Séminaire Dubreil.
Algèbre et théorie
des nombres

Janusz A. Brzozowski

Hierarchies of aperiodic languages

Séminaire Dubreil. Algèbre et théorie des nombres, tome 28, nº 1 (1974-1975), exp. nº 19,
p. 1-11

<http://www.numdam.org/item?id=SD_1974-1975__28_1_A13_0>

Séminaire P. DUBREIL
(Algèbre)
28e année, 1974/75, n° 19, 11 p.

14 avril 1975

HIERARCHIES OF APERIODIC LANGUAGES

by Janusz A. BRZOZOWSKI

[Univ. Waterloo, and (1974/75) Univ. Paris-VI and Paris-VII]

Abstract. - In recent years, considerable attention has been given to the family of aperiodic languages, also known as star-free languages and noncounting regular languages. Several interesting subfamilies of aperiodic languages have been studied and characterized by the properties of the corresponding syntactic monoids. The study of such families has been systematized by examining the position of each family in certain natural hierarchies. This paper gives a brief survey of results in this area.

1. Introduction.

If A is a finite, non-empty alphabet, A^+ (respectively A^*) is the free semigroup (respectively free monoid) generated by A. The empty word is denoted by 1, and Φ is the empty set. Any subset L of A^* is a language. The length of a word $w \in A^*$ is denoted by $|w|$. The cardinality of a set X is denoted $\text{card}\, X$. The symbol $\stackrel{\Delta}{=}$ means "is by definition".

Given languages L, $L' \subset A^*$, the following are also languages : $L \cup L'$ (union), $L \cap L'$ (intersection), $\overline{L} = A^* - L$ (complement),

$$L.L' \stackrel{\Delta}{=} \{w \; ; \; w = uu' \, , \; u \in L \, , \; u' \in L'\} \quad \text{(concatenation or product)} ,$$

$L^+ \stackrel{\Delta}{=} \bigcup_{n \geq 1} L^n$ (the subsemigroup of A^* generated by L), and

$$L^* \stackrel{\Delta}{=} \bigcup_{n \geq 0} L^n = L^+ \cup 1 \quad \text{(the submonoid of } A^* \text{ generated by } L \text{)}.$$

Let $\mathfrak{U}_A$ (or simply $\mathfrak{U}$ when A is understood) be the family of all languages over A. Evidently, $\mathfrak{U}_A$ is a boolean algebra under union, intersection and complement, and a monoid under concatenation.

Let $\mathcal{L}_A \stackrel{\Delta}{=} \{\{a\} \; ; \; a \in A\}$ and let $\mathcal{W}_A \stackrel{\Delta}{=} \{\{x\} \; ; \; w \in A^*\}$. Let $\mathfrak{F}_A$ be the family of all finite languages, and $\mathcal{C}_A \stackrel{\Delta}{=} \{L \subset A^* \; ; \; \overline{L} \in \mathfrak{F}_A\}$ the family of cofinite languages.

For a given family $\mathfrak{X}$ of languages, consider the following properties :

(a) $(L \, , L' \in \mathfrak{X}) \Longrightarrow (L \cup L' \in \mathfrak{X})$,

(b) $(L \in \mathfrak{X}) \Longrightarrow (\overline{L} \in \mathfrak{X})$,

(c) $\{1\} \in \mathfrak{X}$ and $(L \, , L' \in \mathfrak{X}) \Longrightarrow (LL' \in \mathfrak{X})$,

(d) $(L \in \mathfrak{X}) \Longrightarrow (L^* \in \mathfrak{X})$.

It is well known ([6], [9]) that the family of rational or regular languages can be defined as the smallest family containing $\mathcal{L}_A$ and satisfying (a), (c) and (d), and that this family also satisfies (b). Aperiodic languages can be defined as the smallest family containing $\mathcal{L}_A$, and satisfying (a), (b) and (c). (For further reading on aperiodic languages, see [3], [5], [7], [11], [12], [15], [16].)

In the study of aperiodic languages, it is useful to separate the closure under boolean operations from the closure under concatenation. For any family $\mathfrak{X} \subset \mathfrak{U}$, denote by $\mathfrak{X}B$ the boolean algebra generated by $\mathfrak{X}$, i. e. the smallest family containing $\mathfrak{X}$ and satisfying (a) and (b). Similarly, $\mathfrak{X}M$ denotes the monoid generated by $\mathfrak{X}$.

2. Aperiodic languages over a one-letter alphabet.

For $A = \{a\}$, the family $\mathcal{A}_a$ of aperiodic languages is particularly simple (we use $\mathcal{A}_a$ for $\mathcal{A}_{\{a\}}$, etc.) We have $\mathcal{L}_a M = \mathcal{W}_a = \{\{a^n\} ; n \geq 0\}$. Define

$$\mathcal{B}_a \stackrel{\Delta}{=} \mathcal{L}_a MB .$$

One verifies that $\mathcal{B}_a = \mathcal{F}_a \cup \mathcal{C}_a$ and that $\mathcal{B}_a M = \mathcal{B}_a$. This follows because each cofinite language $L \in \mathcal{C}_a$ can be written $L = F \cup a^n a^*$ for some $n \geq 0$ and $F \in \mathcal{F}_a$, and concatenation of languages over a one-letter alphabet is commutative. This implies that $\mathcal{B}_a = \mathcal{A}_a$, i. e. a language over a one-letter alphabet is aperiodic if, and only if, it is either finite or cofinite.

If we start by closing $\mathcal{L}_a$ under Boolean operations, we find

$$\tilde{\mathcal{B}}_a \stackrel{\Delta}{=} \mathcal{L}_a M = \{\emptyset , \{a\} , a^* , a^* - a\} .$$

Next note that

$$\tilde{\mathcal{M}}_a \stackrel{\Delta}{=} \tilde{\mathcal{B}}_a M \supset \{\{a\} , a^*\}M = (\mathcal{L}_a \cup a^*)M = (\mathcal{W}_a \cup a^*)M$$

and

$$\tilde{\mathcal{M}}_a B \supset (\mathcal{W}_a \cup a^*)MB .$$

One verifies that $(\mathcal{W}_a \cup a^*)MB \supseteq \mathcal{F}_a \cup \mathcal{C}_a$, from which it follows that

$$\tilde{\mathcal{M}}_a B \supset \mathcal{B}_a = \mathcal{A}_a .$$

Since obviously $\tilde{\mathcal{M}}_a B \subset \mathcal{A}_a$, we have $\mathcal{A}_a = \mathcal{B}_a = \tilde{\mathcal{M}}_a B$.

These observations are summarized in Fig. 1. For each inclusion, we provide an example of a language which proves the inclusion is proper.

3. Initial phenomena [5].

We now assume that A is fixed, and $\text{card } A > 1$. We use $\mathcal{L}$ for $\mathcal{L}_A$, etc. As in the one-letter case, we have $\mathcal{W} = \mathcal{L}M$ and $\mathcal{B}_0 \stackrel{\Delta}{=} \mathcal{L}MB = \mathcal{F} \cup \mathcal{C}$. However $\mathcal{B}_0 M \neq \mathcal{B}_0$ since (for $A = \{a , b\}$) the language $\{a , b\}^*.a = \Phi.a$ is in $\mathcal{B}_0^2$, but is neither finite nor cofinite. Thus we proceed to define $\mathcal{M}_1 \stackrel{\Delta}{=} \mathcal{B}_0 M$ and $\mathcal{B}_1 \stackrel{\Delta}{=} \mathcal{M}_1 B$. We will return to these families later. For now observe that $\mathcal{B}_1 = (\mathcal{F} \cup \mathcal{C})MB = (\mathcal{W} \cup A^*)MB$, since each cofinite language can be written $L = F \cup A^n A^*$, for some $n \geq 0$ and $F \in \mathfrak{F}$. Hence L can be written as a union of products where each factor is either A^* or it is in $\mathfrak{W}$.

If we close $\mathfrak{L}$ under boolean operations first, we find

$$\tilde{\mathcal{B}}_0 \stackrel{\Delta}{=} \{L \ ; \ L \subset A\} \cup \{L \ ; \ \overline{L} \subset A\} \ .$$

Thus $\tilde{\mathcal{B}}_0$ is a finite boolean algebra with $\mathfrak{L} \cup \overline{A}$ as the set of atoms. Note that $\tilde{\mathcal{B}}_0 \supset \mathfrak{L} \cup A^*$. Next

$$\tilde{\mathcal{M}}_1 \stackrel{\Delta}{=} \tilde{\mathcal{B}}_0 \ M \supset (\mathfrak{L} \cup A^*)M = (\mathfrak{W} \cup A^*)M \ .$$

Thus $\tilde{\mathcal{M}}_1 \ B \supset (\mathfrak{W} \cup A^*)MB = \mathcal{B}_1$. Conversely,

$$\mathcal{B}_1 = \mathfrak{L}MBMB \supset \mathfrak{L}BMB = \tilde{\mathcal{M}}_1 \ B \ ,$$

and $\mathcal{B}_1 = \tilde{\mathcal{M}}_1 \ B$.

These properties are summarized in Fig. 2. It is seen that, except for the few initial differences, it is not important whether $\mathfrak{L}$ is closed under B or M first, since the two sequences coincide from $\mathcal{B}_1$ on.

4. The dot-depth hierarchy [5].

The sequence $(\mathcal{B}_i)$ of boolean algebras, defined below, is called the dot-depth hierarchy. Let

$$\mathcal{B}_0 \stackrel{\Delta}{=} \mathfrak{L}MB$$

$$\mathcal{B}_{n+1} \stackrel{\Delta}{=} \mathcal{B}_n \ MB = \mathcal{B}_0(MB)^n = \mathfrak{L}(MB)^{n+1} \ , \text{ for } n \geq 0 \ .$$

For each aperiodic language L , there exists $n \geq 0$ such that $L \in \mathcal{B}_n$; hence $\mathcal{A} = \bigcup_{n\geq 0} \mathcal{B}_n$.

The "position" of a language in the dot-depth hierarchy can be used as a measure of its complexity. Define the dot-depth (or simply depth) of a language L by

$$d(L) = 0 \ , \text{ if } L \in \mathcal{B}_0$$

$$d(L) = n \ , \text{ if } L \in \mathcal{B}_n - \mathcal{B}_{n-1} \ , \text{ for } n > 0 \ .$$

The depth $d(L)$ corresponds to the minimum number of concatenation levels that must be used to generate L from languages in $\mathcal{B}_0$. Also, $\tilde{\mathcal{B}}_0$ can be used instead of $\mathcal{B}_0$ since $\tilde{\mathcal{B}}_0 \ MB = \mathcal{B}_0 \ MB$; however, $\mathcal{B}_0$ appears to be a more natural starting point (see fig. 3).

The question whether the dot-depth hierarchy is finite or infinite is open (for card $A > 1$). It is known that, for $A = \{a \ , b \ , c\}$, the language

$$L_1 = a^* b\{a \ , b \ , c\}^* = \overline{\overline{\Phi}.\{b \ , c\}.\overline{\Phi}.b.\overline{\Phi}}$$

is of depth 2 , i. e. $L_1 \in \mathcal{B}_2 - \mathcal{B}_1$ [5]. An example over a two letter alphabet is $L_2 = \{ab \ , ba\}^*$. It was shown in SIMON [18] that $L_2 \notin \mathcal{B}_1$, and in McNAUGHTON-PAPERT [11] that $L_2 \in \mathcal{B}_2$.

An upper bound for $d(L)$ has been found as follows [5]. Let n be the number of states in the reduced deterministic finite automaton $\mathfrak{A}_L$ recognizing L. Let $i_a(L)$ be the number of distinct states in input column a, $a \in A$, of the state table of $\mathfrak{A}_L$. Further, let

$$i(L) = \max\{i_a(L) \;;\; a \in A \text{ and } i_a(L) \neq n\}.$$

Then $d(L) \leqslant i(L) + 1$.

This bound is met by L_1 above. On the other hand, let $L_n = a^n a^*$, $n \geqslant 0$ be over $A = \{a\}$. One verifies that $i(L_n) = n - 1$, although L_n is cofinite, and $d(L_n) = 0$.

5. The finite-cofinite hierarchy [4].

In the dot-depth hierarchy, $\mathcal{B}_1 = \mathcal{B}_0 MB$, i. e. a language in $\mathcal{B}_1$ is a boolean function of products of any number of factors from $\mathcal{B}_0$. Thus only one level of concatenation is required, but this concatenation is unlimited in the number of factors. A finer measure of complexity is obtained by limiting the number of factors as follows. Let $\beta_n \stackrel{\Delta}{=} \mathcal{B}_0^n B$ for $n \geqslant 1$. Then

$$\beta_n \subset \beta_{n+1} \quad \text{and} \quad \mathcal{B}_1 = \bigcup_{n \geqslant 1} \beta_n .$$

A number of subfamilies of aperiodic languages that have been studied appear naturally in the sequence $\mathcal{B}_0 = \beta_1 \subset \beta_2 \subset \dots \subset \mathcal{B}_1$ which we refer to as the finite-cofinite hierarchy. We will also need :

$$\beta_{2L} \stackrel{\Delta}{=} (\mathfrak{F}^2 \cup \mathfrak{C}\mathfrak{F} \cup \mathfrak{C}^2)B = (\mathfrak{F} \cup \mathfrak{C}\mathfrak{F} \cup \mathfrak{C})B \subset \beta_2 ,$$

$$\beta_{2R} \stackrel{\Delta}{=} (\mathfrak{F}^2 \cup \mathfrak{F}\mathfrak{C} \cup \mathfrak{C}^2)B = (\mathfrak{F} \cup \mathfrak{F}\mathfrak{C} \cup \mathfrak{C})B \subset \beta_2 .$$

An alternate description of the β families is the following :

$$\beta_1 = \mathcal{B}_0 = \mathfrak{W}B = (\mathfrak{W} \cup A^*)B$$

$$\beta_{2L} = (\mathfrak{W} \cup A^* \mathfrak{W})B$$

$$\beta_{2R} = (\mathfrak{W} \cup \mathfrak{W}A^*)B$$

$$\beta_n = (\mathfrak{W} \cup A^*)^n B = (\mathfrak{F} \cup \mathfrak{C})^n B , \text{ for } n \geqslant 1 ,$$

where $A^* \mathfrak{W} = \{A^* L \;;\; L \in \mathfrak{W}\}$, etc. These claims are easily verified. One can also show [4] that

$$(\mathfrak{W} \cup A^*)^{2n+1} B \supset (\mathfrak{W} \cup A^*)^{2n+2} B , \text{ for } n \geqslant 1 .$$

Therefore $\beta_{2n+2} = \beta_{2n+1}$; however $\beta_{2n+3} \neq \beta_{2n+1}$ for all $n \geqslant 1$ [18].

A language is definite ([1], [9], [13]) if, and only if, it is in β_{2L}, reverse definite ([1], [8]) if, and only if, it is in β_{2R}, generalized definite ([8], [17]) if, and only if, it is in β_2, and locally testable [11] if, and only if, it is in β_3. The original definitions of these families of languages were somewhat different ; however the equivalence of definitions is easily proved [4], and the present formulation appears more natural. We reconsider these families later.

The statements about the finite-cofinite hierarchy are summarized in Fig. 4.

6. The alphabetic hierarchy [18].

The languages introduced here play a key role in the family of depth-one languages. We introduce a family $\alpha_{1,1}$ of languages (The reason for this notation is explained in Section 8.) such that, if $L \in \alpha_{1,1}$, the membership of a word x in L can be determined solely by the set of letters appearing in x. Define

$$x\alpha \overset{\Delta}{=} \{a \in A \; ; \; x = uav \; , \; u \, , v \in A^*\}$$

to be the "alphabet" of $x \in A^*$.

For $x, y \in A^*$, let $x \equiv_\alpha y$ if, and only if, $x\alpha = y\alpha$. The relation $\equiv_\alpha$ is a congruence of finite index on A^*, there being one congruence class $([x]_\alpha)$ for each subset of A. We have

$$[x]_\alpha = (\cap_{a \in x\alpha} A^* aA^*) \cap (\cap_{a \notin x\alpha} \overline{A^* aA^*}) \; .$$

Now define $\alpha_{1,1} \overset{\Delta}{=} \{[x]_\alpha \; ; \; x \in A^*\}B$, and let $A^* \mathcal{L} A^* \overset{\Delta}{=} \{A^* aA^* \; ; \; a \in A\}$. One verifies that $\alpha_{1,1} = (A^* \mathcal{L} A^*)B$. For technical reasons, we use the family $\mathcal{S} \overset{\Delta}{=} A^* \cup A^* \mathcal{L} A^*$ as a generating set for $\alpha_{1,1}$. Note that $\mathcal{S}^m \subset \mathcal{S}^{m+1}$ for $m \geqslant 1$, and we will use the convention $\mathcal{S}^0 = \{\Phi\}$. Let $\alpha_{m,1} \overset{\Delta}{=} \mathcal{S}^m B$ and $\gamma_1 \overset{\Delta}{=} \mathcal{S}MB$. We find $\gamma_1 = \mathcal{S}MB = \cup_{m \geqslant 0} \mathcal{S}^m B = \cup_{m \geqslant 0} \alpha_{m,1}$ and the sequence

$$\alpha_{0,1} \subset \alpha_{1,1} \subset \alpha_{2,1} \subset \dots \subset \gamma_1$$

will be called the <u>alphabetic hierarchy</u>. $L \subset A^*$ is 0-<u>alphabetic</u> if $L \in \alpha_{0,1}$, and it is m-<u>alphabetic</u>, $m \geqslant 1$, if, and only if, $L \in \alpha_{m,1} - \alpha_{m-1,1}$. Finally L is <u>alphabetic</u> if, and only if, $L \in \gamma_1$.

An alternate description of $\alpha_{m,1}$ is obtained by using the "shuffle" operator $\sqcup$ [7]. For $w = a_1 a_2 \dots a_m \in A^*$,

$$A^* \sqcup w = A^* \sqcup (a_1 a_2 \dots a_m) \overset{\Delta}{=} A^* a_1 A^* a_2 A^* \dots a_m A^*$$

and $A^* \sqcup \; = \{A^* \sqcup \mathcal{W} \; ; \; w \in A^*\}$.

Let $\mathcal{W}_{\leqslant m} = \{\{w\} \; ; \; w \in A^*$ and $|w| \leqslant m\}$. One verifies that $\alpha_{m,1} = (A^* \sqcup \mathcal{W}_{\leqslant m})B$, for $m \geqslant 0$ and $\gamma_1 = (A^* \sqcup \mathcal{W})B$.

Over a two-letter alphabet $A = \{a, b\}$, the alphabetic languages can be viewed as a generalization of finite-cofinite languages [2]. Let $\mathcal{L}^\oplus \overset{\Delta}{=} \{\{a\},\{b\},a^+,b^+\}$ and $\mathcal{W}^\oplus \overset{\Delta}{=} \mathcal{L}^\oplus M$ be the generalization of $\mathcal{L}$ and $\mathcal{W}$, respectively. Let $\mathcal{F}^\oplus$ be the closure of the family $\mathcal{W}^\oplus$ under finite unions and let $\mathcal{C}^\oplus \overset{\Delta}{=} \{L \; ; \; \overline{L} \in \mathcal{F}^\oplus\}$. Then it can be shown that

$$\gamma_1 = \mathcal{F}^\oplus \cup \mathcal{C}^\oplus = \mathcal{L}^\oplus MB \; .$$

Furthermore, the initial phenomena of Fig. 2 have their counterpart here, for

$$\mathcal{L}^\oplus BMB = \mathcal{L}^\oplus MBMB \; .$$

7. The locally-testable hierarchy ([4], [18]).

It can be shown that the membership of a word x in a locally-testable language L is determined solely by the first $k-1$ letters of x, the last $k-1$ letters of x and the set of words of length k that appear in x. Formally, $f_k(x)$ (respectively $t_k(x)$) is x, if $|x| \leqslant k$, and it is the prefix (respectively suffix) of x of length k otherwise. Let

$$m_k(x) \stackrel{\Delta}{=} \{w \in A^* ; \quad x = uwv \quad \text{and} \quad |w| = k\} .$$

For $x, y \in A^*$ and $k > 0$, define the congruence

(*) $\quad x \sim_k y$ if and only if, $f_{k-1}(x) = f_{k-1}(y)$, $t_{k-1}(x) = t_{k-1}(y)$ and $m_k(x) = m_k(y)$.

If $[x]_k$ is the congruence class containing x, let $\alpha_{1,k} \stackrel{\Delta}{=} \{[x]_k ; \quad x \in A^*\}B$, be the family of k-testable languages. The reason for this notation will soon be explained. Note however that it is consistent with that of section 6, because L is 1-alphabetic if, and only if, L is 1-testable if, and only if, $L \in \alpha_{1,1}$. One verifies that $\alpha_{1,k} \subset \alpha_{1,k+1}$ and that $\beta_3 = \bigcup_{k \geqslant 1} \alpha_{1,k}$.

If, in the definition (*) of $\sim_k$, we remove the condition $m_k(x) = m_k(y)$, we obtain the family of k-generalized-definite languages which we denote by $\alpha_{0,k}$. One verifies that $\beta_2 = \bigcup_{k \geqslant 1} \alpha_{0,k}$.

8. Simon's depth-one hierarchy [18].

Roughly speaking, the membership of a word x in a language L of depth 1 can be determined by testing $f_{k-1}(x)$, $t_{k-1}(x)$ and the set $\mu_{m,k}$ of m-tuples of words of length k that appear in x. Thus depth-one languages are generalizations of both the k-testable and m-alphabetic languages ; the locally testable and alphabetic hierarchies turn out to be "orthogonal".

Let $W = (w_1, \ldots, w_m)$ be an m-tuple of words of length k. We say that W occurs in x if, and only if, there exist words $u_1, \ldots, u_m, v_1, \ldots, v_m$ such that $|u_1| < |u_2| < \ldots < |u_m|$ and $x_i = u_i w_i v_i$, for $i = 1, \ldots, m$. Let

$$\mu_{m,k}(x) = \{W | W = (w_1, \ldots, w_m), \ |w_1| = \ldots = |w_m| = k \text{ and } W \text{ occurs in } x\} .$$

By convention $\mu_{0,k} = \emptyset$ for all $k \geqslant 1$. Note that $\mu_{m,k}(x) = \emptyset$ if, and only if, $|x| < m + k - 1$.

For $x, y \in A^*$, $m \geqslant 0$, $k \geqslant 1$ define $x \; {}_m{\sim}_k \; y$ if, and only if,

(a) $\qquad x = y$ if $|x| < m + k - 1$

or

(b) $f_{k-1}(x) = f_{k-1}(y)$, $t_{k-1}(x) = t_{k-1}(y)$ and $\mu_{m,k}(x) = \mu_{m,k}(y)$, otherwise.

The relation ${}_m{\sim}_k$ is a congruence of finite index on A^*. Let

$$\alpha_{m,k} \stackrel{\Delta}{=} \{{}_m[x]_k ; \quad x \in A^*\}B .$$

One verifies that this is consistent with the previous definitions.

The hierarchy defined by $\underset{m}{\sim}_k$ is illustrated in Fig. 4, where $\gamma_i = \bigcup_{m\geq 0} \alpha_{m,k}$ and (one verifies that) $\beta_{2m+1} = \bigcup_{k\geq 1} \alpha_{m,k}$ for $m \geq 1$ (The case of β_2 is somewhat degenerate). All the hierarchies shown are known to be infinite.

9. Syntactic monoids.

For $L \subset A^*$ the syntactic congruence, $\equiv_L$, is defined by $x \equiv_L y$ if, and only if, for all $u, v \in A^*$, $(u x v \in L) \iff (uyv \in L)$. The quotient monoid $A^*/\equiv_L$ is called the syntactic monoid M_L of L, and $A^+/\equiv_L$ is the syntactic semigroup, S_L.

It is well-known that L is rational (or regular) if, and only if, M_L is finite. It has been shown by SCHÜTZENBERGER ([15], [16]) that L is aperiodic if, and only if, M_L is group-free, (contains no groups other than the trivial one-element groups).

A number of families of languages in Simon's hierarchy have been characterized by the properties of their syntactic monoids. In this connection, the alphabetic hierarchy plays a key role. The following is known [18] :

(1) $L \in \alpha_{0,1} = \{\emptyset , A^*\}$ if, and only if, $M_L = 1$.

(2) $L \in \alpha_{1,1}$ if, and only if, M_L is idempotent and commutative.

(3) $L \in \gamma_1$ if, and only if, M_L is $\mathcal{J}$-trivial, i. e. for all $m, m' \in M_L$

$$(M_L mM_L = M_L m' M_L) \text{ implies } (m = m') .$$

These properties appear to carry over to the finite-cofinite hierarchy as follows:

" $L \in \alpha_{m,1}$ if, and only if, M_L has property P " seems to correspond to " $L \in \beta_{2m+1}$ if, and only if, for each idempotent $e \in S_L$ the submonoid $e S_L e$ has property P ".

The following evidence supports this statement :

(1^*) $L \in \beta_2$ if, and only if, $e S_L e = e$ ([4], [14], [19]).

(2^*) $L \in \beta_3$ if, and only if, $e S_L e$ is idempotent and commutative ([4], [10], [19], [20], [21]).

(3^*) If $L \in \mathcal{B}_1$, then $e S_L e$ is $\mathcal{J}$-trivial [18].

As can be seen, the results are quite fragmentary, and the proofs of these results are quite complex. This approach appears to be very fruitful not only for classifying languages, but also monoids. For a more detailed account of the relationship between languages and monoids, see [7].

For the sake of brevity, we have not touched upon the relationship between aperiodic languages and finite automata. There exist also characterizations of the families of depth-one languages by the properties of the corresponding finite automata. We refer the reader to [18].

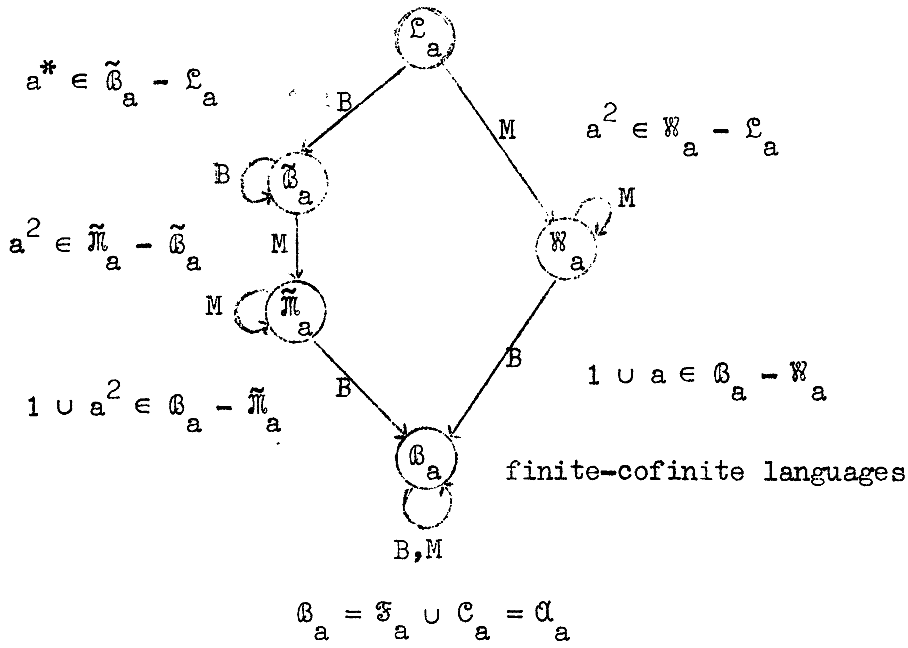

Fig. 1 : Aperiodic languages over a one-letter alphabet

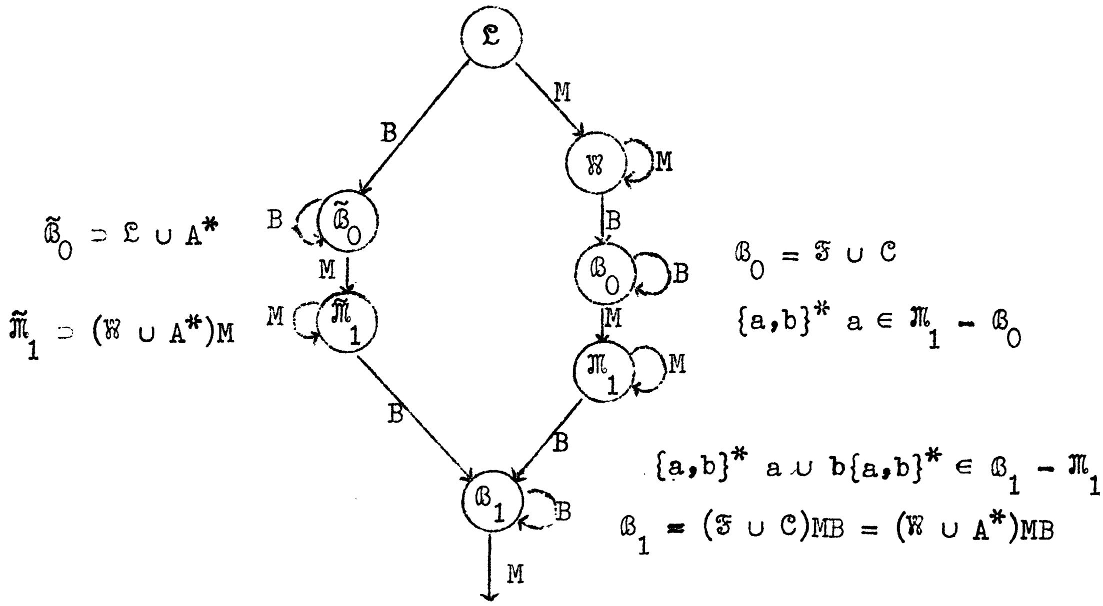

Fig. 2 : Initial families for card $A > 2$

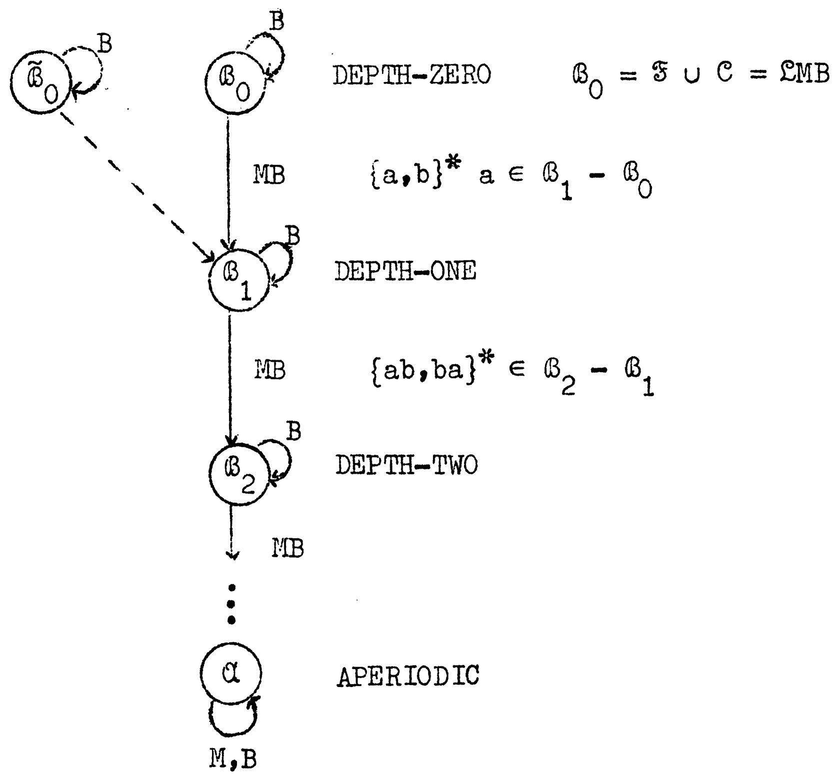

Fig. 3 : The dot-depth hierarchy

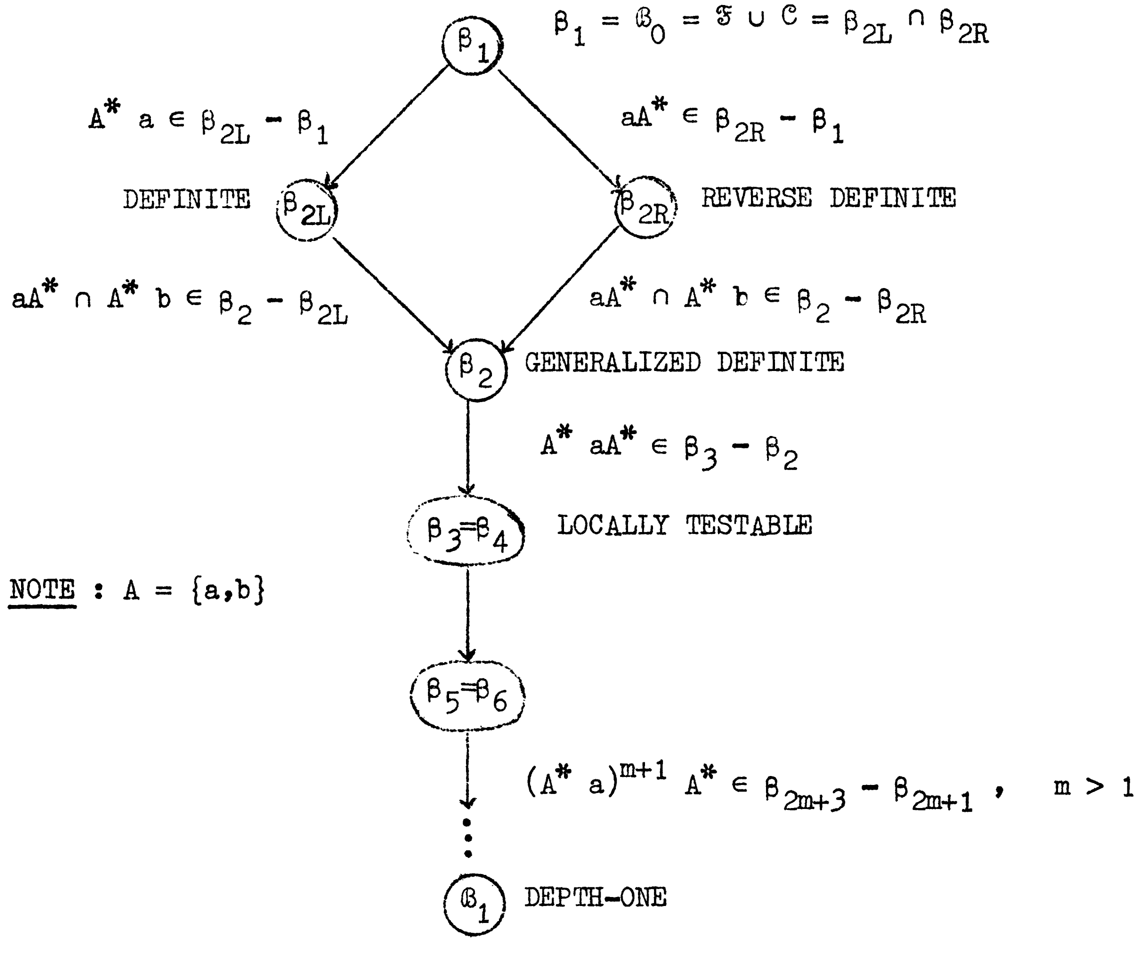

Fig. 4 : The finite-cofinite hierarchy

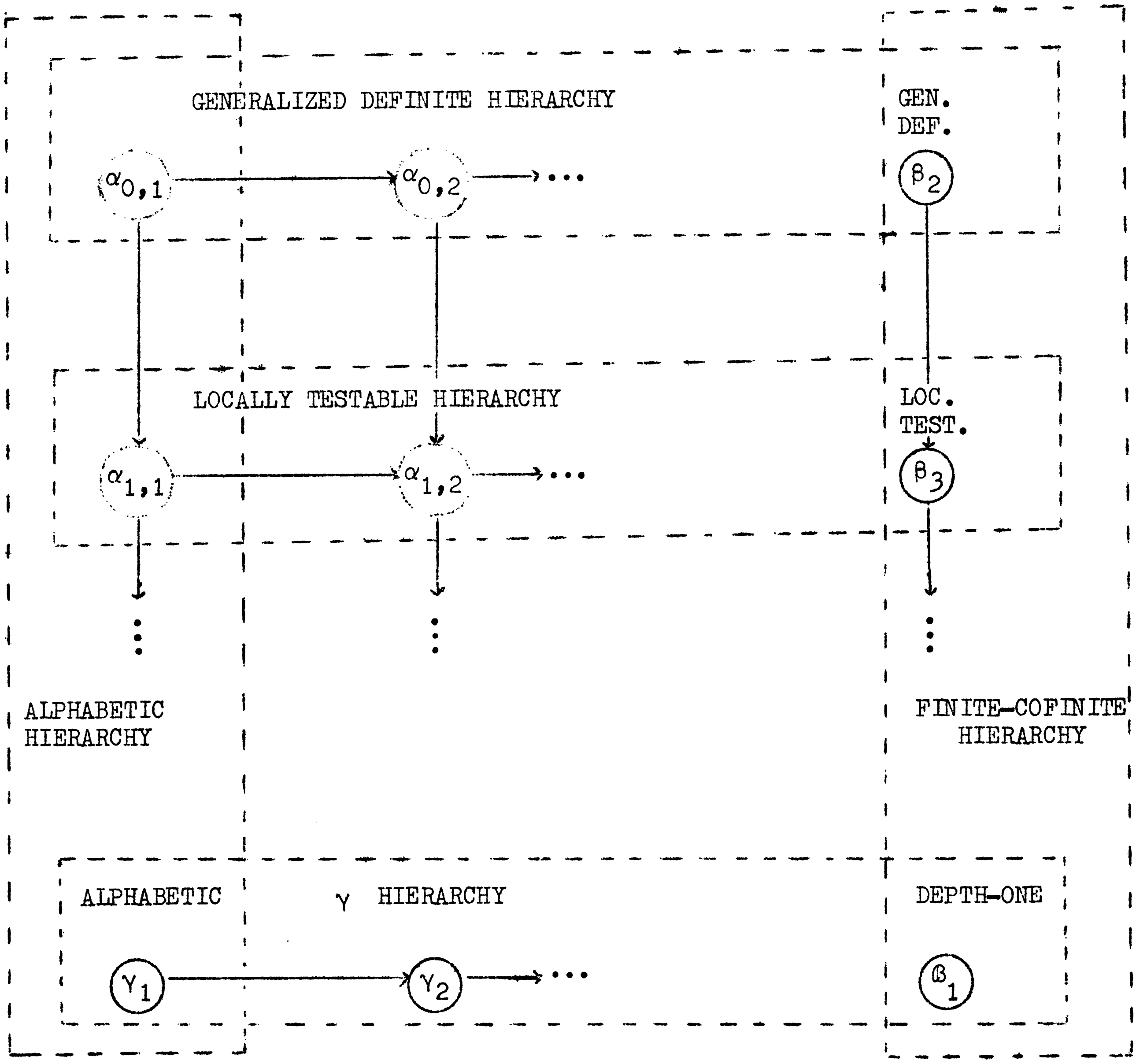

Fig. 5 : Simon's depth-one hierarchy

REFERENCES

[1] BRZOZOWSKI (J. A.). - Canonical regular expressions and minimal state graphs for definite events, "Mathematical theory of automata" New York, 1962, p. 529-561. - Brooklyn, Polytechnic Institute of Brooklyn, 1963 (Symposia Series, 12).

[2] BRZOZOWSKI (J. A.). - J-trivial monoids and alphabetic languages over two letters (unpublished manuscript, February 1975).

[3] BRZOZOWSKI (J. A.), CULIK II (K.) and GABRIELIAN (A.). - Classification of noncounting events, J. Computer and System Sc., t. 5, 1971, p. 41-53.

[4] BRZOZOWSKI (J. A.) and SIMON (I.). - Characterisations of locally testable events, Discrete Mathematics, t. 4, 1973, p. 243-271.