



**HAL**  
open science

# An explicit formula for the intersection of two polynomials of regular languages

Jean-Eric Pin

► **To cite this version:**

Jean-Eric Pin. An explicit formula for the intersection of two polynomials of regular languages. Developments in Language Theory, - 17th International Conference, DLT 2013, Jun 2013, Champs sur Marne, France. pp.31-45. hal-01247099

**HAL Id: hal-01247099**

**<https://hal.science/hal-01247099>**

Submitted on 21 Dec 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# An explicit formula for the intersection of two polynomials of regular languages

Jean-Éric Pin<sup>1\*</sup>

LIAFA, University Paris-Diderot and CNRS, France.

**Abstract.** Let  $\mathcal{L}$  be a set of regular languages of  $A^*$ . An  $\mathcal{L}$ -polynomial is a finite union of products of the form  $L_0 a_1 L_1 \cdots a_n L_n$ , where each  $a_i$  is a letter of  $A$  and each  $L_i$  is a language of  $\mathcal{L}$ . We give an explicit formula for computing the intersection of two  $\mathcal{L}$ -polynomials. Contrary to Arfi's formula (1991) for the same purpose, our formula does not use complementation and only requires union, intersection and quotients. Our result also implies that if  $\mathcal{L}$  is closed under union, intersection and quotient, then its polynomial closure, its unambiguous polynomial closure and its left [right] deterministic polynomial closure are closed under the same operations.

## 1 Introduction

Let  $\mathcal{L}$  be a set of regular languages of  $A^*$ . An  $\mathcal{L}$ -polynomial is a finite union of products of the form  $L_0 a_1 L_1 \cdots a_n L_n$ , where each  $a_i$  is a letter of  $A$  and each  $L_i$  is a language of  $\mathcal{L}$ . The *polynomial closure* of  $\mathcal{L}$ , denoted by  $\text{Pol}(\mathcal{L})$ , is the set of all  $\mathcal{L}$ -polynomials.

It was proved by Arfi [1] that if  $\mathcal{L}$  is closed under Boolean operations and quotient, then  $\text{Pol}(\mathcal{L})$  is closed under intersection. This result was obtained by giving an explicit formula for computing the intersection of two polynomials of regular languages.

It follows from the main theorem of [6] that Arfi's result can be extended to the case where  $\mathcal{L}$  is only closed under union, intersection and quotient. However, this stronger statement is obtained as a consequence of a sophisticated result involving profinite equations and it is natural to ask for a more elementary proof.

The objective of this paper is to give a new explicit formula for computing the intersection of two  $\mathcal{L}$ -polynomials. Contrary to the formula given in [1], our formula only requires using union, intersection and quotients of languages of  $\mathcal{L}$ . Our proof is mainly combinatorial, but relies heavily on the notion of syntactic ordered monoid, a notion first introduced by Schützenberger [14] (see also [10]). The main difficulty lies in finding appropriate notation to state the formula, but then its proof is merely a verification.

Our result also leads to the following result, that appears to be new: if  $\mathcal{L}$  is closed under union, intersection and quotient, then its unambiguous polynomial

---

\* Work supported by the project ANR 2010 BLAN 0202 02 FREC.

closure and its left [right] deterministic polynomial closure are closed under the same operations.

Let us mention also that our algorithm can be readily extended to the setting of infinite words by using syntactic ordered  $\omega$ -semigroups [8].

## 2 Background and notation

### 2.1 Syntactic order

The *syntactic congruence* of a language  $L$  of  $A^*$  is the congruence on  $A^*$  defined by  $u \sim_L v$  if and only if, for every  $x, y \in A^*$ ,

$$xuy \in L \iff xvy \in L$$

The monoid  $M = A^*/\sim_L$  is the *syntactic monoid* of  $L$  and the natural morphism  $\eta : A^* \rightarrow M$  is called the *syntactic morphism* of  $L$ . It is a well-known fact that a language is regular if and only if its syntactic monoid is finite.

The *syntactic preorder*<sup>1</sup> of a language  $L$  is the relation  $\leq_L$  over  $A^*$  defined by  $u \leq_L v$  if and only if, for every  $x, y \in A^*$ ,  $xuy \in L$  implies  $xvy \in L$ . The associated equivalence relation is the syntactic congruence  $\sim_L$ . Further,  $\leq_L$  induces a partial order on the syntactic monoid  $M$  of  $L$ . This partial order  $\leq$  is compatible with the product and can also be defined directly on  $M$  as follows: given  $s, t \in M$ , one has  $s \leq t$  if and only if, for all  $x, y \in M$ ,  $xsy \in \eta(L)$  implies  $xt y \in \eta(L)$ . The ordered monoid  $(M, \leq)$  is called the *syntactic ordered monoid* of  $L$ .

Let us remind an elementary but useful fact: if  $v \in L$  and  $\eta(u) \leq \eta(v)$ , then  $u \in L$ . This follows immediately from the definition of the syntactic order by taking  $x = y = 1$ .

### 2.2 Quotients

Recall that if  $L$  is a language of  $A^*$  and  $x$  is a word, the *left quotient* of  $L$  by  $x$  is the language  $x^{-1}L = \{z \in A^* \mid xz \in L\}$ . The *right quotient*  $Ly^{-1}$  is defined in a symmetrical way. Right and left quotients commute, and thus  $x^{-1}Ly^{-1}$  denotes either  $x^{-1}(Ly^{-1})$  or  $(x^{-1}L)y^{-1}$ . For each word  $v$ , let us set

$$\begin{aligned} [L]_{\uparrow v} &= \{u \in A^* \mid \eta(v) \leq \eta(u)\} \\ [L]_{=v} &= \{u \in A^* \mid \eta(u) = \eta(v)\} \end{aligned}$$

---

<sup>1</sup> In earlier papers [6,10,13], I used the opposite preorder, but it seems preferable to go back to Schützenberger's original definition.

**Proposition 2.1.** *The following formulas hold:*

$$[L]_{\uparrow v} = \bigcap_{\{(x,y) \in A^* \times A^* \mid v \in x^{-1}Ly^{-1}\}} x^{-1}Ly^{-1} \quad (1)$$

$$[L]_{=v} = [L]_{\uparrow v} - \bigcup_{\eta(v) < \eta(u)} [L]_{\uparrow u} \quad (2)$$

$$[L]_{\uparrow v} = \bigcup_{\eta(v) \leq \eta(u)} [L]_{=u} \quad (3)$$

*Proof.* A word  $u$  belongs to the right hand side of (1) if and only if the condition  $v \in x^{-1}Ly^{-1}$  implies  $u \in x^{-1}Ly^{-1}$ , which is equivalent to stating that  $v \leq_L u$ , or  $\eta(v) \leq \eta(u)$ , or yet  $u \in [L]_{\uparrow v}$ . This proves (1). Formulas (2) and (3) are obvious.  $\square$

Let us make precise a few critical points. First,  $v$  always belongs to  $[L]_{\uparrow v}$ . This is the case even if  $v$  cannot be completed into a word of  $L$ , that is, if  $v$  does not belong to any quotient  $x^{-1}Ly^{-1}$ . In this case, the intersection on the right hand side of (1) is indexed by the empty set and is therefore equal to  $A^*$ .

Secondly, the intersection occurring on the right hand side of (1) and the union occurring on the right hand side of (2) are potentially infinite, but they are finite if  $L$  is a regular language, since a regular language has only finitely many quotients.

### 3 Infiltration product and infiltration maps

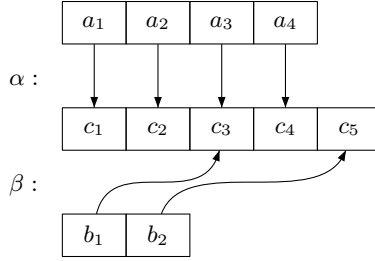
The definition below is a special case of a more general definition given in [7]. A word  $c_1 \cdots c_r$  belongs to the *infiltration product* of two words  $a_1 \cdots a_p$  and  $v = b_1 \cdots b_q$ , if there are two order preserving maps  $\alpha : \{1, \dots, p\} \rightarrow \{1, \dots, r\}$  and  $\beta : \{1, \dots, q\} \rightarrow \{1, \dots, r\}$  such that

- (1) for each  $i \in \{1, \dots, p\}$ ,  $a_i = c_{\alpha(i)}$ ,
- (2) for each  $i \in \{1, \dots, q\}$ ,  $b_i = c_{\beta(i)}$ ,
- (3) the union of the ranges of  $\alpha$  and  $\beta$  is  $\{1, \dots, r\}$ .

For instance, the set  $\{ab, aab, abb, aabb, abab\}$  is the infiltration product of  $ab$  and  $ab$  and the set  $\{aba, bab, abab, abba, baab, baba\}$  is the infiltration product of  $ab$  and  $ba$ .

A pair of maps  $(\alpha, \beta)$  satisfying Conditions (1)–(3) is called a *pair of infiltration maps*. Note that these conditions imply that  $p + q \leq r$ .

In the example pictured in Figure 1, one has  $p = 4$ ,  $q = 2$  and  $r = 5$ . The infiltration maps  $\alpha$  and  $\beta$  are given by  $\alpha(1) = 1$ ,  $\alpha(2) = 2$ ,  $\alpha(3) = 3$ ,  $\alpha(4) = 4$  and  $\beta(1) = 3$ ,  $\beta(2) = 5$ .



**Fig. 1.** A pair of infiltration maps.

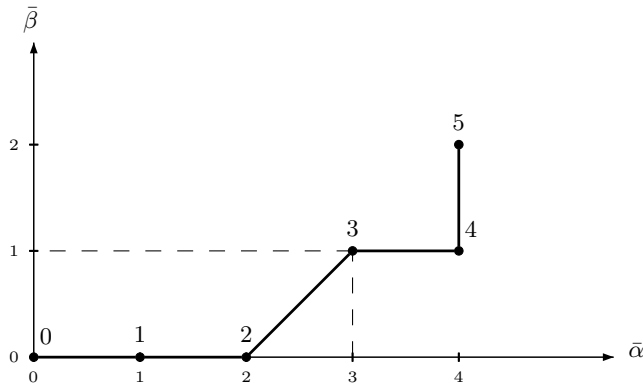
In order to state our main theorem in a precise way, we need to handle the intervals of the form  $\{\alpha(i)+1, \dots, \alpha(i+1)-1\}$ , but also the two extremal intervals  $\{1, \dots, \alpha(1)-1\}$  and  $\{\alpha(p)+1, \dots, r\}$ . As a means to get a uniform notation, it is convenient to extend  $\alpha$  and  $\beta$  to mappings  $\alpha : \{0, \dots, p+1\} \rightarrow \{0, \dots, r+1\}$  and  $\beta : \{0, \dots, q+1\} \rightarrow \{0, \dots, r+1\}$  by setting  $\alpha(0) = \beta(0) = 0$  and  $\alpha(p+1) = \beta(q+1) = r+1$ . The two extremal intervals are now of the standard form  $\{\alpha(i)+1, \dots, \alpha(i+1)-1\}$ , with  $i = 0$  and  $i = p$ , respectively. Further, we introduce the two maps  $\bar{\alpha} : \{0, \dots, r\} \rightarrow \{0, \dots, p\}$  and  $\bar{\beta} : \{0, \dots, r\} \rightarrow \{0, \dots, q\}$  defined by

$$\bar{\alpha}(i) = \max\{k \mid \alpha(k) \leq i\} \quad \text{and} \quad \bar{\beta}(i) = \max\{k \mid \beta(k) \leq i\}.$$

For instance, one gets for our example:

$$\begin{array}{cccccc} \bar{\alpha}(0) = 0 & \bar{\alpha}(1) = 1 & \bar{\alpha}(2) = 2 & \bar{\alpha}(3) = 3 & \bar{\alpha}(4) = 4 & \bar{\alpha}(5) = 4 \\ \bar{\beta}(0) = 0 & \bar{\beta}(1) = 0 & \bar{\beta}(2) = 0 & \bar{\beta}(3) = 1 & \bar{\beta}(4) = 1 & \bar{\beta}(5) = 2 \end{array}$$

These two functions are conveniently represented in Figure 2



**Fig. 2.** Graphs of  $\bar{\alpha}$  and  $\bar{\beta}$ : for instance,  $\bar{\alpha}(3) = 3$  and  $\bar{\beta}(3) = 1$ .

The next lemmas summarize the connections between  $\alpha$  and  $\bar{\alpha}$ . Of course, similar properties hold for  $\beta$  and  $\bar{\beta}$ .

**Lemma 3.1.** *The following properties hold:*

- (1)  $\bar{\alpha}(\alpha(k)) = k$ , for  $0 \leq k \leq p$ .
- (2)  $\bar{\alpha}(s+1) \leq \bar{\alpha}(s) + 1$ , for  $0 \leq s \leq r-1$ .
- (3)  $k \leq \bar{\alpha}(s)$  if and only if  $\alpha(k) \leq s$ , for  $0 \leq k \leq p$  and  $0 \leq s \leq r$ .
- (4)  $k \geq \bar{\alpha}(s)$  if and only if  $\alpha(k+1) \geq s+1$ , for  $0 \leq k \leq p-1$  and  $0 \leq s \leq r-1$ .

*Proof.* These properties follow immediately from the definition of  $\bar{\alpha}$ .  $\square$

**Lemma 3.2.** *For  $0 \leq s \leq r-1$ , the conditions  $\bar{\alpha}(s+1) = \bar{\alpha}(s) + 1$  and  $\alpha(\bar{\alpha}(s+1)) = s+1$  are equivalent.*

*Proof.* Put  $k = \bar{\alpha}(s)$  and suppose that  $\bar{\alpha}(s+1) = k+1$ . Since  $k+1 \leq \bar{\alpha}(s+1)$ , Lemma 3.1 (3) shows that  $\alpha(k+1) \leq s+1$ . Further, since  $k \geq \bar{\alpha}(s)$ , Lemma 3.1 (4) shows that  $\alpha(k+1) \geq s+1$ . Therefore  $\alpha(k+1) = s+1$  and finally  $\alpha(\bar{\alpha}(s+1)) = s+1$ .

Conversely, suppose that  $\alpha(\bar{\alpha}(s+1)) = s+1$ . Putting  $\bar{\alpha}(s+1) = k+1$ , one gets  $\alpha(k+1) = s+1$  and Lemma 3.1 (4) shows that  $k \geq \bar{\alpha}(s)$ . By Lemma 3.1 (2), one gets  $\bar{\alpha}(s+1) \leq \bar{\alpha}(s) + 1$  and hence  $k \leq \bar{\alpha}(s)$ . Thus  $\bar{\alpha}(s) = k$  and  $\bar{\alpha}(s+1) = \bar{\alpha}(s) + 1$ .  $\square$

Let us denote by  $P_\alpha(s)$  the property  $\bar{\alpha}(s+1) = \bar{\alpha}(s) + 1$ .

**Lemma 3.3.** *For  $0 \leq s \leq r-1$ , one of  $P_\alpha(s)$  or  $P_\beta(s)$  holds.*

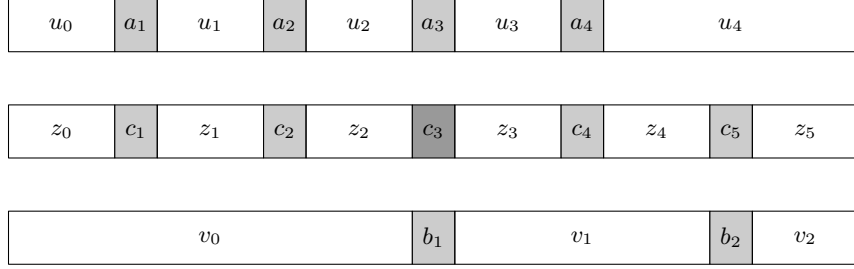
*Proof.* Since the union of the ranges of  $\alpha$  and  $\beta$  is  $\{1, \dots, r\}$ , there is an integer  $k \geq 0$  such that either  $\alpha(k+1) = s+1$  or  $\beta(k+1) = s+1$ . In the first case, one gets  $\bar{\alpha}(s+1) = \bar{\alpha}(\alpha(k+1)) = k+1$  and Lemma 3.1 (3) shows that  $\bar{\alpha}(s) \leq k$ . Since  $\bar{\alpha}(s+1) \leq \bar{\alpha}(s) + 1$  by Lemma 3.1 (2), one also has  $k \leq \bar{\alpha}(s)$  and finally  $\bar{\alpha}(s) = k$ , which proves  $P_\alpha(s)$ . In the latter case, one gets  $P_\beta(s)$  by a similar argument.  $\square$

## 4 Main result

Let  $a_1, \dots, a_p, b_1, \dots, b_q$  be letters of  $A$  and let  $K_0, \dots, K_p, L_0, \dots, L_q$  be languages of  $A^*$ . Let  $K = K_0 a_1 K_1 \cdots a_p K_p$  and  $L = L_0 b_1 L_1 \cdots b_q L_q$ .

A word of  $K \cap L$  can be factorized as  $u_0 a_1 u_1 \cdots a_p u_p$ , with  $u_0 \in K_0, \dots, u_p \in K_p$  and as  $v_0 b_1 v_1 \cdots b_q v_q$ , with  $v_0 \in L_0, \dots, v_q \in L_q$ . These two factorizations can be refined into a single factorization of the form  $z_0 c_1 z_1 \cdots c_r z_r$ , where  $c_1 \cdots c_r$  belongs to the infiltration product of  $a_1 \cdots a_p$  and  $b_1 \cdots b_q$ .

For instance, for  $p = 4$  and  $q = 2$ , one could have  $r = 5$ , with the relations  $c_1 = a_1, c_2 = a_2, c_3 = a_3 = b_1, c_4 = a_4$  and  $c_5 = b_2$ , leading to the factorization  $z_0 c_1 z_1 c_2 z_2 c_3 z_3 c_4 z_4 c_5 z_5$ , as pictured in Figure 3.



**Fig. 3.** A word of  $K \cap L$  and its factorizations.

The associated pair of infiltration maps  $(\alpha, \beta)$  is given by

$$\begin{aligned} \alpha(1) &= 1 & \alpha(2) &= 2 & \alpha(3) &= 3 & \alpha(4) &= 4 \\ \beta(1) &= 3 & \beta(2) &= 5 \end{aligned}$$

Two series of constraints will be imposed on the words  $z_i$ :

$$\begin{aligned} z_0 \in K_0, z_1 \in K_1, z_2 \in K_2, z_3 \in K_3 \text{ and } z_4 c_5 z_5 \in K_4, \\ z_0 c_1 z_1 c_2 z_2 \in L_0, z_3 c_4 z_4 \in L_1 \text{ and } z_5 \in L_2. \end{aligned}$$

We are now ready to state our main result. Let us denote by  $I(p, q)$  the set of pairs of infiltration maps  $(\alpha, \beta)$  with domain  $\{1, \dots, p\}$  and  $\{1, \dots, q\}$ , respectively. Since  $r \leq p + q$ , the set  $I(p, q)$  is finite.

**Theorem 4.1.** *Let  $K = K_0 a_1 K_1 \dots a_p K_p$  and  $L = L_0 b_1 L_1 \dots b_q L_q$  be two products of languages. Then their intersection is given by the formulas*

$$K \cap L = \bigcup_{(\alpha, \beta) \in I(p, q)} U(\alpha, \beta) \quad (4)$$

where

$$U(\alpha, \beta) = \bigcup_{(z_0, \dots, z_r) \in C(\alpha, \beta)} U_0 c_1 U_1 \dots c_r U_r \quad (5)$$

and, for  $0 \leq i \leq r$ ,

$$U_i = [K_{\bar{\alpha}(i)}]_{\uparrow z_i} \cap [L_{\bar{\beta}(i)}]_{\uparrow z_i} \quad (6)$$

and  $C(\alpha, \beta)$  is the set of  $(r + 1)$ -tuples  $(z_0, \dots, z_r)$  of words such that

- (C<sub>1</sub>) for  $0 \leq k \leq p$ ,  $z_{\alpha(k)} c_{\alpha(k)+1} z_{\alpha(k)+1} \dots c_{\alpha(k+1)-1} z_{\alpha(k+1)-1} \in K_k$ ,
- (C<sub>2</sub>) for  $0 \leq k \leq q$ ,  $z_{\beta(k)} c_{\beta(k)+1} z_{\beta(k)+1} \dots c_{\beta(k+1)-1} z_{\beta(k+1)-1} \in L_k$ .

For instance, if  $(\alpha, \beta)$  is the pair of infiltration maps of our example, one would have

$$\begin{aligned} U(\alpha, \beta) = & \bigcup_{(z_0, \dots, z_5) \in C(\alpha, \beta)} ([K_0]_{\uparrow z_0} \cap [L_0]_{\uparrow z_0}) a_1 ([K_1]_{\uparrow z_1} \cap [L_0]_{\uparrow z_1}) a_2 \\ & ([K_2]_{\uparrow z_2} \cap [L_0]_{\uparrow z_2}) b_1 ([K_3]_{\uparrow z_3} \cap [L_1]_{\uparrow z_3}) a_4 ([K_4]_{\uparrow z_4} \cap [L_1]_{\uparrow z_4}) b_2 ([K_4]_{\uparrow z_5} \cap [L_2]_{\uparrow z_5}) \end{aligned}$$

and the conditions (C<sub>1</sub>) and (C<sub>2</sub>) would be

$$(C_1) \quad z_0 \in K_0, z_1 \in K_1, z_2 \in K_2, z_3 \in K_3, z_4 c_5 z_5 \in K_4,$$

$$(C_2) \quad z_0 c_1 z_1 c_2 z_2 \in L_0, z_3 c_4 z_4 \in L_1 \text{ and } z_5 \in L_2.$$

Before proving the theorem, it is important to note that if the languages  $K_0, \dots, K_p, L_0, \dots, L_q$  are regular, the union indexed by  $C(\alpha, \beta)$  is actually a finite union. Indeed, Proposition 2.1 shows that, if  $R$  is a regular language, there are only finitely many languages of the form  $[R]_z$ .

*Proof.* Let  $U$  be the right hand side of (4). We first prove that  $K \cap L$  is a subset of  $U$ . Let  $z$  be a word of  $K \cap L$ . Then  $z$  can be factorized as  $u_0 a_1 u_1 \cdots a_p u_p$ , with  $u_0 \in K_0, \dots, u_p \in K_p$  and as  $v_0 b_1 v_1 \cdots b_q v_q$ , with  $v_0 \in L_0, \dots, v_q \in L_q$ . The common refinement of these two factorizations leads to a factorization of the form  $z_0 c_1 z_1 \cdots c_r z_r$ , where each letter  $c_k$  is either equal to some  $a_i$  or to some  $b_j$  or both. This naturally defines a pair of infiltration maps  $\alpha : \{1, \dots, p\} \rightarrow \{1, \dots, r\}$  and  $\beta : \{1, \dots, q\} \rightarrow \{1, \dots, r\}$ . Conditions (C<sub>1</sub>) and (C<sub>2</sub>) just say that the factorization  $z_0 c_1 z_1 \cdots c_r z_r$  is a refinement of the two other ones. Now, since, for  $0 \leq i \leq r$ , the word  $z_i$  belongs to  $[K_{\bar{\alpha}(i)}]_{\uparrow z_i} \cap [L_{\bar{\beta}(i)}]_{\uparrow z_i}$ , the word  $z$  belongs to  $U$ . Thus  $K \cap L \subseteq U$ .

We now prove the opposite inclusion. Let  $r \leq p + q$  be an integer, let  $\alpha : \{1, \dots, p\} \rightarrow \{1, \dots, r\}$  and  $\beta : \{1, \dots, q\} \rightarrow \{1, \dots, r\}$  be two infiltration maps and let  $(z_0, \dots, z_r) \in C(\alpha, \beta)$  and  $c_1, \dots, c_r$  satisfying (C<sub>1</sub>) and (C<sub>2</sub>). It suffices to prove that  $U_0 c_1 U_1 \cdots c_r U_r$  is a subset of  $K \cap L$ . We need a stronger version of (C<sub>1</sub>) and (C<sub>2</sub>).

**Lemma 4.2.** *The following relations hold:*

$$(C_3) \quad \text{for } 0 \leq k \leq p, U_{\alpha(k)} c_{\alpha(k)+1} U_{\alpha(k)+1} \cdots c_{\alpha(k+1)-1} U_{\alpha(k+1)-1} \subseteq K_k,$$

$$(C_4) \quad \text{for } 0 \leq k \leq q, U_{\beta(k)} c_{\beta(k)+1} U_{\beta(k)+1} \cdots c_{\beta(k+1)-1} U_{\beta(k+1)-1} \subseteq L_k.$$

Coming back once again to our main example, these conditions would be

$$(C_3) \quad U_0 \subseteq K_0, U_1 \subseteq K_1, U_2 \subseteq K_2, U_3 \subseteq K_3, U_4 c_4 U_5 \subseteq K_4,$$

$$(C_4) \quad U_0 c_1 U_2 c_2 U_2 \subseteq L_0, U_3 c_4 U_4 \subseteq L_1, U_5 \subseteq L_5.$$

*Proof.* Let  $\eta_k$  be the syntactic morphism of  $K_k$ . To simplify notation, let us set  $i = \alpha(k) + 1$  and  $j = \alpha(k + 1) - 1$ . Since  $\alpha(k) = i - 1 < i < \dots < j < \alpha(k + 1)$ , one gets  $\bar{\alpha}(i - 1) = \bar{\alpha}(i) = \dots = \bar{\alpha}(j) = k$ . Let  $u_{i-1} \in U_{i-1}, u_i \in U_i, \dots, u_j \in U_j$ . Then  $u_{i-1} \in [U_k]_{\uparrow z_{i-1}}, u_i \in [U_k]_{\uparrow z_i}, \dots, u_j \in [U_k]_{\uparrow z_j}$  and by definition,  $\eta_k(z_{i-1}) \leq \eta_k(u_{i-1}), \eta_k(z_i) \leq \eta_k(u_i), \dots, \eta_k(z_j) \leq \eta_k(u_j)$ . Therefore we get

$$\begin{aligned} \eta_k(z_{i-1} c_i z_i \cdots c_j z_j) &= \eta_k(z_{i-1}) \eta_k(c_i) \eta_k(z_i) \cdots \eta_k(c_j) \eta_k(z_j) \\ &\leq \eta_k(u_{i-1}) \eta_k(c_i) \eta_k(u_i) \cdots \eta_k(c_j) \eta_k(u_j) = \eta_k(u_{i-1} c_i u_i \cdots c_j u_j) \end{aligned}$$

Now, since  $z_{i-1} c_i z_i \cdots c_j z_j \in K_k$  by (C<sub>1</sub>), we also get  $u_{i-1} c_i u_i \cdots c_j u_j \in K_k$ , which proves (C<sub>3</sub>). The proof of (C<sub>4</sub>) is similar.  $\square$

Now, since  $\bar{\alpha}$  and  $\bar{\beta}$  are surjective, Lemma 4.2 shows that  $U_0 c_1 U_1 \cdots c_r U_r$  is a subset of  $K \cap L$ , which concludes the proof of the theorem.  $\square$



*Example 4.3.* Let  $K = b^*aA^*ba^*$  and  $L = a^*bA^*ab^*$ . The algorithm described in Theorem 4.1 gives for  $K \cap L$  the expression  $aa^*bA^*ba^*a \cup bb^*aA^*ba^*a \cup aa^*bA^*ab^*b \cup bb^*aA^*ab^*b \cup aa^*ba^*a \cup bb^*ab^*b$ .

**Corollary 4.4.** *Let  $\mathcal{L}$  be a lattice of regular languages closed under quotient. Then its polynomial closure is also a lattice closed under quotient.*

## 5 Some variants of the product

We consider in this section two variants of the product introduced by Schützenberger in [15]: unambiguous and deterministic products. These products were also studied in [2,3,4,5,9,11,12,13].

### 5.1 Unambiguous product

The marked product  $L = L_0a_1L_1 \cdots a_nL_n$  of  $n$  nonempty languages  $L_0, L_1, \dots, L_n$  of  $A^*$  is *unambiguous* if every word  $u$  of  $L$  admits a unique factorization of the form  $u_0a_1u_1 \cdots a_nu_n$  with  $u_0 \in L_0, u_1 \in L_1, \dots, u_n \in L_n$ . We require the languages  $L_i$  to be nonempty to make sure that subfactorizations remain unambiguous:

**Proposition 5.1.** *Let  $L_0a_1L_1 \cdots a_nL_n$  be an unambiguous product and let  $i_1, \dots, i_k$  be a sequence of integers satisfying  $0 < i_1 < \dots < i_k < n$ . Finally, let  $R_0 = L_0a_1L_1 \cdots a_{i_1-1}L_{i_1-1}$ ,  $R_1 = L_{i_1}a_{i_1+1}L_{i_1+1} \cdots a_{i_2-1}L_{i_2-1}$ ,  $\dots$ ,  $R_k = L_{i_k}a_{i_k+1}L_{i_k+1} \cdots a_nL_n$ . Then the product  $R_0a_{i_1}R_1 \cdots a_{i_k}R_k$  is unambiguous.*

*Proof.* Trivial.

The *unambiguous polynomial closure* of a class of languages  $\mathcal{L}$  of  $A^*$  is the set of languages that are finite unions of unambiguous products of the form  $L_0a_1L_1 \cdots a_nL_n$ , where the  $a_i$ 's are letters and the  $L_i$ 's are elements of  $\mathcal{L}$ . The term *closure* actually requires a short justification.

**Proposition 5.2.** *Any unambiguous product of unambiguous products is unambiguous.*

*Proof.* Let

$$\begin{aligned} L_0 &= L_{0,0}a_{1,0}L_{1,0} \cdots a_{k_0,0}L_{k_0,0} \\ L_1 &= L_{0,1}a_{1,1}L_{1,1} \cdots a_{k_1,1}L_{k_1,1} \\ &\vdots \\ L_n &= L_{0,n}a_{1,n}L_{1,n} \cdots a_{k_n,n}L_{k_n,n} \end{aligned} \tag{7}$$

be unambiguous products and let  $L = L_0b_1L_1 \cdots b_nL_n$  be an unambiguous product. We claim that the product

$$L_{0,0}a_{1,0}L_{1,0} \cdots a_{k_0,0}L_{k_0,0}b_1L_{0,1}a_{1,1}L_{1,1} \cdots b_nL_{0,n}a_{1,n}L_{1,n} \cdots a_{k_n,n}L_{k_n,n}$$

is unambiguous. Let  $u$  be a word of  $L$  with two factorizations

$$x_{0,0}a_{1,0}x_{1,0} \cdots a_{k_0,0}x_{k_0,0}b_1x_{0,1}a_{1,1}x_{1,1} \cdots b_nx_{0,n}a_{1,n}x_{1,n} \cdots a_{k_n,n}x_{k_n,n}$$

and

$$y_{0,0}a_{1,0}y_{1,0} \cdots a_{k_0,0}y_{k_0,0}b_1y_{0,1}a_{1,1}y_{1,1} \cdots b_ny_{0,n}a_{1,n}y_{1,n} \cdots a_{k_n,n}y_{k_n,n}$$

with  $x_{0,0}, y_{0,0} \in L_{0,0}, \dots, x_{k_n,n}, y_{k_n,n} \in L_{k_n,n}$ . Setting

$$\begin{aligned} x_0 &= x_{0,0}a_{1,0}x_{1,0} \cdots a_{k_0,0}x_{k_0,0} & y_0 &= y_{0,0}a_{1,0}y_{1,0} \cdots a_{k_0,0}y_{k_0,0} \\ x_1 &= x_{0,1}a_{1,1}x_{1,1} \cdots a_{k_1,1}x_{k_1,1} & y_1 &= y_{0,1}a_{1,1}y_{1,1} \cdots a_{k_1,1}y_{k_1,1} \\ &\vdots & &\vdots \\ x_n &= x_{0,n}a_{1,n}x_{1,n} \cdots a_{k_n,n}x_{k_n,n} & y_n &= y_{0,n}a_{1,n}y_{1,n} \cdots a_{k_n,n}y_{k_n,n} \end{aligned} \quad (8)$$

we get two factorizations of  $u$ :  $x_0b_1x_1 \cdots b_nx_n$  and  $y_0b_1y_1 \cdots b_ny_n$ . Since the product  $L_0b_1L_1 \cdots a_nL_n$  is unambiguous, we have  $x_0 = y_0, \dots, x_n = y_n$ . Each of these words has now two factorizations given by (8) and since the products of (7) are unambiguous, these factorizations are equal. This proves the claim and the proposition.  $\square$

We now consider the intersection of two unambiguous products.

**Theorem 5.3.** *If the products  $K = K_0a_1K_1 \cdots a_pK_p$  and  $L = L_0b_1L_1 \cdots b_qL_q$  are unambiguous, the products occurring in Formula (4) are all unambiguous.*

*Proof.* Let  $(\alpha, \beta)$  be a pair of infiltration maps, and let  $U_i = [K_{\bar{\alpha}(i)}]_{\uparrow z_i} \cap [L_{\bar{\beta}(i)}]_{\uparrow z_i}$ , for  $0 \leq i \leq r$ . We claim that the product  $U = U_0c_1U_1 \cdots c_rU_r$  is unambiguous. Let

$$u = u_0c_1u_1 \cdots c_ru_r = u'_0c_1u'_1 \cdots c_ru'_r \quad (9)$$

be two factorizations of a word  $u$  of  $U$  such that, for  $0 \leq i \leq r$ ,  $u_i, u'_i \in U_i$ . We prove by induction on  $s$  that  $u_s = u'_s$ .

*Case  $s = 0$ .* By the properties of  $\alpha$  and  $\beta$ , we may assume without loss of generality that  $\alpha(1) = 1$ , which implies that  $c_1 = a_1$ . It follows from (C<sub>3</sub>) that  $U_0 \subseteq K_0$ . Now the product  $K_0a_1(K_1a_2K_2 \cdots a_pK_p)$  is unambiguous by Proposition 5.1, and by (C<sub>3</sub>),  $U_1c_2U_2 \cdots c_ru_r$  is contained in  $K_1a_1K_2 \cdots a_pK_p$ . Therefore,  $u$  admits the two factorizations  $u_0a_1(u_1c_2u_2 \cdots c_ru_r)$  and  $u'_0a_1(u'_1c_2u'_2 \cdots c_ru'_r)$  in this product. Thus  $u_0 = u'_0$ .

*Induction step.* Let  $s > 0$  and suppose by induction that  $u_i = u'_i$  for  $0 \leq i \leq s-1$ . If  $s = r$ , then necessarily  $u_s = u'_s$ . If  $s < r$ , we may assume without loss of generality that  $s$  is in the range of  $\alpha$ . Thus  $\alpha(k) = s$  for some  $k$  and  $c_s = a_k$ . We now consider two cases separately.

If  $\alpha(k+1) = s+1$  (and  $c_{s+1} = a_{k+1}$ ), it follows from (C<sub>3</sub>) that  $u$  has two factorizations

$$\begin{aligned} (u_0c_1u_1 \cdots c_{s-1}u_{s-1})a_ku_s a_{k+1}(u_{s+1}c_{s+1}u_{s+2} \cdots c_ru_r) \text{ and} \\ (u_0c_1u_1 \cdots c_{s-1}u_{s-1})a_ku'_s a_{k+1}(u'_{s+1}c_{s+1}u'_{s+2} \cdots c_ru'_r) \end{aligned}$$

over the product  $(K_0 a_1 K_1 \cdots a_{s-1} K_{s-1}) a_k K_s a_{k+1} (K_{s+1} a_{k+2} K_{s+2} \cdots a_p K_p)$ . Since this product is unambiguous by Proposition 5.1, we get  $u_s = u'_s$ .

If  $\alpha(k+1) \neq s+1$ , then  $s+1 = \beta(t+1)$  for some  $t$  and  $c_{s+1} = b_{t+1}$ . Setting  $i = \beta(t)$ , we get  $c_i = b_t$  and it follows from (C<sub>4</sub>) that  $u$  has two factorizations

$$(u_0 c_1 u_1 \cdots c_{i-1} u_{i-1}) b_t (u_i c_{i+1} u_{i+1} \cdots c_s u_s) b_{t+1} (u_{s+1} c_{s+2} u_{s+2} \cdots c_r u_r) \text{ and} \\ (u_0 c_1 u_1 \cdots c_{i-1} u_{i-1}) b_t (u'_i c_{i+1} u'_{i+1} \cdots c_s u'_s) b_{t+1} (u'_{s+1} c_{s+2} u'_{s+2} \cdots c_r u'_r)$$

over the product  $(L_0 b_1 L_1 \cdots b_{t-1} L_{t-1}) b_t L_t b_{t+1} (L_{t+1} b_{t+1} L_{t+2} \cdots b_p L_p)$ . This product is unambiguous by Proposition 5.1, and thus

$$u_i c_{i+1} u_{i+1} \cdots c_s u_s = u'_i c_{i+1} u'_{i+1} \cdots c_s u'_s$$

Now the induction hypothesis gives  $u_i = u'_i, \dots, u_{s-1} = u'_{s-1}$  and one finally gets  $u_s = u'_s$ .  $\square$

We state separately another interesting property.

**Theorem 5.4.** *Let  $K = K_0 a_1 K_1 \cdots a_p K_p$  and  $L = L_0 b_1 L_1 \cdots b_q L_q$  be two unambiguous products and let  $(\alpha, \beta)$  and  $(\alpha', \beta')$  be two pairs of infiltration maps of  $I(p, q)$ . If the sets  $U(\alpha, \beta)$  and  $U(\alpha', \beta')$  meet, then  $\alpha = \alpha'$  and  $\beta = \beta'$ .*

*Proof.* Suppose that a word  $u$  belongs to  $U(\alpha, \beta)$  and to  $U(\alpha', \beta')$ . Then  $u$  has two decompositions of the form

$$u = u_0 c_1 u_1 \cdots c_r u_r = u'_0 c'_1 u'_1 \cdots c'_r u'_r$$

Condition (C<sub>1</sub>) [(C<sub>2</sub>)] and the unambiguity of the product  $K_0 a_1 K_1 \cdots a_p K_p$  [ $L_0 b_1 L_1 \cdots b_q L_q$ ] show that, for  $0 \leq i \leq p$  and for  $0 \leq j \leq q$ ,

$$u_{\alpha(i)} c_{\alpha(i)+1} u_{\alpha(i)+1} \cdots c_{\alpha(i+1)-1} u_{\alpha(i+1)-1} = \\ u'_{\alpha'(i)} c'_{\alpha'(i)+1} u'_{\alpha'(i)+1} \cdots c'_{\alpha'(i+1)-1} u'_{\alpha'(i+1)-1} \in K_i \quad (10)$$

$$u_{\beta(j)} c_{\beta(j)+1} u_{\beta(j)+1} \cdots c_{\beta(j+1)-1} u_{\beta(j+1)-1} = \\ u'_{\beta'(j)} c'_{\beta'(j)+1} u'_{\beta'(j)+1} \cdots c'_{\beta'(j+1)-1} u'_{\beta'(j+1)-1} \in L_j \quad (11)$$

We prove by induction on  $s$  that, for  $1 \leq s \leq \min(r, r')$ , the following properties hold:

$$E_1(s) : u_{s-1} = u'_{s-1} \text{ and } c_s = c'_s,$$

$$E_2(s) : \bar{\alpha}(s) = \bar{\alpha}'(s) \text{ and } \bar{\beta}(s) = \bar{\beta}'(s),$$

$$E_3(s) : \text{for } i \leq \bar{\alpha}(s), \alpha(i) = \alpha'(i) \text{ and for } j \leq \bar{\beta}(s), \beta(j) = \beta'(j).$$

*Case  $s = 1$ .* We know that either  $\alpha(1) = 1$  or  $\beta(1) = 1$  and that either  $\alpha'(1) = 1$  or  $\beta'(1) = 1$ . Suppose that  $\alpha(1) = 1$ . We claim that  $\alpha'(1) = 1$ . Otherwise, one has  $\beta'(1) = 1$ . Now, Formula (10) applied to  $i = 0$  gives

$$u_0 = u'_0 c'_1 u'_1 \cdots c'_{\alpha'(1)-1} u'_{\alpha'(1)-1}$$

and Formula (11) applied to  $j = 0$  gives

$$u_0 c_1 u_1 \cdots c_{\beta(1)-1} u_{\beta(1)-1} = u'_0.$$

Therefore  $u_0 = u'_0$  and  $\alpha'(1) = 1$ , which proves the claim. It follows also that  $a_1 = c_{\alpha(1)} = c_{\alpha'(1)}$  and thus  $c_1 = c'_1$ . We also have in this case  $\bar{\alpha}(1) = \bar{\alpha}'(1) = 1$ . A similar argument shows that if  $\alpha'(1) = 1$ , then  $\alpha(1) = 1$ . Therefore, the conditions  $\alpha(1) = 1$  and  $\alpha'(1) = 1$  are equivalent and it follows that  $\bar{\alpha}(1) = \bar{\alpha}'(1)$ . A dual argument would prove that the conditions  $\beta(1) = 1$  and  $\beta'(1) = 1$  are equivalent and that  $\bar{\beta}(1) = \bar{\beta}'(1)$ .

*Induction step.* Let  $s$  be such that  $1 \leq s+1 \leq \min(r, r')$  and suppose by induction that the properties  $E_1(i)$ ,  $E_2(i)$ ,  $E_3(i)$  hold for  $1 \leq i \leq s$ .

**Lemma 5.5.** *Suppose that  $P_\alpha(s)$  holds and let  $k = \bar{\alpha}(s)$ . Then*

$$s \leq \alpha'(k+1) - 1 \tag{12}$$

and

$$u_s = u'_s c'_{s+1} u'_{s+1} \cdots c_{\alpha'(k+1)-1} u'_{\alpha'(k+1)-1} \tag{13}$$

*Proof.* Applying (10) with  $i = k$ , we get

$$u_{\alpha(k)} c_{\alpha(k)+1} u_{\alpha(k)+1} \cdots c_s u_s = u'_{\alpha'(k)} c'_{\alpha'(k)+1} u'_{\alpha'(k)+1} \cdots c_{\alpha'(k+1)-1} u'_{\alpha'(k+1)-1} \tag{14}$$

Since  $\bar{\alpha}(s) = \bar{\alpha}'(s)$  by  $E_2(s)$ , one has  $\bar{\alpha}'(s) = k$  and  $\alpha'(k+1) \geq s+1$  by Lemma 3.1, which gives (12). Further, since  $k = \bar{\alpha}(s)$ , it follows from  $E_3(s)$  that  $\alpha(k) = \alpha'(k)$ . Now, for  $i \leq s$ ,  $E_1(i)$  implies that  $u_{i-1} = u'_{i-1}$  and  $c_i = c'_i$ . It follows that the word  $u_{\alpha(k)} c_{\alpha(k)+1} u_{\alpha(k)+1} \cdots c_s$  is a prefix of both sides of (14). Therefore, this prefix can be deleted from both sides of (14), which gives (13).  $\square$

We now establish  $E_1(s+1)$ .

**Lemma 5.6.** *One has  $u_s = u'_s$  and  $c_{s+1} = c'_{s+1}$ . Further,  $P_\alpha(s)$  and  $P_{\alpha'}(s)$  are equivalent and  $P_\beta(s)$  and  $P_{\beta'}(s)$  are equivalent.*

*Proof.* Let us prove that  $u'_s$  is a prefix of  $u_s$ . By Lemma 3.3, either  $P_\alpha(s)$  or  $P_\beta(s)$  holds. Suppose that  $P_\alpha(s)$  holds. Then by Lemma 5.5,  $u'_s$  is a prefix of  $u_s$ . If  $P_\beta(s)$  holds, we arrive to the same conclusion by using (11) in place of (10) in the proof of Lemma 5.5.

Now, a symmetrical argument using the pair  $(\bar{\alpha}', \bar{\beta}')$  would show that  $u_s$  is a prefix of  $u'_s$ . Therefore,  $u_s = u'_s$ . Coming back to (13), we obtain  $\alpha'(k+1) = s+1$  and since by  $E_2(s)$ ,  $k = \bar{\alpha}(s) = \bar{\alpha}'(s)$ , one gets  $\alpha'(\bar{\alpha}'(s) + 1) = s+1$ , which, by Lemma 3.2, is equivalent to  $P_{\alpha'}(s)$ . Thus  $P_\alpha(s)$  implies  $P_{\alpha'}(s)$  and a dual argument would prove the opposite implication.

We also have  $c_{s+1} = c_{\alpha(k+1)} = a_{k+1} = c'_{\alpha'(k+1)} = c'_{s+1}$  and thus  $c_{s+1} = c'_{s+1}$ . Finally, a similar argument works for  $\beta$ .  $\square$

We now come to the proof of  $E_2(s+1)$  and  $E_3(s+1)$ . Since  $P_\alpha(s)$  and  $P_{\alpha'}(s)$  are equivalent, the next two lemma cover all cases.

**Lemma 5.7.** *If neither  $P_\alpha(s)$  nor  $P_{\alpha'}(s)$  hold, then  $\bar{\alpha}(s+1) = \bar{\alpha}'(s+1)$  and for  $i \leq \bar{\alpha}(s+1)$ ,  $\alpha(i) = \alpha'(i)$ . Similarly, if neither  $P_\beta(s)$  nor  $P_{\beta'}(s)$  hold, then  $\bar{\beta}(s+1) = \bar{\beta}'(s+1)$  and for  $i \leq \bar{\beta}(s+1)$ ,  $\beta(i) = \beta'(i)$ .*

*Proof.* We just prove the “ $\alpha$  part” of the lemma. If neither  $P_\alpha(s)$  nor  $P_{\alpha'}(s)$  hold, then  $\bar{\alpha}(s+1) = \bar{\alpha}(s)$  and  $\bar{\alpha}'(s+1) = \bar{\alpha}'(s)$ . Since  $\bar{\alpha}(s) = \bar{\alpha}'(s)$  by  $E_2(s)$ , one gets  $\bar{\alpha}(s+1) = \bar{\alpha}'(s+1)$ . The second property is an immediate consequence of  $E_3(s)$ .  $\square$

**Lemma 5.8.** *If both  $P_\alpha(s)$  and  $P_{\alpha'}(s)$  hold, then  $\bar{\alpha}(s+1) = \bar{\alpha}'(s+1)$  and for  $i \leq \bar{\alpha}(s+1)$ ,  $\alpha(i) = \alpha'(i)$ . Similarly, if both  $P_\beta(s)$  and  $P_{\beta'}(s)$  hold, then  $\bar{\beta}(s+1) = \bar{\beta}'(s+1)$  and for  $i \leq \bar{\beta}(s+1)$ ,  $\beta(i) = \beta'(i)$ .*

*Proof.* Again, we just prove the “ $\alpha$  part” of the lemma. If both  $P_\alpha(s)$  and  $P_{\alpha'}(s)$  hold, then  $\bar{\alpha}(s+1) = \bar{\alpha}(s) + 1$  and  $\bar{\alpha}'(s+1) = \bar{\alpha}'(s) + 1$ . Since  $\bar{\alpha}(s) = \bar{\alpha}'(s)$  by  $E_2(s)$ , one gets  $\bar{\alpha}(s+1) = \bar{\alpha}'(s+1)$ . Property  $E_3(s)$  shows that for  $i \leq \bar{\alpha}(s)$ ,  $\alpha(i) = \alpha'(i)$ . Since  $\bar{\alpha}(s+1) = \bar{\alpha}(s) + 1$ , it just remains to prove that

$$\alpha(\bar{\alpha}(s+1)) = \alpha'(\bar{\alpha}'(s+1)) \quad (15)$$

But Lemma 3.2 shows that  $\alpha(\bar{\alpha}(s+1)) = s+1$  and  $\alpha'(\bar{\alpha}'(s+1)) = s+1$ , which proves (15) since  $\bar{\alpha}(s+1) = \bar{\alpha}'(s+1)$ .  $\square$

This concludes the induction step and the proof of Theorem 5.4.  $\square$

**Corollary 5.9.** *Let  $\mathcal{L}$  be a lattice of regular languages closed under quotient. Then its unambiguous polynomial closure is also a lattice closed under quotient.*

If  $\mathcal{L}$  is a Boolean algebra, then one can be more precise.

**Corollary 5.10.** *Let  $\mathcal{L}$  be a Boolean algebra of regular languages closed under quotient. Then its unambiguous polynomial closure is also a Boolean algebra closed under quotient.*

Let us conclude with an example which shows that, under the assumptions of Theorem 5.4, the sets  $U(\alpha, \beta)$  cannot be further decomposed as a disjoint union of unambiguous products.

Let  $K = K_0aK_1$  and  $L = L_0aL_1$  with  $K_0 = L_1 = 1 + b + c + c^2$  and  $L_0 = K_1 = a + ab + ba + ac + ca + ac^2 + bab + cac + cac^2$ . Then

$$\begin{aligned} K \cap L = & aa + aab + aba + aac + aca + aac^2 + abab + acac + acac^2 + \\ & baa + baab + baba + baac + baac^2 + babab + caa + \\ & caab + caac + caca + caac^2 + cacac + cacac^2 \end{aligned}$$

One can write for instance  $K \cap L$  as  $(1 + b + c)aa(1 + b + c + c^2) + (1 + b)a(1 + b)a(1 + b) + (1 + c)a(1 + c)a(1 + c + c^2)$  but the three components of this language

are not disjoint, since they all contain  $aa$ . Note that the words  $acab, abac, baca$  and  $caba$  are not in  $K \cap L$ .

The syntactic ordered monoid of  $K_0$  and  $L_1$  has 4 elements  $\{1, a, b, c\}$  and is presented by the relations  $a = ba = b^2 = bc = ca = cb = 0$  and  $c^2 = b$ . Its syntactic order is defined by  $a < b < c < 1$ .

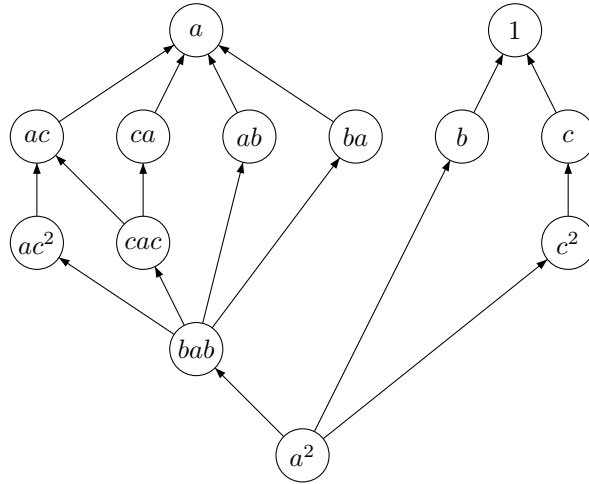
The syntactic ordered monoid of  $L_0$  and  $K_1$  has 13 elements:

$$\{1, a, b, c, a^2, ab, ac, ba, ca, c^2, ac^2, bab, cac\}$$

and is defined by the relations  $cac^2 = bab$  and

$$b^2 = bc = cb = a^2 = aba = aca = bac = cab = c^2a = c^3 = 0.$$

The syntactic order is:



There is only one pair of infiltration maps  $(\alpha, \beta)$  of  $I(1, 1)$  that defines a nonempty set  $U(\alpha, \beta)$ . This pair is defined as follows:  $\alpha(1) = 1$  and  $\beta(1) = 2$ . The triples  $(z_0, z_1, z_2)$  of  $C(\alpha, \beta)$  are exactly the triples of words such that  $z_0az_1az_2 \in K \cap L$ . In particular,  $z_0 \in \{1, b, c\}$ ,  $z_1 \in \{1, b, c\}$  and  $z_2 \in \{1, b, c, c^2\}$ . Now, one has

$$\begin{array}{lll} [K_0]_{\uparrow 1} = 1 & [K_0]_{\uparrow b} = 1 + b + c + c^2 & [K_0]_{\uparrow c} = 1 + c \\ [K_1]_{\uparrow 1} = 1 & [K_1]_{\uparrow b} = 1 + b & [K_1]_{\uparrow c} = 1 + c \quad [K_1]_{\uparrow c^2} = 1 + c + c^2 \\ [L_0]_{\uparrow 1} = 1 & [L_0]_{\uparrow b} = 1 + b & [L_0]_{\uparrow c} = 1 + c \\ [L_1]_{\uparrow 1} = 1 & [L_1]_{\uparrow b} = 1 + b + c + c^2 & [L_1]_{\uparrow c} = 1 + c \quad [L_1]_{\uparrow c^2} = 1 + b + c + c^2 \end{array}$$

which gives the following possibilities for the triples  $(U_0, U_1, U_2)$ , for the following triples  $z = (z_0, z_1, z_2)$ :

$z = (1, 1, 1)$	$U_0 = 1$	$U_1 = 1$	$U_2 = 1$
$z = (b, b, b)$	$U_0 = 1 + b$	$U_1 = 1 + b$	$U_2 = 1 + b$
$z = (c, c, c)$	$U_0 = 1 + c$	$U_1 = 1 + c$	$U_2 = 1 + c$
$z = (b, c, c^2)$	$U_0 = 1 + b$	$U_1 = 1 + c$	$U_2 = 1 + c + c^2$
$z = (c, c, c^2)$	$U_0 = 1 + c$	$U_1 = 1 + c$	$U_2 = 1 + c + c^2$

## 5.2 Deterministic product

The marked product  $L = L_0 a_1 L_1 \cdots a_n L_n$  of  $n$  nonempty languages  $L_0, L_1, \dots, L_n$  of  $A^*$  is *left deterministic* [*right deterministic*] if, for  $1 \leq i \leq n$ , the set  $L_0 a_1 L_1 \cdots L_{i-1} a_i [a_i L_i \cdots a_n L_n]$  is a prefix [suffix] code. This means that every word of  $L$  has a unique prefix [suffix] in  $L_0 a_1 L_1 \cdots L_{i-1} a_i [a_i L_i \cdots a_n L_n]$ . It is observed in [3, p. 495] that the marked product  $L_0 a_1 L_1 \cdots a_n L_n$  is deterministic if and only if, for  $1 \leq i \leq n$ , the language  $L_{i-1} a_i$  is a prefix code. Since the product of two prefix codes is a prefix code, we get the following proposition.

**Proposition 5.11.** *Any left [right] deterministic product of left [right] deterministic products is left [right] deterministic.*

*Proof.* This follows immediately from the fact that the product of two prefix codes is a prefix code.  $\square$

Factorizing a deterministic product also gives a deterministic product. More precisely, one has the following result.

**Proposition 5.12.** *Let  $L_0 a_1 L_1 \cdots a_n L_n$  be a left [right] deterministic product and let  $i_1, \dots, i_k$  be a sequence of integers satisfying  $0 < i_1 < \dots < i_k < n$ . Finally, let  $R_0 = L_0 a_1 L_1 \cdots a_{i_1-1} L_{i_1-1}, \dots, R_k = L_{i_k} a_{i_k+1} L_{i_k+1} \cdots L_{n-1} a_n L_n$ . Then the product  $R_0 a_{i_1} R_1 \cdots a_{i_k} R_k$  is left [right] deterministic.*

*Proof.* Trivial.  $\square$

The *left [right] deterministic polynomial closure* of a class of languages  $\mathcal{L}$  of  $A^*$  is the set of languages that are finite unions of left [right] deterministic products of the form  $L_0 a_1 L_1 \cdots a_n L_n$ , where the  $a_i$ 's are letters and the  $L_i$ 's are elements of  $\mathcal{L}$ .

We can now state the counterpart of Theorem 5.3 for deterministic products.

**Theorem 5.13.** *If the products  $K = K_0 a_1 K_1 \cdots a_p K_p$  and  $L = L_0 b_1 L_1 \cdots b_q L_q$  are deterministic, the products occurring in Formula (4) are all deterministic.*

*Proof.* Let  $i \in \{0, \dots, r\}$ . By construction, there exists  $k \geq 0$  such that  $i + 1 = \alpha(k + 1)$  or  $i + 1 = \beta(k + 1)$ . By Lemma 4.2, there exists  $j \leq i$  such that either  $U_j c_{j+1} U_{j+1} \cdots U_i \subseteq K_k$  and  $c_{\alpha(k+1)} = a_{k+1}$  or  $U_j c_{j+1} U_{j+1} \cdots U_i \subseteq L_k$  and

$c_{\alpha(k+1)} = b_{k+1}$ . Suppose we are in the first case and that  $U_i c_{i+1}$  is not a prefix code. Then  $U_j c_{j+1} U_{j+1} \cdots U_i c_{i+1}$  is not a prefix code and thus  $K_k a_{k+1}$  is not a prefix code. This yields a contradiction since the product  $K_0 a_1 K_1 \cdots a_p K_p$  is deterministic.  $\square$

**Corollary 5.14.** *Let  $\mathcal{L}$  be a lattice of regular languages closed under quotient. Then its deterministic polynomial closure is also closed under quotient.*

## Acknowledgements

I would like to thank Mário J. J. Branco for his careful reading of a first version of this article.

## References

1. M. ARFI, Opérations polynomiales et hiérarchies de concaténation, *Theoret. Comput. Sci.* **91** (1991), 71–84.
2. M. J. J. BRANCO, On the Pin-Thérien expansion of idempotent monoids, *Semigroup Forum* **49**,3 (1994), 329–334.
3. M. J. J. BRANCO, The kernel category and variants of the concatenation product, *Internat. J. Algebra Comput.* **7**,4 (1997), 487–509.
4. M. J. J. BRANCO, Two algebraic approaches to variants of the concatenation product, *Theoret. Comput. Sci.* **369**,1-3 (2006), 406–426.
5. M. J. J. BRANCO, Deterministic concatenation product of languages recognized by finite idempotent monoids, *Semigroup Forum* **74**,3 (2007), 379–409.
6. M. J. J. BRANCO AND J.-E. PIN, Equations for the polynomial closure, in *ICALP 2009, Part II*, S. Albers and W. Thomas (eds.), pp. 115–126, *Lect. Notes Comp. Sci.* vol. 5556, Springer, Berlin, 2009.
7. M. LOTHAIRE, *Combinatorics on words*, *Cambridge Mathematical Library*, Cambridge University Press, Cambridge, 1997.
8. D. PERRIN AND J.-E. PIN, *Infinite Words*, *Pure and Applied Mathematics* vol. 141, Elsevier, 2004. ISBN 0-12-532111-2.
9. J.-E. PIN, Propriétés syntactiques du produit non ambigu, in *7th ICALP*, pp. 483–499, *Lect. Notes Comp. Sci.* n° 85, Springer, Berlin, 1980.
10. J.-E. PIN, A variety theorem without complementation, *Russian Mathematics (Iz. VUZ)* **39** (1995), 80–90.
11. J.-E. PIN, H. STRAUBING AND D. THÉRIEN, Locally trivial categories and unambiguous concatenation, *J. of Pure and Applied Algebra* **52** (1988), 297–311.
12. J.-E. PIN AND D. THÉRIEN, The bideterministic concatenation product, *Internat. J. Algebra Comput.* **3** (1993), 535–555.
13. J.-E. PIN AND P. WEIL, Polynomial closure and unambiguous product, *Theory Comput. Systems* **30** (1997), 1–39.
14. M.-P. SCHÜTZENBERGER, Une théorie algébrique du codage, in *Séminaire Dubreil-Pisot, année 1955-56, Exposé No. 15, 27 février 1956, 24 pages*, Inst. H. Poincaré, Paris, 1956.
15. M.-P. SCHÜTZENBERGER, Sur le produit de concaténation non ambigu, *Semigroup Forum* **18** (1976), 331–340.