the label of v is x_i. (If any language L_i is empty, h has no P-section.) We usually just say 'section', when the poset P is clear from context.

For each P-section w of h, let P_w denote the labelled poset obtained from P by relabelling the vertex v by the word w_v.

Definition 2.1. The value of P on h is given by

$$P_{lang}(h) := \bigcup_w \mathbf{Tr}(P_w),$$

the set of all traces of all posets P_w, as w varies over all P-sections of h. When the variables labelling vertices of P are in the set $\{x_1, \ldots, x_n\}$, we sometimes write $P_{lang}(h)$ as

$$P_{lang}(L_1, \ldots, L_n).$$

Proposition 2.2. Each poset operation on languages is order-preserving; *i.e.*, if $P \in \mathbf{Pos}(V)$, $x_i h \subseteq x_i h'$, all i, then

$$P_{lang}(h) \subseteq P_{lang}(h').$$

Proof. Let $S(h)$ denote the set of all sections of h, and $S(h')$ the set of sections of h'. Then $S(h) \subseteq S(h')$, so

$$\begin{aligned} P_{lang}(h) &= \bigcup_{w \in S(h)} \mathbf{Tr}(P_w) \\ &\subseteq \bigcup_{w \in S(h')} \mathbf{Tr}(P_w) \\ &= P(h'). \end{aligned}$$

□

Example 2.3. If P is the two element chain $x_1 \cdot x_2$, where the minimal element is labelled x_1 and the maximal element is labelled x_2, then

$$P_{lang}(L_1, L_2) = L_1 \cdot L_2 = \{uv : u \in L_1, v \in L_2\}.$$

If Q is the two-element poset $x_1 \otimes x_2$ consisting of two unrelated elements, one labelled x_1, the other labelled x_2, then

$$Q_{lang}(L_1, L_2) = L_1 \otimes L_2,$$

where, for languages L, L', $L \otimes L'$ is their shuffle product:

$$L \otimes L' = \{u_1 v_1 \cdots u_n v_n : u_1 \cdots u_n \in L, v_1 \cdots v_n \in L', n \geqslant 0\}.$$

Any set may take the role of V; if P is an A-pomset, and $h : A \longrightarrow \mathscr{P}(\Sigma^*)$ is any function, the value $P_{lang}(h)$ may be defined just as in Definition 2.1. If we regard $h : A \longrightarrow \mathscr{P}(\Sigma^*)$ as fixed, we obtain the function $h : \mathbf{Pos}(A) \longrightarrow \mathscr{P}(\Sigma^*)$, defined by

$$P \mapsto P_{lang}(h). \tag{2}$$

Perhaps we should give a different name to this extension of h, but the context should eliminate any possible confusion.

A collection $\mathscr{C}$ of languages in Σ^* is closed under the poset operation P_{lang} if $P_{lang}(h) \in \mathscr{C}$

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

whenever $x_i h \in \mathscr{C}$, for all labels x_i occurring in P. Clearly, $\mathscr{P}(\Sigma^*)$ is closed under any poset operation. For any set X of V-pomsets and for any alphabet Σ, we call the structure

$$\Sigma_X := (\mathscr{P}(\Sigma^*), P_{lang})_{P \in X}$$

an X**-algebra of languages** or a **language** X**-algebra**. The **variety** generated by all of the X-algebras of languages Σ_X is denoted $\mathbf{Lang}(X)$.

The following fact has a routine proof.

Proposition 2.4. For any alphabet Σ, the collection of regular subsets of Σ^* is closed under any poset operation.

3. Poset operations on posets

For any sets A, B, each poset P in $\mathbf{Pos}(B)$ determines an operation on $\mathbf{Pos}(A)$ as follows. Suppose that

$$g : b \mapsto Q_b$$

is a function mapping B to $\mathbf{Pos}(A)$.

Definition 3.1. For $P \in \mathbf{Pos}(B)$, the poset $P_{pos}(g) \in \mathbf{Pos}(A)$ is obtained from P by simultaneously replacing all vertices v of P by (disjoint copies of) the posets Q_b, with the restriction that any vertex labelled $b \in V$ in P is replaced by the poset Q_b. Given two vertices u, v in $P_{pos}(g)$, $u < v$ in $P_{pos}(g)$ iff both are inside the same poset Q_b and $u < v$ in Q_b, or u occurs in a poset Q_b determined by the vertex $u' \in P$ and v occurs in a poset $Q_{b'}$ determined by the vertex $v' \in P$, and $u' < v'$ in P. If we hold the function g fixed, we obtain the function $g : \mathbf{Pos}(B) \longrightarrow \mathbf{Pos}(A)$:

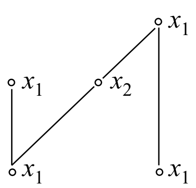
$$P \mapsto P_{pos}(g). \tag{3}$$

If $B = V$ and the labels of P are in the set $\{x_1, \ldots, x_n\}$, we may write $P_{pos}(Q_1, \ldots, Q_n)$ for $P_{pos}(g)$.

Proposition 3.2. Suppose that P is a traceable V-pomset, and $x_i h = Q_i$ is a nonempty traceable A-pomset, for each $x_i \in V$. Then $P_{pos}(Q_1, \ldots, Q_n)$ is also traceable.

If we allow some $Q_i = x_i g$ to have the value 1 for some $x_i \in V$, where 1 is the empty poset, it is possible that $P_{pos}(Q_1, \ldots, Q_n)$ is not traceable.

Example 3.3. Suppose that: P is a 5-element poset with vertices [5]. Suppose also that: 1,2 are minimal; 4,5 are maximal; and $\mathit{pred}(4) = \{1, 3\}$, $\mathit{pred}(3) = \mathit{pred}(5) = \{2\}$. Suppose the label of 3 is x_2 and that of every other vertex is x_1.

$P =$ 

Then P is traceable, but if $Q_1 = a, Q_2 = 1$, $P(Q_1, Q_2)$ is not traceable.

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

The poset and language operations are closely related. In order to describe this relationship we need some notation. Let ${}^U W$ denote the set of all functions from the set U to the set W; if $f : W \longrightarrow W'$, then ${}^U f : {}^U W \longrightarrow {}^U W'$ is the function taking $g : U \longrightarrow W$ to $gf : U \longrightarrow W'$.

Lemma 3.4. For any $h : A \longrightarrow \mathscr{P}(\Sigma^*)$, and any poset $P \in \mathbf{Pos}(V)$,

$$\begin{array}{ccc} {}^V\mathrm{Pos}\,(A) & \xrightarrow{\;{}^V h\;} & {}^V\mathscr{P}\,(\Sigma^*) \\ \Big\downarrow{\scriptstyle P_{pos}} & & \Big\downarrow{\scriptstyle P_{lang}} \\ \mathrm{Pos}\,(A) & \xrightarrow[\;h\;]{} & \mathscr{P}\,(\Sigma^*) \end{array} \qquad (4)$$

where $h : \mathbf{Pos}(A) \longrightarrow \mathscr{P}(\Sigma^*)$ was defined in (2) above.

Suppose that $P \in \mathbf{Pos}(V)$. We say a collection $\mathscr{C}$ of posets in $\mathbf{Pos}(A)$ is closed under the poset operation P_{pos} if $P_{pos}(Q_1, \ldots, Q_n) \in \mathscr{C}$ whenever the posets Q_i are in C, $x_i \in V$.

For a fixed set X of nonempty traceable V-pomsets, and any set A, let

$$\mathbf{F}_X(A)$$

denote the least collection of nonempty traceable posets in $\mathbf{Pos}(A)$ containing all singletons $a \in A$, closed under all of the operations P_{pos}, $P \in X$. Then $\mathbf{F}_X(A)$ is an X-algebra of posets.

Notation: When A has n-elements, we sometimes write $\mathbf{F}_X(n)$ for $\mathbf{F}_X(A)$.

Corollary 3.5. For any function $h : A \longrightarrow \mathscr{P}(\Sigma^*)$, the restriction to $\mathbf{F}_X(A)$ of $h : \mathbf{Pos}(A) \longrightarrow \mathscr{P}(\Sigma^*)$ is the unique function $g : \mathbf{F}_X(A) \longrightarrow \mathscr{P}(\Sigma^*)$ such that $ag = ah$ for each $a \in A$, and such that for any poset $P \in X$, the diagram

$$\begin{array}{ccc} {}^V\mathrm{F}_x\,(A) & \xrightarrow{\;{}^V g\;} & {}^V\mathscr{P}\,(\Sigma^*) \\ \Big\downarrow{\scriptstyle P_{pos}} & & \Big\downarrow{\scriptstyle P_{lang}} \\ \mathrm{F}_x\,(A) & \xrightarrow[\;g\;]{} & \mathscr{P}\,(\Sigma^*) \end{array} \qquad (5)$$

commutes.

Remark 3.6. There is no finite set X of V-labelled traceable posets such that $\mathbf{F}_X(A)$ contains all traceable A-pomsets (unless $A = \varnothing$).

4. Free algebras in $\mathbf{Lang}(X)$

In this section we prove the following theorem.

Theorem 4.1. Suppose that A is a fixed set and X is a set of nonempty traceable posets in $\mathbf{Pos}(V)$. The A-generated free algebra in $\mathbf{Lang}(X)$ is $\mathbf{F}_X(A)$.

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

Proof. First we show $\mathbf{F}_X(A)$ belongs to $\mathbf{Lang}(X)$. Indeed, let $h_0 : A \longrightarrow \mathscr{P}(\Sigma(A))$ be the function that maps the letter $a \in A$ to the infinite set of two-letter words

$$a_0\overline{a}_0, a_1\overline{a}_1, \ldots$$

Then, by Corollary 3.5, h_0 is a morphism from $\mathbf{F}_X(A)$ to $\Sigma(A)_X$. From Proposition 1.8, it follows that h_0 is injective on the traceable posets. The fact that any function $h : A \longrightarrow \mathscr{P}(\Sigma^*)$ extends to a unique homomorphism from $\mathbf{F}_X(A) \longrightarrow \Sigma_X$ follows from Corollary 3.5. □

Corollary 4.2. (Tschantz 1994; Bloom and Ésik 1996) The collection of A-labelled series-parallel posets is the A-generated free algebra in the variety generated by the language structures $(\mathscr{P}(\Sigma^*), \cdot, \otimes)$.

The X-algebras Σ_X are also **ordered algebras** with respect to the order of set inclusion, by Proposition 2.2. Let $\mathbf{Lang}(X)_\leqslant$ denote the variety of ordered algebras generated by the X-algebras of languages $(\Sigma_X, \subseteq)$.

The relation on $\mathbf{F}_X(A)$ defined by

$$Q \leqslant Q' \Leftrightarrow Q_{lang}(h_0) \subseteq Q'_{lang}(h_0)$$

is clearly a preorder. Since h_0 is injective, this relation is a partial order. By Corollary 3.5, this partial order is preserved by all of the poset operations P_{pos}, $P \in X$. Let $(\mathbf{F}_X(A), \leqslant)$ denote the resulting ordered X-algebra.

In Bloom and Ésik (1996), it was shown how the value $P(g)$ is determined by g and the set of words $P(h_0)$, for any poset $P \in \mathbf{Pos}(A)$ and for any $g : A \longrightarrow \mathscr{P}(\Sigma^*)$. It follows that

$$P(h_0) \subseteq Q(h_0) \Rightarrow P(g) \subseteq Q(g), \qquad (6)$$

for all $P, Q \in \mathbf{Pos}(A)$, all alphabets Σ and all functions $g : A \longrightarrow \mathscr{P}(\Sigma^*)$.

Theorem 4.3. $(\mathbf{F}_X(A), \leqslant)$ is the A-generated free ordered algebra in $\mathbf{Lang}(X)_\leqslant$.

Proof. By Theorem 4.1, each function $A \longrightarrow \mathscr{P}(\Sigma^*)$ extends to a unique X-algebra homomorphism $\mathbf{F}_X(A) \longrightarrow \mathscr{P}(\Sigma^*)$. By (6), the extension preserves the ordering. □

Corollary 4.4. (Bloom and Ésik 1996) The collection of A-labelled series-parallel posets ordered as above is the A-generated free ordered algebra in the variety of ordered algebras generated by the language structures $(\mathscr{P}(\Sigma^*), \cdot, \otimes, \leqslant)$.

5. The union-enriched variety Lang(X)

Let $\Sigma_X = (\mathscr{P}(\Sigma^*), P_{lang})_{P \in X}$ be an X-algebra in $\mathbf{Lang}(X)$. Since the powerset $\mathscr{P}(\Sigma^*)$ is closed under the union operation, we may consider the structures

$$\Sigma_X^+ = (\mathscr{P}(\Sigma^*), +, 0, P_{lang})_{P \in X}$$

where $L + L' = L \cup L'$ and 0 is the empty set.

We want to describe the free algebras in the variety $\mathbf{Lang}^+(X)$ generated by the structures Σ_X^+.

A poset $P \in \mathbf{F}_X(A)$ belongs to the **closure** of the subset C of $\mathbf{F}_X(A)$, in symbols,

$$P \in cl(C)$$

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

if for all alphabets Σ and all $h : A \longrightarrow \mathscr{P}(\Sigma^*)$, $P(h) \subseteq \bigcup_{Q \in C} Q(h)$. By (6), it follows that P is in the closure of C iff $P(h_0) \subseteq \bigcup_{Q \in C} Q(h_0)$. Note that $cl(C)$ is finite whenever C is. *Write* $\mathbf{I}_\omega(A)$ *for the collection of subsets of* $\mathbf{F}_X(A)$ *of the form* $cl(C)$ *for finite* $C \subseteq \mathbf{F}_X(A)$.

Remark 5.1. By definition, if $P \in cl(C)$, then $P \in \mathbf{F}_X(A)$. But if $C \subseteq \mathbf{F}_X(A)$, $P \in \mathbf{Pos}(A)$ and $h_0(P) \subseteq h_0(C)$, it is not always the case that $P \in \mathbf{F}_X(A)$. Indeed, let $A = \{a\}$ and let X consist just of the 5-element poset P of Example 3.3, with vertex i labelled $v_i, i \in [5]$. Then, if $h : V \longrightarrow \mathbf{F}_X(A)$ takes each letter to the singleton a, let $C = \{P'\}$, where $P' = P(h)$. Then there is a traceable poset Q with $h_0(Q) \subseteq h_0(P')$ that is not in $\mathbf{F}_X(A)$; indeed, one such Q has vertex set [5] with: 1, 2 minimal; 4, 5 maximal; and

$$pred(5) = \{3\}$$
$$pred(3) = pred(4) = \{1, 2\}.$$

Now we define poset operations on $\mathbf{I}_\omega(A)$. For any finite set of functions g_j: $V \longrightarrow \mathbf{F}_X(A),\ j \in J$, define

$$x_i g = cl(\bigcup_{j \in J} x_i g_j),$$

all x_i in V. For each poset $P \in \mathbf{Pos}(V)$, define

$$P_\omega(g) := cl(\{P_{lang}(g_j) : j \in J\}).$$

Further, define the $+$ operation on $\mathbf{I}_\omega(A)$ by:

$$C + C' := cl(C \cup C').$$

In this way, with 0 defined as the empty set, the structure

$$(\mathbf{I}_\omega(A), +, 0, P_\omega)_{P \in X}$$

is the same as the language algebras Σ_X^+.

By the general arguments in Bloom and Ésik (1996), we obtain the following theorem.

Theorem 5.2. For each set A, the algebra $(\mathbf{I}_\omega(A), +, 0, P_\omega)_{P \in X}$ is the A-generated free algebra in the variety $\mathbf{Lang}^+(X)$.

Corollary 5.3. (Bloom and Ésik 1996) The collection of finitely generated closed subsets of the A-labelled series-parallel posets with the indicated operations is the algebra freely generated by A in the variety generated by the language structures $(\mathscr{P}(\Sigma^*), \cdot, \otimes, +, 0)$.

If we relax the requirement that the closed subsets of $\mathbf{F}_X(A)$ be finitely generated, we obtain a structure with an infinitary sum; all the poset operations distribute over the sums, and the resulting structure $\mathbf{I}(A)$ is freely generated by A in the variety generated by the language structures enriched by arbitrary unions. The least substructure of $\mathbf{I}(A)$ containing the singletons, closed under $\cdot, \otimes$ binary $+$ and the operation $x \mapsto 1 + x + x^2 + \ldots$ is freely generated by A in the variety generated by the language structures $(\mathscr{P}(\Sigma^*), P_{lang}, +, ^*, 0, 1)_{P \in X}$. We omit the details, which follow exactly the model in Bloom and Ésik (1996).

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

6. $\mathbf{F}_X(n)$ is a subalgebra of $\mathbf{F}_X(1)$

In this section we show that for 'reasonable' sets X, the n-generated free algebra in both $\mathbf{Lang}(X)$ and $\mathbf{Lang}(X)_{\leqslant}$ is a subalgebra of the 1-generated free algebra.

Definition 6.1. A set X of traceable posets in $\mathbf{Pos}(V)$ is **reasonable** if for each $i \geqslant 1$, the poset

$$R_i = x_1 \cdot (\overbrace{x_1 \otimes \ldots \otimes x_1}^{i})$$

belongs to $\mathbf{F}_X(1)$.

Note that if both posets $x_1 \cdot x_2$ and $x_1 \otimes x_2$ belong to X, then X is reasonable.

Definition 6.2. Suppose that $B = \{b_1, \ldots, b_n\}$ and $A = \{a\}$. We define $\varphi : \mathbf{Pos}(B) \longrightarrow \mathbf{Pos}(A)$ as the function determined by the map

$$b_i \mapsto R_i \tag{7}$$

as in (3).

Proposition 6.3. The function $\varphi : \mathbf{Pos}(B) \longrightarrow \mathbf{Pos}(A)$ is injective.

Proof. It is clear how to recover Q from $Q\varphi$. □

Proposition 6.4. Let P be any poset in $\mathbf{Pos}(V)$ whose labels are among the letters $\{x_1, \ldots, x_n\}$. For any $Q_i \in \mathbf{Pos}(B)$,

$$P_{pos}(Q_1, \ldots, Q_n)\ \varphi = P_{pos}(Q_1\varphi, \ldots, Q_n\varphi).$$

Proof. This statement follows from the associativity of substitution. The poset on the left is obtained by first substituting Q_i for x_i in P, and then replacing b_i by $b_i\varphi = R_i$; the one on the right is obtained by first substituting R_i for b_i in each Q_j, and then replacing x_i by $Q_i\varphi$ in P. □

Corollary 6.5. If X is reasonable, the n-generated free algebra $\mathbf{F}_X(n)$ is isomorphic to a subalgebra of the 1-generated free algebra $\mathbf{F}_X(1)$.

Remark 6.6. The assumption that X is reasonable is used to ensure that when $Q \in \mathbf{F}_X(n)$ we have $Q\varphi \in \mathbf{F}_X(1)$.

Remark 6.7. The definition of 'reasonable' is certainly not forced, but some assumption on the set X is necessary in order to obtain the result in Corollary 6.5. Indeed, if X consists only of the poset $x_1 \cdot x_2$, then $\mathbf{F}_X(2)$ is isomorphic to the nonempty words on the two letter alphabet x_1, x_2, and $\mathbf{F}_X(1)$ is isomorphic to the nonempty words on x_1. For any morphism $\varphi : \mathbf{F}_X(2) \longrightarrow \mathbf{F}_X(1)$,

$$(x_1 \cdot x_2)\varphi = (x_2 \cdot x_1)\varphi.$$

Thus, φ is not injective.

7. The variety V_2

Let V_2 denote the subvariety of $\mathbf{Lang}(X)$ generated by the one X-algebra

$$\{a, b\}_X = (\mathscr{P}(\{a, b\}), P_{lang})_{P \in X}$$

of languages on a two-letter alphabet. In this section, we show that $V_2 = \mathbf{Lang}(X)$.

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

Theorem 7.1. Suppose that $A = \{a\}$. Define the function $h : A \longrightarrow \mathscr{P}(\{a,b\})$ by

$$a \mapsto \{ab, a^2b^2, a^4b^4, \dots, a^nb^n, \dots : n = 2^k,\ k \geqslant 0\}. \tag{8}$$

If P, Q are traceable posets in $\mathbf{Pos}(A)$ and $Ph = Qh$, then P and Q are isomorphic.

Proof. In this argument, we assume familiarity with the notion of a distinguishing trace, introduced in Bloom and Ésik (1996) and discussed above in Proposition 1.10. We will show that one can recover a distinguishing trace from the set Ph. It follows that Ph determines P up to isomorphism.

Suppose that n is a nonnegative integer. Let $(n)_2$ denote the binary representation of n, and let $B(n)$ denote the number of 1's in $(n)_2$. For any integers m, j, say that j **occurs in** m if $(m)_2$ has a 1 in position j.

Lemma 7.2. Suppose that P is a poset in $\mathbf{Pos}(A)$ with n vertices. Then

$$n = max\{B(|w|) : w \in Ph\}.$$

Indeed, the length of any word in Ph is the sum of the lengths of the words in a section of h, and is thus a sum of powers of 2. The number of 1's will be maximum when these powers are all distinct. There are at most n distinct powers possible, for any section of h.

If $w \in Ph$ and $B(|w|) = n$, the shortest binary representation of the length of w is

$$\overbrace{11\dots1}^{n}0.$$

There are many words $w \in Ph$ whose length has this binary representation. In order to obtain such a word, one vertex must be labelled ab, one is labelled a^2b^2, *etc.* Let w_0 be the alphabetically least word with

$$(|w_0|)_2 = \overbrace{11\dots1}^{n}0$$

(assuming the letters a and b are ordered as usual with $a < b$). Then

$$w_0 = a^{n_1}b^{m_1}a^{n_2}b^{m_2}\dots a^{n_k}b^{m_k},$$

where $\sum n_i = \sum m_i$, and n_1 is as large as possible, m_1 is as small as possible, etc., in order to obtain a trace. Since w_0 is a trace, if we write each superscript as a sum of distinct powers of 2, we may identify the vertices involved. Since w_0 is alphabetically least, if there are three minimal vertices, say, they will be labelled $a^{2^{n-1}}b^{2^{n-1}}$, $a^{2^{n-2}}b^{2^{n-2}}$ and $a^{2^{n-3}}b^{2^{n-3}}$. Now $m_1 \leqslant n_1$, and if j occurs in m_1, then j occurs in n_1; similarly, if j occurs in m_k, j occurs in some $n_{k'}$, for some $k' \leqslant k$. Thus, we may correlate with this word a trace on the alphabet

$$a_0, \overline{a}_0, a_1, \overline{a}_1, \dots.$$

Indeed, if $n_j = 2^{j_1} + \dots 2^{j_s}$, we replace a^{n_j} by $a_{j_1}\cdots a_{j_s}$; if $m_j = 2^{j_1} + \dots 2^{j_s}$, replace b^{m_j} by $\overline{a}_{j_1}\cdots\overline{a}_{j_s}$. Since w_0 is minimal, it is clear that the resulting word is a distinguishing trace of P. Thus, the alphabetically least word $w \in Ph$ with $B(|w|)$ a maximum, and the length of $(|w|)_2$ a minimum, determines P up to isomorphism. The proof of Theorem 7.1 is complete. □

https://doi.org/10.1017/S0960129597002442 Published online by Cambridge University Press

Corollary 7.3. The function $h : \mathbf{F}_X(1) \longrightarrow \{a,b\}_X$ is injective.

We can extend this embedding to the case that the set A has more than one element. Let p_i, $i \geqslant 0$ denote the i-th prime, with $p_0 = 2$. Suppose that $A = \{a_0, a_1, \ldots, a_r\}$. Then define $h : A \longrightarrow \mathscr{P}(\{a,b\}^*)$ by

$$a_i \mapsto \{a^k b^k : k = 2^{p_i^n}, 0 \leqslant i \leqslant r, n \geqslant 1\}.$$

The same argument given for Theorem 7.1 proves the following proposition.

Proposition 7.4. For any $P, Q \in \mathbf{Pos}(A)$, if $Ph = Qh$, then P and Q are isomorphic. In fact, if P, Q have at most m vertices, and $Ph_m = Qh_m$, then P and Q are isomorphic, where

$$a_i h_m := \{a^k b^k : k = 2^{p_i^n}, 0 \leqslant i \leqslant r, 1 \leqslant n \leqslant m\}.$$

Corollary 7.5. For any set X of nonempty, traceable posets, each finitely generated free algebra $\mathbf{F}_X(n)$ in $\mathbf{Lang}(X)$ belongs to V_2.

Proof. The proof follows from Corollary 7.4. □

Corollary 7.6. The varieties V_2 and $\mathbf{Lang}(X)$ coincide.

Proof. The proof is immediate from Corollary 7.5. □

Corollary 7.7. The variety V_2 is generated by the structure $(\mathscr{P}_\omega(\{a,b\}, P_{lang})_{P \in X}$, where $\mathscr{P}_\omega(\{a,b\}$ is the collection of finite languages on the two-letter alphabet $\{a,b\}$.

Proof. This fact follows from the second statement of Proposition 7.4. □

8. Open problems

The main open problem is whether the results of the last two sections can be extended to the variety $\mathbf{Lang}(X)_\leqslant$ of ordered algebras. In detail,

- Is there a notion of 'reasonable', so that for reasonable sets X of traceable posets, each finitely generated free algebra in $\mathbf{Lang}(X)_\leqslant$ is an ordered subalgebra of the 1-generated free algebra? The map φ, defined in (7), is not order-reflecting: $(b_1^3)\varphi \leqslant (b_2 \otimes b_2)\varphi$ but b_1^3 is not related to $b_2 \otimes b_2$.
- Is the embedding h in (8) order reflecting? If not, is there any order-reflecting embedding of $\mathbf{F}_X(1)$ into $\{a,b\}_X$?

Acknowledgment

The authors want to thank Klaus Sutner for his many useful comments and suggestions.

References

Aceto, L. (1991) Full abstraction for series-parallel pomsets. In: Proceedings of TAPSOFT '91. *Springer-Verlag Lecture Notes in Computer Science* **493** 1–40.

Aceto, L. and Hennessy, M. (1993) Towards action refinement in process algebras. *Information and Computation* **103** (2) 204–269.

Bloom, S. L. (1976) Varieties of ordered algebras. *Journal of Computer and System Sciences* **45** 200–212.

Bloom, S. L. and Ésik, Z. (1995) Free Shuffle Algebras in Language Varieties: Extended Abstract. Proceedings of LATIN '95. *Springer-Verlag Lecture Notes in Computer Science* **911** 99–111.

Bloom, S. L. and Ésik, Z. (1996) Free Shuffle Algebras in Language Varieties. *Theoretical Computer Science* **163** 55–98.

Gischer, J. L. (1984) *Partial Orders and the Axiomatic Theory of Shuffle*, Ph.D. thesis, Stanford University, Computer Science Dept.

Gischer, J. L. (1988) The equational theory of pomsets. *Theoretical Computer Science* **61** 199–224.

Pratt, V. (1986) Modeling concurrency with partial orders. *International Journal of Parallel Processing* **15** (1) 33–71.

Tschantz, S. T. (1994) Languages under concatenation and shuffling. *Mathematical Structures in Computer Science* **4** 505–511.