FILOMAT (Niš) **10** (1996), 73–86

# ABSOLUTELY ELEMENTARY BIORDERED LANGUAGES

**Yu Bingjun**

**Abstract.** A language is called (elementary) biordered, shortened as (EBL) BL, if it is recognized by (a single idempotent) some idempotents of a monoid. In this paper, the sufficient and necessary conditions for a language to be (EBL) BL are given. Further, the concept of absolutely elementary biordered language (AEBL) is defined. The precise structure and the syntactic monoids of AEBLs are described. Some important relations among them are investigated. Finally, a structure of any BL is given via AEBLs.

## 1. Introduction and Preliminaries

As is well known, an important task of the theory of formal languages is to characterize the structure of languages by their syntactic monoids, or more generally, by any monoid recognizing them. On the other hand, it is well known also that idempotents in a monoid play a central role in deciding the properties and structure of the monoid. The theory of biordered sets initiated by Nambooripad is a typical model in this aspect. This paper is a part of an attempt to combine these two to investigate the properties and structure of languages via those decided by idempotents.

This section presents the basic terminology and notations concerning languages and monoids we will need.

Let $A$ be a nonempty set (finite or infinite) called an *alphabet* whose elements are called *letters*. Define a *word* over $A$ as a nonempty finite sequence, written $a_1a_2\dots a_n$, of elements of $A$. The *length* of $w$ denoted by $l(w)$ is defined as the number of occurrences of letters of $A$ in $w$. It is well-known that the set $A^+$ of all words over $A$ with the concatenation of words forms a semigroup called the *free semigroup over A*. The *free monoid over A* denoted by $A^*$ is obtained from $A^+$ by adjoining an identity, the *empty word* 1 whose lenght is 0.

Received August 19, 1995

1991 *Mathematics Subject Classification*: 20M05, 68Q45.

The author would like to thank Prof. Nambooripad for his many helps.

Any subset $L \subseteq A^+$ is called a *language* over $A$. Given a language $L$ over $A$, the *syntactic congruence* $P_L$ of $L$ is defined as follows:

$$(x,y) \in P_L \Leftrightarrow (uxv \in L \Leftrightarrow uyv \in L,\ \forall u,v \in A^*),\ \forall x,y \in A^+.$$

One verifies that $L$ is a union of classes of $P_L$. In fact $P_L$ is the largest congruence on $A^*$ for which $L$ is union of classes. The quotient monoid $A^+/P_L$ is the *syntatic monoid* of $L$ and is denoted by $M(L)$. The canonical homomorphism from $A^*$ onto $M(L)$ is the *syntactic homomorphism* of $L$ and is denoted by $\phi_L$. For any language $L \subseteq A^+$, if there exist a monoid $M$ and a surjective homomorphism $\theta : A^* \longrightarrow M$ such that $L$ is union of $\ker\theta$-classes we say that $\theta$, or $\ker\theta$, or $M$ recognizes $L$, where $\ker\theta$ is the congruence on $A^*$ induced by $\theta$. If this is the case, there exists a unique homomorphism $\sigma : M \longrightarrow M(L)$ such that $\phi_L = \theta\sigma$. This means, the syntactic monoid $M(L)$ is the smallest monoid recognizing $L$.

Given two languages $L, K \subseteq A^+$, the product of them is defined by $LK = \{uv : u \in L, v \in K\}$. The product operation of languages is associative. The star language generated by a language $L$ is defined as $L^* = \bigcup_{n=0}^{\infty} L^n$, where $L^0 = \{1\}, L^2 = LL$ and $L^n = L^{n-1}L, \forall n \geq 2$. Moreover, $L^+ = L^* - \{1\}$.

## 2. Definition and Criterions

The following definition is the main definition of this paper:

**Definition 1.** A language $L \subseteq A^+$ is called a *biordered language* over $A$, shortened as BL, if there exist a monoid $M$ and a surjective morphism $\theta : A^+ \longrightarrow M$ such that $L = P\theta^{-1}$, for some $P \subseteq E(M)$, where $E(M)$ is the set of all idempotents of $M$. If this is the case we say that $M(\ker\theta, \theta)$ recognizes $L$ by its idempotents. Further, $L$ is called an *elementary biordered language*, shortened as EBL, if $P$ is a singleton. Especially $L$ is called *identity biordered language*, shortened as IBL, if $L = 1\theta^{-1}$, where 1 is the identity of the monoid $M$. The set of all BLs (res. EBLs, IBLs) over $A$ is denoted by $L(A)$ (res. $EL(A), IL(A)$).

Obviously, any BL $L$ over $A$ can be written as a union of some disjoint EBLs: $L = \bigcup\{L_i : i \in I\}$ such that the monoid recognizing $L$ by $P \subseteq E(M)$ recognizes each $L_i, i \in I$, by a single idempotent in $P$. This decomposition of $L$ is called an *effective decomposition* of $L$.

The following criterion is a consequence of Theorem 10.6 in [2].

**Theorem 1.** *A language $L \subseteq A^*$ is a BL over $A$ if and only if the following conditions are satisfied:*

(i) $L = \bigcup\{L_i : i \in I\},\quad L_i \cap L_j = 0,\ \forall i,j \in I,\ i \neq j;$

(ii) *$L_i$ is a subsemigroup of $A^*$, for each $i \in I$;*

(iii) *$uL_iv \cap L_j \neq 0$ implies $uL_iv \subseteq L_j, \forall i,j \in I$ and $u,v \in A^*$.*

*In particular, $L \in EL(A)$ if and only if $L$ is a subsemigroup of $A^*$ and that $uLv \cap L \neq 0$ implies $uLv \subseteq L, \forall u, v \in A^*$.*

From this theorem, any two-sided ideal of $A^*$ is an EBL over $A$. In fact, any EBL over $A$ can be viewed as a generalized two-sided ideal of $A^*$. Moreover, using Theorem 1 one verifies that any left (res. right) principal ideal of $A^*$ is also an EBL over $A$ (but not every left (res. right) ideal of $A^*$ is) and that the following proposition holds.

**Proposition 2.** *The set of all languages in $EL(A)$ which are two-sided (res. left, right) ideals of $A^*$ is an order ideal under the inverse inclusion.*

The next proposition is also immediate from Theorem 1:

**Proposition 3.** *For any $L \subseteq L(A)$ $(EL(A))$, if $w \in L$ is of the shortest length in $L$, then $L' = L - \{w\} \in L(A)$ $(EL(A))$.* □

The converse of this proposition is not true. For example, taking $A = \{a, b\}$ and $L = (ab)^+$, one verifies that $L \in EL(A)$ but $(ab)^* = L \cup \{1\} \notin EL(A)$.

Now we proceed to establish another criterion theorem via congruences on $A^*$. For this purpose we need some new notations.

By $C$ we denote the lattice of all congruence on $A^*$, where $A$ is an alphabet. For convenience we will abuse any $\rho \in C$ and the morphism induced by it. So given any language $L \subseteq A^*$, we have

$$L_{\rho\rho^{-1}} = \{w \in A^* : w\phi \in L_\phi\},$$

where $\phi : A^* \longrightarrow A^*/\rho$ is the canonical morphism induced by $\rho$.

Clearly, $C^L$ is the set of all congruences on $A^*$ recognizing $L$.

Moreover, denoting $\delta(L) = \{(w, w^2) : w \in L\}$ and $\delta'(L) = L \times L$, we define

$$C_L = \{\rho \in C : \delta(L) \subseteq \rho\} \text{ and } C' = \{\rho \in C : \delta'(L) \subseteq \rho\}.$$

We also denote by $\delta_L$ and $\delta'_L$ the congruences generated by $\delta(L)$ and $\delta(L)$ respectively.

**Lemma 4.** *For any $\rho, \sigma \in C$, we have*

(i) *$\rho \in C_L$ $(C'_L)$ and $\rho \subseteq \sigma$ together imply $\sigma \in C$ $(C'_L)$;*

(ii) *$\rho \in C^L$ and $\sigma \subseteq \rho$ together imply $\sigma \in C^L$;*

(iii) *$C_L$ $(C'_L)$ is a complete sublattice of $C$ with the 1-element $A^* \times A^*$ and the 0-element $\delta_L$ $(\delta'_L)$;*

(iv) *$C^L$ is a complete sublattice of $C$ with the 1-element $P_L$, the syntactic congruence of $L$, and the 0-element $1_{A^*}$.*

*Proof.* We prove only (ii) and (iv). The proof for (i) and (iii) is similar but much more easier.

Let $\rho \in \mathcal{C}^L$ and $\sigma \subseteq \rho$. So $L_{\rho\rho^{-1}} = L$ by definition. For any $w \in L_{\sigma\sigma^{-1}}$, there is $u \in L$ such that $(w,u) \in \sigma \subseteq \rho$. Therefore we have $w \in u\rho\rho^{-1} \subseteq L_{\rho\rho^{-1}} = L$. This proves $L_{\sigma\sigma^{-1}} \subseteq L$ which implies $L_{\sigma\sigma^{-1}} = L$, since $L \subseteq L_{\sigma\sigma^{-1}}$ is trivial.

For (iv), it suffices to prove that $\rho = \bigvee\{\rho_i : i \in I\} \in \mathcal{C}^L$ for any family $\{\rho_i \in \mathcal{C}^L : i \in I\}$. Indeed, if $w \in L_{\rho\rho^{-1}}$, there exists $u \in L$ such that $(w,u) \in \rho = (\bigcup\{\rho_i : i \in I\})^\infty$. This means that there exist an integer $n \geq 1$, $i_j \in I$ and $x_j \in A^*, j = 1, \ldots, n$, satisfying

$$w\rho_{i_1}x_1\rho_{i_2}x_2\cdots\rho_{i_n}x_n = u.$$

Now we have $u \in L$ and $L\rho_{i_j}\rho_{i_j}{}^{-1} = L$, since $\rho_{i_j} \in \mathcal{C}^L$, for each $j = 1, \ldots, n$. A simple induction deduces $w \in L$ which implies $L_{\rho\rho^{-1}} = L$. This completes the proof. □

Now we define

$$\mathcal{C}(L) = \mathcal{C}_L \cap \mathcal{C}^L \text{ and } \mathcal{C}'(L) = \mathcal{C}'_L \cap \mathcal{C}^L.$$

We have the following:

**Theorem 5.** *For any language $L \subseteq A^*$, we have*

(i) *$L \in L(A) \Leftrightarrow \mathcal{C}(L) \neq 0$;*
(ii) *$L \in EL(A) \Leftrightarrow L$ is a subsemigroup of $A^*$ and $\mathcal{C}'(L) \neq 0$;*
(iii) *If $\mathcal{C}(L) \neq 0$, it is a complete sublattice of $\mathcal{C}$ with the 1-element $P_L$ and the 0-element $\delta_L$. More precisely, $\mathcal{C}(L) = \{\rho \in \mathcal{C} : \delta_L \subseteq \rho \subseteq P_L\}$;*
(iv) *If $\mathcal{C}'(L) \neq 0$, it is a complete sublattice of $\mathcal{C}$ with the 1-element $P_L$ and the 0-element $\delta'_L$. More precisely, $\mathcal{C}'(L) = \{\rho \in \mathcal{C} : \delta'_L \subseteq \rho \subseteq P_L\}$.*

*Proof.* The statement (i) and (ii) are immediate by definition. We prove (iii) for $\mathcal{C}(L)$. The proof of (iv) for $\mathcal{C}'(L)$ is similar.

If $\mathcal{C}(L) \neq 0$, it is clearly a complete sublattice of $\mathcal{C}$, since the intersection of complete sublattices is also complete. Now let $\rho \in \mathcal{C}(L)$, then $\delta_L \subseteq \rho \subseteq P_L$ by definition. Because $\delta(L) \subseteq \rho$ and $L_{\rho\rho^{-1}} = L$, we have $\delta(L) \subseteq P_L$ and $L_{\delta_L\delta_L{}^{-1}} = L$ by Lemma 4 which implies that both of $P_L$ and $\delta_L$ are in $\mathcal{C}(L)$. The last equality is merely a simple consequence of Lemma 4. □

Now we see that a language $L \subseteq A^*$ is a BL (an EBL) if and only if there exists a family of monoids recognizing $L$ by their idempotents (a single idempotent). In this family there is the smallest one, $M(L)$, the syntactic monoid of $L$ which is a homomorphic image of any other member. Meanwhile, there is also the largest one, $A^*/\delta_L$, $(A^*/\delta'_L)$, of which every member in this family is a homomorphic image. In particular, we have the following two immediate corollaries.

**Corollary 6.** *For any language $L \subseteq A^*$, the following statements are equivalent:*

(i) *$L$ is BL;*
(ii) *$\delta(L) \subseteq P_L$, i.e. $M(L)$ recognizes $L$ by its idempotents;*
(iii) *$L_{\delta_L \delta_L^{-1}} = L$.*

**Corollary 7.** *For any language $L \subseteq A^*$, the following statements are equivalent:*

(i) *$L$ is an EBL;*
(ii) *$L$ is a subsemigroup of $A^*$ and $\delta'(L) \subseteq P_L$, i.e. $M(L)$ recognizes $L$ by its single idempotent.*
(iii) *$L$ is a subsemigroup of $A^*$ and $L_{\delta'_L \delta'_L{}^{-1}} = L$.*

The next corollary is also immediate:

**Corollary 8.** *$L \subseteq A^*$ is an IBL if and only if $L = 1{P_L}^{-1}$, where 1 is the identity of $M(L)$, that is, $L$ is the IBL recognized by its syntactic monoid.*

## 3. Boolean Operations

In this section, we investigate the Boolean operations of BLs over $A$. First, we have the following:

**Proposition 9.** *For any alphabet $A$, the intersection of any family of BLs (res. EBLs, IBLs) over $A$ is also a BL (res. an EBL, an IBL) over $A$.*

*Proof.* Let $\{L_i \subseteq A^* : i \in I\}$ be a family of BLs and $L = \bigcap\{L_i : i \in I\}$. Denote $\rho = \bigcap\{P_{L_i} : i \in I\}$, where $P_{L_i}$ is the syntactic congruence of $L_i$, $i \in I$. It suffices to prove that $L = L_{\rho\rho^{-1}}$ and $\delta(L) \subseteq \rho$.

For any $w \in L_{\rho\rho^{-1}}$, there exists $u \in L$ such that $(w, u) \in \rho$. Since $L \subseteq L_i$, $\rho \subseteq P_{L_i}$ and $L_{i\rho\rho^{-1}} = L_i$ by Lemma 4, we have $w \in u\rho\rho^{-1} \subseteq L_{i\rho\rho^{-1}} = L_i$ for each $i \in I$. This shows $L = L_{\rho\rho^{-1}}$.

Moreover, we have $\delta(L_i) \subseteq P_{L_i}$, since $L_i$ are BLs by Theorem 5, $i \in I$. Meanwhile we have also

$$\delta(L) = \{(w, w^2) : w \in L = \bigcap\{L_i : i \in I\}\}$$
$$= \bigcap\{\{(w, w^2) : w \in L_i\} : i \in I\} = \bigcap\{\delta(L_i) : i \in I\}.$$

Since $\delta(L_i) \subseteq \delta_{L_i} \subseteq P_{L_i}$, for each $i \in I$, we have

$$\delta(L) \subseteq \bigcap\{P_{L_i} : i \in I\} = \rho.$$

This proves that $L$ is a BL.

When $L_i$ are EBLs, $i \in I$, we have $L_i \times L_i \subseteq P_{L_i}$ for all $i \in I$. Therefore

$$L \times L \subseteq \bigcap\{L_i \times L_i : i \in I\} \subseteq \bigcap\{P_{L_i} : i \in I\} = \rho$$

which proves that $L$ is an EBL, since that $L$ is a subsemigroup of $A^*$ is obvious. The assertion for IBL is merely a consequence of that for EBL. □

From this proposition we can see that, for any $L \subseteq A^*$, there exists a unique BL (res. EBL, IBL), which is the smallest one containing $L$, i.e. the intersection of all BLs (res. EBLs, IBLs) containing $L$, called the *generated BL (res. EBL, IBL) by L*. Therefore each of $L(A), EL(A)$ and $IL(A)$ forms a complete lattice under inclusion of languages and $IL(A)$ is a complete sublattice of $EL(A)$, where, for a family of BLs (res. EBLs, IBLs) $\{L_i : i \in I\}$, the join $\bigvee\{L_i : i \in I\}$ is defined to be the BL (res. EBL, IBL) generated by $L = \bigcup\{L_i : i \in I\}$. Note that $EL(A)$ is not a sublattice of $L(A)$ if $A$ contains at least two letters $a, b$, since $(ab)^* = \{1\} \vee (ab)^+$ holds thue in $L(A)$ but not in $EL(A)$. One verifies that, for any $w \in A^+$, $w^+$ is the BL as well as EBL generated by $w$ and that $w^*$ is the IBL generated by $w$ if and only if $w = a^n$, for some $a \in A$ and $n \geq 1$.

As for the union operation, the situation is much more complicated. Firstly, for any family $\{L_i : i \in I\}$ of BLs over $A$, if $\bigcap\{\mathcal{C}(L_i) : i \in I\} \neq 0$, the union $\bigcup\{L_i : i \in I\}$ is clearly a BL over $A$. Howerer, when this intersection is empty, all possibilities appear. For example, $L_1 = (a^2)^+$ and $L_2 = (a^3)^+$ are both BLs over $A = \{a\}$, but $L_1 \cup L_2$ is not. On the other hand, if $I_1$ and $I_2$ are different two-sided ideal of $A^+ = a^+$, then $I = I_1 \cup I_2$ is an EBL over $A$, since $I$ is also a two-sided ideal of $A^*$. But $\mathcal{C}(I_1) \cap \mathcal{C}(I_2) = 0$, because if there were a monoid $M$ recognizing both of them by its idempotents, then they should be the inverse homomorphic image of the zero of $M$ so that $I_1 = I_2$, contradiction. Anyhow we have the following:

**Proposition 10.** *Let $L_1, L_2 \subseteq A^*$ by BLs with $L_1 \cap L_2 = 0$. Then $L_1 \cup L_2$ is a BL if and only if $\mathcal{C}(L_1) \cap \mathcal{C}(L_2) \neq 0$.*

*Proof.* It suffices to prove necessity. Denote $\delta_i = \delta_{L_i}$, $i = 1, 2$, and $\delta = \delta_{L_1 \cup L_2}$.

First we prove $\delta = \delta_1 \vee \delta_2$. By definition $\delta(L_i) \subseteq \delta$. So $\delta_i \subseteq \delta$, since $\delta_i$ is the smallest congruence containing $\delta(L_i)$, $i = 1, 2$. Hence $\delta_1 \vee \delta_2 \subseteq \delta$. Conversely, we have $\delta(L_1 \cup L_2) = \delta(L_1) \cup \delta(L_2)$. Therefore $\delta(L_1 \cap L_2) \subseteq \delta_1 \cup \delta_2 \subseteq \delta_1 \vee \delta_2$ which implies $\delta \subseteq \delta_1 \vee \delta_2$ by the definition of $\delta$.

Next we have $L_1 \cup L_2 = (L_1 \cup L_2)\delta\delta^{-1}$ from the assumption of the proposition and $\delta = \delta_{L_1 \cup L_2}$. Now we proceed to prove $\delta \subseteq P_{L_i}$, $i = 1, 2$, i.e., $L_i = L_i\delta\delta^{-1}$ for each $i = 1, 2$, which will imply $\delta \in \mathcal{C}(L_1) \cap \mathcal{C}(L_2)$.

Indeed, for any $w \in L_1\delta\delta^{-1}$, there is $u \in L_i$ such that $(w,u) \in \delta = \delta_1 \vee \delta_2 = (\delta_1 \circ \delta_2)^\infty$. This means that there exist $n \geq 1$ and $x_1, \ldots, x_{2n-1} \in A^*$ satisfying

$$w\delta_1 x_1 \delta_2 x_2 \delta_1 x_3 \cdots x_{2n-2}\delta_1 x_{2n-1}\delta_2 u.$$

From $\delta_2 \subseteq \delta_1 \vee \delta_2 = \delta$ we have

$$x_{2n-1} \in \dot{u}\delta\delta^{-1} \subseteq (L_1 \cup L_2)\delta\delta^{-1} = L_1 \cup L_2.$$

If $x_{2n-1} \in L_2$, we would have

$$u \in x_{2n-1}\delta_2{\delta_2}^{-1} \subseteq L_2\delta_2{\delta_2}^{-1} = L_2,$$

which contradicts the condition $L_1 \cap L_2 = 0$. Hence $x_{2n-1} \in L_1$ so that $x_{2n-2} \in x_{2n-1}\delta_1{\delta_1}^{-1} \subseteq L_1$. By a simple induction we deduce $w \in L_1$, that is, $L_1\delta\delta^{-1} = L_1$. The equality $L_2\delta\delta^{-1} = L_2$ can be proved similary. □

We note that the Proposition 10 can not be extended to any family of Bls as showed later in an example after Proposition 19.

## 4. AEBLs and the Absolute Decomposition of BLs

For any $L \in L(A)$, $\delta_L \subseteq \delta'_L$ always holds. The equality holds if and only if each monoid recognizing $L$ as a BL recognizes it by a single idempotent. In other words, $L$ is a BL if and only if it is an EBL. We call a BL with such a property *absolutely elementary biordered language*, shortened as AEBL. In this section, we discuss the structure and some properties of AEBLs. We also discuss a special role they play in constructing any BLs. For this we need the following two concepts.

A word $p \in A^+$ is called *primitive*, if $p = u^n$ implies $n = 1$ for any $u \in A^*$. The set of all primitive words in $A^*$ is denoted by $Q(A)$. The following results are well-known (cf. [7]).

(1) *For any word $w \in A^+$, there exist unique $p \in Q(A)$ and $n \geq 1$ such that $w = p^n$. This $p$ is called the primitive root of $w$.*

(2) *For any $u, v \in A^+$, $uv = vu$ implies that $u$ and $v$ have the same primitive root.*

Moreover, two words $w, w' \in A^*$ are said to be conjugate, if $w = uv$ and $w' = vu$ for some $u, v \in A^*$. The conjugacy is an equivalence relation on $A^*$. Further, $w$ and $w'$ are conjugate if and only if their primitive roots are conjugate and $l(w) = l(w')$. A word $w$ is primitive if and only if its conjugate is primitive and a primitive word of length $l$ has exactly $l$ conjugates.

One checks that $\{1\}$ and any language of the form $w^t w^*$, $w \in A^+$, $t \geq 1$ are AEBLs over $A$. On the other hand, any IBL $L$ containing a nonempty word is not that.

The following three results characterize the structure of all AEBLs but $\{1\}$.

**Lemma 11.** *Let $L \in L(A)$ and $1 \notin L$. If all the words in $L$ are powers of the some word, then $L$ is an EBL of the form $p^{tn}(p^n)^*$, for some $p \in Q(A)$ and $t, n \geq 1$.*

*Proof.* Notice that the condition of the lemma is equivalent to that all the words in $L$ have the same primitive root $p \in Q(A)$. We first prove that $L$ is actually an EBL. In fact, $L$ can be written as a disjoint union of some EBLs: $L = \bigcup\{L_i : i \in I\}$ by definition. For any two words $p^{m_j} \in L$, $j = 1, 2$, there are EBLs $L_{i_j}$ such that $p^{m_j} \in L_{i_j}$ which implies that $(p^{m_j})^+ \subseteq L_{i_j}$, since $L_{i_j}$ are subsemigroups of $A^*$, $j = 1, 2$. If $i_1 \neq i_2$, we should have $(p^{m_1})^+ \cap (P^{m_2})^+ = 0$ which contradicts the fact $p^{m_1 m_2} \in (p^{m_1})^+ \cap (P^{m_2})^+$.

Now let $p^{m_0}$ and $p^{m_0+n}$ be the shortest words in $L$ and $L - \{p^{m_0}\}$, respectively. We have immediately $n > 0$ and $p^{m_0}(p^n)^* \subseteq L$, since $L$ is an EBL. If there were $p^m \in L - p^{m_0}(p^n)^*$, there would be $k, n_1$ with $k \geq 0$, $0 < n_1 < n$ such that $m = m_0 + kn + n_1$, then we should have $p^{n_1} L \cap L \neq 0$ which implies $p^{n_1} L \subseteq L$ so that $p^{m_0+n_1} \in L - \{p^{m_0}\}$ which contradicts the definition of $n$. This proves $L = p^{m_0}(p^n)^*$. Further we have $p^{2m_0} \in L$ which implies $m_0 = tn$, for some $t \geq 1$. □

**Proposition 12.** *Let $L$ be an AEBL over $A$ and $p$ the primitive root of a word of shortest length in $L$. Then all the words in $L$ are powers of $p$.*

*Proof.* Let $p^{m_0}$ be a word of shortest length in $L$ and $w$ an arbitrary word in $L$. Because $L$ is an AEBL we have $p^{m_0} \delta_L w$, that is, there exist $n \geq 1$, $z_0, z_i, u_i, v_i \in A^*$ and $x_i \in L$, $i = 1, \ldots, n$ such that

$$(*) \qquad \begin{aligned} &p^{m_0} = z_0, \qquad w = z_n, \text{ and} \\ &z_{i-1} = u_i {x_i}^2 v_i, \quad z_i = u_i x_i v_i, \text{ or} \\ &z_{i-1} = u_i x_i v_i, \quad z_i = u_i {x_i}^2 v_i, \\ &\qquad i = 1, \ldots, n. \end{aligned}$$

We prove by induction on $i$ that $z_i$ is a power of $p$. The result is trivially true for $i = 0$. Assuming $z_{i-1} = p^m$, we prove that $z_i$ is also a power of $p$ by induction on $l(z_i)$ as follows:

Let $l(z_i)$ be shortest length of the words in $L$. Since $x_i \in L$, we certainly have $u_i v_i = 1$, $z_i = x_i$ and $z_{i-1} = p^m = {x_i}^2$ by the equalities of (*). This implies that $z_i$ is a power of $p$. Assume that any word in $L$ is a power of $p$ whenever it is of a length strictly less than $l(z_i)$. We investigate the equalities (*) again. If $u_i v_i = 1$, using the same argument as above, we can get that $z_i$ is a power of $p$. So let $u_i v_i \neq 1$. Then $l(x_i) < l(z_i)$ which implies $x_i = p^s$ for some $s \geq 1$ by the induction assumption. Hence $l(u_i v_i)$ is divided by $l(p)$ from (*). Without less of generality we assume $0 \leq l(u_i)$, $l(v_i) \leq l(p)$. We claim that $0 < l(u_i) < l(p)$ is impossible. In fact, if this were this case, we should have $0 < l(v_i) < l(p)$ and $l(u_iv_i) = l(p)$ which would lead to $u_iv_i = p$. But we have already $p^m = u_ip^sv_i$. So we should have $u_iv_i = p = v_iu_i$ which contradicts the fact that $p$ is a primitive word. Therefore $u_i = 1$ or $p$. Similarly $v_i = 1$ or $p$. Thus $z_i$ has $p$ its primitive root, i.e., $z_i$ is a power of $p$

By induction on $i$ we obtain that $w$ is a power of $p$. □

**Theorem 13.** *For any $L \subseteq A^+$, the following are equivalent:*

(i) *$L$ is an AEBL;*
(ii) *$L \in L(A)$ and all the words in $L$ are powers of the same word;*
(iii) *$L = p^{tn}(p^n)^*$, for some $p \in Q(A)$ and $t, n \geq 1$;*
(iv) *$L \in L(A)$ and $\delta_L{}' \subseteq \delta_L$.*

*Proof.* (i) ⇒ (ii) by Proposition 12. (ii) ⇒ (iii) by Lemma 11, (iii) ⇒ (iv) was checked by the reader, and (iv) ⇒ (i) by definition. □

Now we discuss some relations among AEBLs. For this purpose we fix an integer $n \geq 1$ and $p \in Q(A)$. Let $l \geq 1$ be the length of $p$. Denote $l$ conjugates of $p$ by $p_1 = p, p_2, \ldots, p_l$. Further, for any $i \geq 1$, we denote $L_{i,j} = p_j{}^{in}(p_j{}^n)^*$, $P_{ij} = P_{L_{ij}}$, the syntactic congruence of $L_{ij}$, and $\delta_{ij} = \delta_{L_{ij}}$.

**Proposition 14.** *With the notations above, we have:*

(i) *$P_{1,j} = P_{1,j'}$, for any $j, j' = 1, \ldots, l$;*
(ii) *For any $i > 1$ and $j, j' = 1, \ldots, l$, $P_{i,j} \subseteq P_{i-1,j'}$ and the inclusion is strict;*
(iii) *For any $i > 1$ and $j, j' = 1, \ldots, l$ with $j \neq j'$, $\delta_{i-1,j'} \subseteq P_{i,j}$.*

*Proof.* Without loss of generality, we prove $P_{1,j} = P_{1,1}$, $\forall j = 1, \ldots, l$. By the symmetry of conjugacy, it suffices to prove $P_{1,1} \subseteq P_{1,j}$, $\forall j = 1, \ldots, l$. Now we have $L_{1,1} = (p^n)^+$ and $L_{1,j} = (p_j{}^n)^+$, $p = uv$, $p_j = vu$ for some $u, v \in A^*$. Let $(w, w') \in P_{1,1}$, we prove $(w, w') \in P_{1,j}$ as follows:

First one checks that $\{1\}$ is a single $P_{1,1}$-class, so we assume $w, w' \in A^+$. For any $x, y \in A^*$ with $xwy \in (p_j^n)^+$, we have

$$(uv)^{n-1}uxwyv \in (uv)^{n-1}u((vu)^n)^+v \subseteq ((uv)^n)^+ = (p^n)^+.$$

Thus we have $(uv)^{n-1}uxw'yv = (p^n)^s = (uv)^{ns}$ for some $s > 1$, because $(w, w') \in P_{1,1}$ and $w' \neq 1$. Hence $xw'y = (vu)^{(s-1)n} \in (p_j^n)^+$. Similarly, $xw'y \in (p_j^n)^+$ implies $xwy \in (p_j^n)^+$. Therefore $(w, w') \in P_{1,j}$, $\forall j = 1, \ldots, l$.

Now let $i > 1$. For (ii) it suffices to prove $P_{i,1} \subseteq P_{i-1,j}$, for all $j = 1, \ldots, l$. Suppose $(w, w') \in P_{i,1}$ and $w, w' \in A^+$. For any $x, y \in A^*$, if

$xwy \in p_j^{(i-1)n}(p_j^n)^*$, we have

$$(uv)^{n-1}uxwyv \in (uv)^{n-1}u(vu)^{(i-1)n}((vu)^n)^+v = p^{in}(p^n)^+,$$

thus $(uv)^{n-1}uxw'yv = p^{(i+s)n} = (uv)^{(i+s)n}$ for some $s \geq 0$ from $(w,w') \in P_{i,1}$. This implies $xw'y \in p_j^{(i-1)n}(p_j^n)^*$. Similarly, $xw'y \in p_j^{(i-1)n}(p_j^n)^*$ implies $xwy \in p_j^{(i-1)n}(p_j^n)^*$. This proves $P_{i,1} \subseteq P_{i-1,j}$, for all $j = 1, \ldots ; l$.

For (iii) it suffices to prove

$$\delta_{i-1,j} \subseteq P_{i,1}, \ \forall j \neq 1,$$

that is, $(p_j^{kn}, p_j^{k'n}) \in P_{i,1}$, for any $k, k' \geq i-1$. Indeed, if $xp_j^{kn}y = p^{(i+s)n}$ for some $x, y \in A^*$ and $s \geq 0$, then $l(xy)$ is divided by $nl(p)$. Without loss of generality, we assume $0 \leq l(x), l(y) \leq nl(p)$. Since $p \neq p_j$, we have $0 < l(x), l(y) < nl(p)$ which together with the equation above implies $xy = p^n = (uv)^n$. Thus we have

$$x = (uv)^\alpha x', \quad y = y'(uv)^\beta$$

for some $\alpha, \beta \geq 0$ and $x', y' \in A^*$ with

$$0 < l(x'), \ l(y') < l(p) \text{ and } x'y' = uv.$$

Now it is easily seen that $n = \alpha + \beta + 1$ and $x'(vu)^{kn}y' = (uv)(uv)^{(i+s-1)n}$. Using the equidivisibility of $A^*$ and the primitivity of $p$ we have $x' = u$ and $y' = v$ so that

$$\begin{aligned} xp_j^{k'n}y &= (uv)^\alpha u(vu)^{k'n}v(uv)^\beta \\ &= (uv)^{k'n+\alpha+\beta+1} \\ &= (uv)^{(k'+1)n}, \end{aligned}$$

which implies $xp_j^{k'n}y \in p^{in}(p^n)^*$, because we have $k'+1 \geq i$. Similarly, $xp_j^{k'n}y \in p^{in}(p^n)^*$ implies $xp_j^{kn}y \in p^{in}(p^n)^*$ for any $x, y \in A^*$. This proves $\delta_{i-1,j} \subseteq P_{i,1}$.

For any $j = 1, \ldots, l$, $P_{i,j} \neq P_{i-1,j}$ is clear, because $L_{i,j}$ and $L_{i-1,j}$ are a single classes of the two congruences respectively, but they are different from each other and have a nonempty intersection. Let $j \neq j'$. For $i = 2$ we have $P_{1,j} = P_{1,j'}$, so $P_{2,j} \subseteq P_{1,j'}$ is strict. Now let $i > 2$. We have $P_{i,j} \subseteq P_{i-1,j'} \subseteq P_{i-2,j}$ and $\delta_{i-2,j} \subseteq P_{i-1,j'}$ by the preceding argument. If $P_{i,j} = P_{i-1,j'}$, then we should have $\delta_{i-2,j} \subseteq P_{i,j} \subseteq P_{i-2,j}$. This is not possible. □

From this proposition we see that $P_{i,j} \in \mathcal{C}(L_{i-1,j'})$, i.e. the monoid $M(L_{i,j})$ recognizes $L_{i-1,j'}$, for all $i \geq 2$, $j, j' = 1, \ldots, l$, $j \neq j'$. Further, the following corollary shows that for any $t \geq 1$, the languages $L_{t,j}$, $j = 1, \ldots, l$ are recognized by the same monoid too.

**Corollary 15.** *With the notions above, we have*

$$\bigcap_{j=1}^{l} P_{t,j} \in \bigcap_{j=1}^{l} \mathcal{C}(L_{t,j}),$$

*for any* $l, t \geq 1$.

*Proof.* The result is trivially true for $l = 1$ or $t = 1$. Now assume $l, t > 1$. From Proposition 14 $P_{i,j} \subseteq P_{i-1,j'}$ for any $i > 1$ and $j, j' = 1, \ldots, l$ which implies

$$P_{i,j} \subseteq \bigcap_{j'=1}^{l} P_{i-1,j'} = p \subseteq P_{i-1,j},$$

for all $j = 1, \ldots, l$. Further, for any $j' = 1, \ldots, l$, there is $j = 1, \ldots, l$ with $j \neq j'$, since $l > 1$. So by Proposition 14 again we have

$$\delta_{i-1,j'} \subseteq P_{i,j} \subseteq p \subseteq P_{i-1,j'}.$$

Taking $i = t + 1$, the result follows. □

The next three results characterize the syntatctic monoids of AEBLs.

**Proposition 16.** *Let $K$ be an EBL recognized by the syntactic monoid $M(L)$ of the AEBL $L = p^{tn}(p^n)^*$ by its nonidentity and nonzero idempotent, then there is a conjugate $q$ of $p$ such that*

$$K = \begin{cases} (q^n)^*, & \text{if } t = 1 \\ q^{(t-1)n}(q^n)^*, & \text{if } t > 1. \end{cases}$$

*Proof.* First any $w \in K$ is of a length divided by $nl(p)$, because $K$ is recognized by $M(L)$ by its nonzero idempotent, there are $x, y \in A^*$ such that $xwy,\ xw^2y \in L$. Let $q$ be the primitive root of $w$, we prove, $w$ is of the form $q^{mn}$, $m \geq 1$. In fact, there are $x, y \in A^*$ and $s \geq 0$ such that $xwy = p^{(t+s)n}$. This implies that $l(xy)$ is divided by $nl(p)$. Without loss generality, assume $0 \leq l(x),\ l(y) \leq nl(p)$. If anyone of the equalities in them holds, we have $q = p$ and $w = p^{mn}$. Now suppose $0 < l(x),\ l(y) < nl(p)$. Then we have $l(xy) = nl(p)$ and $xwy = p^{(t+s)n}$. This implies $xy = p^n$. Thus $x = p^\alpha u,\ y = vp^\beta$ for some $\alpha, \beta \geq 0$ and $u, v \in A^*$ with $0 < l(u), l(v) < l(p)$. If any equality in them holds, we have again $q = p$ and $w = p^{mn}$. So let $u, v \in A^+$. Then we have $uv = p$ and $\alpha + \beta = 1 = n$ by $xy = p^n$. Form $xwy = p^{(t+s)n}$ we have immediately $w = (vu)^{(t+s-1)n}$. Since $uv = p$ is primitive, $vu$ is also primitive so that $q = vu$ and $w = q^{mn}$. Since $M(L)$ recognizes $K$ by its nonidentity idempotent, $w \neq 1$. So $m \geq 1$.

If $t = 1$, we have $P_{(p^n)^+} = P_{(q^n)^+}$ by the Proposition 14 which means $(q^n)^+$ is recognized by $M(L)$ by its single idempotent. But now $K \bigcap (q^n)^+ \neq 0$, this ensures $K = (q^n)^+$. If $t > 1$, we have $P_{t,1} \in \mathcal{C}(L_{t-1,j})$ for all $j = 1, \ldots, l$, this implies that $q^{(t-1)n}(q^n)^*$ is an EBL recognized by $M(L)$ because $q = p_j$, for some $j = 1, \ldots, l$. But we have also $K \bigcap q^{(t-1)n}(q^n)^* \neq 0$, thus $K = q^{(t-1)n}(q^n)^*$ holds. □

**Corollary 17.** *For any $L = p^{tn}(p^n)^*$, $p \in Q(A)$, $t, n \geq 1$, the idempotents of $M(L)$ which are neither identity nor zero are exactly the images of the AEBLs of the form $(q^n)^+$ (whent $= 1$) or $q^{(t-1)n}(q^n)^*$ (whent $> 1$) under the syntactic morhism $\phi_L$, where $q$ runs over all the conjugates of $p$.*

**Proposition 18.** *Any AEBL $L$ is a rational language and $E(M(L))$ is a semilattice.*

*Proof.* Let $L = p^{tn}(p^n)^*$, $p \in Q(A)$, $t, n \geq 1$. $L$ is rational, because $L = (p^n)^+ - \{p^{in} : 1 \leq i < t\}$ and the later two are both rational.

To prove $E(M(L))$ is a semilattice we discuss two casses: (i) $l(p) = 1$. One easily checks that $E(M(L))$ is a chain of two or three elements, (ii) $l(p) > 1$. We first prove that, for any $q, q' \in A^+$ with $l(q) = l(q') = l(p)$ and $q \neq q'$, we have $A^*qq'A^* \bigcap L = 0$. In fact, if $xqq'y = p^m$ for some $x, y \in A^*$ and $m \geq 1$, $l(xy)$ is divided by $l(p)$. Without loss of generality, assume $0 \leq l(x), l(y) \leq l(p)$. If any one of the equalities holds, we have $x, y = 1$ or $p$ which together with $l(q) = l(q') = l(p)$ and $xqq'y = p^m$ implies $q = p = q'$, a contradiction. Thus $0 < l(x), l(y) < l(p)$, and we have $xy = p$ by $l(xy) = l(p)$. From this we have $qq' = (yx)^{m-1}$ which implies $q = yx = q'$, also a contradiction. Now $p$ has at least two different conjugates and the idempotents of $M(L)$ which are neither identity nor zero are exactly the images of AEBLs consisting of the powers of these conjugates by Corollary 17. The product of any two such idempotents must be zero of $M(L)$ from the above, that is, $E(M(L))$ is a semilattice. □

The following proposition shows that AEBLs play a special role in constructing any BLs.

**Proposition 19.** *For any $L \in L(A)$, $L$ can be uniquely decomposed as a disjoint union of AEBLs over $A$.*

*Proof.* Assume $1 \notin L$. For any $p \in Q(A)$, if $L \bigcap p^+ \neq 0$, it is a BL by Proposition 9. Further it is an AEBL of the form $p^{tn}(p^n)^*$ by Theorem 13. Denote $Q_L = \{p \in Q(A) : L \bigcap p^+ \neq 0\} = \{p_i : i \in I\}$. Obviously we have $L = \bigcup \{p^{t_i n_i}(p^{n_i})^* : p_i \in Q_L\}$, where, $p_i^{t_i n_i}$ is the shortest word in $L \bigcap p_i^+$.

That this decomposition is unique and disjoint is clear. The result for the case that $1 \in L$ is a natural cosequence of the above argument. □

The decomposition of any BL $L \in L(A)$ in the above proposition is called the *absolute decomposition*, shortened as the AD, of $L$. Unfortunately, the AD of any BL is not necessarily effective, i.e., the AEBL components of this decomposition are not necessarily recognized by the same monoid. For example, taking $A = \{a, b\}$, the AD of $A^+ \in L(A)$ is $A^+ = \bigcup\{p^+ : p \in Q(A)\}$ which has infinite AEBL components of the form $pp^*$, $p \in Q(A)$. However, one easily verifies that $A^*/\delta_{A^+}$ is a 7-element idempotent monoid, which implies that the AD of $A^+$ can not be effective. For any language $L$, $Q_L$ defined in the proof of Proposition 19 is called *the primitive spectrum of L*. In what follows we prove that, for any BL with finite primitive spectrum, its AD is effective.

**Lemma 20.** *Let L be a subsemigroup of $A^*$. If $|Q_L| > 1$, then $|Q_L| = \infty$. In particular, an EBL L which is not AEBL has an infinite primitive spectrum.*

*Proof.* It is well-known that the equation $x^m y^n = z^p$, $m, n, p \geq 2$ in $A^*$ has only trivial solutions, i.e., $x, y, z$ must be powers of the same word (cf. [4]). So if $p, q \in Q_L$ with $p \neq q$ and $p^m, q^n \in L$, then $p^{tm} q^{kn} \in L \bigcap Q(A)$ for all $t, k \geq 2$. □

**Proposition 21.** *If $L \in L(A)$ has a finite primitive spectrum, the AD of L is effective and $M(L)$ is a finite monoid with commuting idempotents.*

*Proof.* Since $L \in L(A)$, it can be decomposed as a disjoint union of some EBLs: $L = \bigcup\{L_i : i \in I\}$ which is effective by definition. Because $Q_L$ is finite, every $L_i$ in this decomposition is an AEBL by Lemma 20. So this decomposition is the AD of $L$, because of the uniqueness of the AD. Further $|I|$ must be finite, so $L$ is a union of finite rational languages with commuting idempotents which implies that $L$ itself is a rational language with commuting idempotents from [1] and [5]. □

## References


[1] C. J. Ash, *Finite Semigroups with Commuting Idempotents*, J. Austral. Math.Soc. (Series A) **43** (1987), 81–90.

[2] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups*, Math.Surveys No.7, Amer.Math.Soc., Providence, R.I. **I (1961), II (1967)**.

[3] G. Lallement, *John Wiley and Sons, New York*, 1979.

[4] R. C. Lyndon and M. P. Schutzenberger, *The Equation $a^m = b^n = c^p$ in a Free Monoid*, Michigan Math. J **9** (1962), 289–298.

[5] S. W. Margolis and J. E. Pin, *Inverse Semigroups and Varieties of Finite Semigroups*, Journal of Algebra **110** (1987), 306–323.

[6] J. E. Pin, *North Oxford Academic, London*, 1986.

[7] H. J. Shyr, *Institute of Applied Mathematics, National Chung-Haing University, Taichang, Taiwan*, 1991.

DEPART MATH., SICHUAN NORMAL UNIVER., 610066, PRC