KYBERNETIKA — VOLUME 13 (1977), NUMBER 5

On Families Recognizable by Finite Branching Automata

VÁCLAV BENDA, KAMILA BENDOVÁ

Various topics concerning families of languages recognizable by recently introduced finite branching automata are investigated. The attention is paid to questions with answers similar to results from the "classical" automata theory, e.g. characterization of recognizable family by means of finite number of regular languages or its algebraic decomposition (cf. Section 3 and 4), as well as to questions with different answers, e.g. the class of all recognizable families is not closed under union and complement (cf. Section 2). Moreover, some results are presented which have no natural counterpart in the classical theory, e.g. about infinite cardinalities, about strong and well-recognizable families (cf. Section 3 and 5).

INTRODUCTION

Finite automata were introduced more than two decades ago as a formal model of discrete systems characterized by a finite number of states. In subsequent years a great deal of research has been concerned with this model and a number of its variants and extensions were proposed, motivated by different areas of theoretical computer science as well as by some application fields. One such extension, connected with the notion of a *finite branching automaton*, was recently suggested ([1]) in an attempt to provide a formal mathematical model for a generalized state space, which is a structure useful for formal problem solving and plan formation in the framework of artificial intelligence. This novel approach yields a variety of new problems, some of them interesting from the point of view of the mentioned motivation, other merely for their unusual flavor comparing to the classical problems of automata theory and of formal languages.

The state space — a finite set of states together with a set of actions transforming one state into another — seems to be the most important underlying structure not only for formal problem solving, but also for other related disciplines that may be grouped under a common heading of "mathematics of action". A straight-forward approach would treat any sequence of actions leading from an initial state to one of specified goal states as a solution to a problem, or equivalently, as a plan how to achieve the goal. Such an idea of *straight-line plans* is, however, not quite satisfactory when one considers planning as an imaginative activity without sufficient apriori knowledge of the real circumstances in the future. This leads to the idea of *branching plans* where possible alternatives have to be always considered explicitely, with "success" being just one additional alternative.

In the automata-theoretic treatment, the straight-line plans may be represented by successful, or accepted, strings of letters in a classical finite automaton, while the branching plans are represented by formal languages (sets of strings), accepted by a finite branching automaton. Correspondingly, the collection of all successful straight-line plans is in the former case represented by the *language recognized by the* (classical) *automaton* and the collection of all branching plans is in the latter case represented by the *family of languages*, which is *recognized by the finite branching automaton.*

The aim of the present paper*) is to answer various particular questions – some of them posed as open problems in [1] – concerning families of formal languages recognizable by finite branching automata. We concentrate both on questions where the answers directly refer to those from the classical theory of regular languages and on questions which are specific to the theory of finite branching automata and have no obvious classical counterparts**).

The paper is divided into five sections. After this introduction we present in Section 1 necessary preliminaries including a survey of those results from [1] that we shall need later. The second section deals with the results specific to the theory of finite branching automata, namely, that the class of recognizable families of languages is, in contrast to the class of regular languages, not closed under the usual operations of union, complement and intersection, though these operations still have similar intuitive interpretation here as in the classical theory.

The third section mainly attempts to establish characterization of recognizable families of languages with the help of the regular languages. Specific problems are discussed in subsection devoted to "strong" families. It is shown that while in the classical theory it is possible to add strings to any language in such a way that the resulting language is regular this is not always possible in the case of recognizable families. From methodological point of view it is perhaps interesting to mention that in the theory of finite branching automata (where we work with families of languages whose cardinality can be $\aleph_1$***) and the class of all families has cardinality $\aleph_2$) mere comparison of infinite cardinals can yield interesting results.

*) Some of the results of this paper were presented at the MFCS'76 Symposium in Gdaňsk (cf. [2]).

**) Some other results of this character appeared also in another paper [3].

***) We tacitly use the continuum hypothesis.

The fourth section presents a decomposition theorem for recognizable families and also prepares ground for the following section.

The most extensive fifth section deals with an important subclass of the class of recognizable families: the class of families with recognizable complements. The complement of recognizable family represents a collection of plans which may fail. The following natural interpretation suggests itself: families with recognizable complements are just the families with "explicite" goal, i.e. with goal which characterizes the successful trends independently from possible configurations of state space (thus not substantiating the slogan: "ends justify the means" which – as it is known – has not originated from Jesuits).

We conclude this introduction by defining some notions of general usage. In the present context an *alphabet* Σ is an arbitrary, but fixed finite non-empty set of objects called *letters* (usually denoted $a, b, c, \ldots$). We denote by Σ^* the set of all finite sequences of letters (the free monoid generated by Σ under concatenation). The elements of Σ^* are called *strings* and usually denoted $u, v, w \ldots$. The unit element in Σ^* is the *empty string* $\Lambda \in \Sigma^*$. We denote $\Sigma_\Lambda = \Sigma \cup \{\Lambda\}$ and $\Sigma^+ = \Sigma^* \setminus \{\Lambda\}$. For $u \in \Sigma^*$, $\lg(u)$ denotes the length of u (the number of occurences of letters in u). In particular, $\lg(\Lambda) = 0$. $\mathscr{L}(\Sigma)$ is the set of all non-empty subsets of Σ^*, elements of $\mathscr{L}(\Sigma)$ are called *languages* (usually denoted L). Any $X \subseteq \mathscr{L}(\Sigma)$ will be called a *family of languages* (*over* Σ). Note that we admit empty family of languages but not families with empty element. In the following L is always a non-empty language. *Singleton* is a family $\{L\}$ consisting of only one language L.

For $u, v \in \Sigma^*$ we define:

1) $u \leqq v \equiv (\exists w \in \Sigma^*)\,[uw = v]$, (i.e. u is a *prefix* of v);

2) $u \parallel v \equiv (u \nleqq v)\,\&\,(v \nleqq u)$;

for $u \in \Sigma^*$, $L \in \mathscr{L}(\Sigma)$ we define:

3) the *derivative* of L with respect to u

$$\partial_u L = \{v;\ v \in \Sigma^*\ \&\ uv \in L\}\ ;$$

4) the *maximal strings* of L

$$\operatorname{Max}(L) = \{u;\ \partial_u L = \{\Lambda\}\}\ ;$$

5) the *prefix closure* of L

$$\operatorname{Pref}(L) = \{u;\ (\exists v \in L)\,[u \leqq v]\}\ ;$$

6) the set of *first letters* of L

$$\operatorname{Fst}(L) = \operatorname{Pref}(L) \cap \Sigma\ ;$$

7) $\operatorname{Fst}_\Lambda(L) = \operatorname{Fst}(L) \cup (\{\Lambda\} \cap L)$ (in [5] denoted by $\Delta_L(\Lambda)$).

Note that the operations defined in points 5) and 7) produce from languages necessarily again languages but other operations produce languages only if the result is non-empty. Besides we shall use current logical connectives and quantifiers and usual symbols $\in$, $\subseteq$ etc.

1. FINITE BRANCHING AUTOMATA: BASIC DEFINITIONS AND PROPERTIES

1.1. Definition. A *finite branching automaton* $\mathscr{B} = \langle Q, \Sigma, \delta, q_0, B\rangle$ consists of an ordinary finite deterministic automaton $\langle Q, \Sigma, \delta, q_0\rangle$ without final states (Q is a set of states, Σ is an alphabet, δ is the transition function and $q_0 \in Q$ is the initial state) and of branching relation $B \subseteq Q \times 2^{\Sigma_\Lambda}$.

We shall use the term "to accept" and "to recognize" with natural distinction according to which a finite automaton accepts strings and recognizes languages; an analogous distinction is made in the case of a finite branching automaton which, however, accepts (non-empty) languages and recognizes families of languages.

1.2. Definition. A language $L \in \mathscr{L}(\Sigma)$ is *accepted* by a finite branching automaton $\mathscr{B} = \langle Q, \Sigma, \delta, q_0, B\rangle$ if for each $u \in \mathrm{Pref}(L)$,

$$\langle \delta(q_0, u), \mathrm{Fst}_\Lambda(\partial_u L)\rangle \in B .$$

We denote by $T(\mathscr{B})$ the family of all languages accepted by $\mathscr{B}$.

1.3. Definition. A family $X \subseteq \mathscr{L}(\Sigma)$ is *recognizable* if $X = T(\mathscr{B})$ for some finite branching automaton $\mathscr{B}$.

1.4. Definition. The *derivative of a family X with respect to a string u* is the family

$$\partial_u X = \{\partial_u L;\ L \in X\} \setminus \{\emptyset\} .$$

We denote $\mathscr{D}(X) = \{\partial_u X;\ u \in \Sigma^*\}$ and we say that X is *finitely derivable* if $\mathscr{D}(X)$ is finite.

1.5. Definition. For every $u \in \Sigma^*$ we define a binary *replacement operator* R_u on $\mathscr{L}(\Sigma)$ as follows: for every $L_1, L_2 \in \mathscr{L}(\Sigma)$,

$$R_u(L_1, L_2) = (L_1 - u\Sigma^*) \cup uL_2 .$$

1.6. Definition. A family $X \subseteq \mathscr{L}(\Sigma)$ has the *replacement property* if for each $L_1, L_2 \in X$ and each $u \in \mathrm{Pref}(L_1) \cap \mathrm{Pref}(L_2)$,

$$R_u(L_1, \partial_u L_2) \in X .$$

The intuitive meaning of the replacement property is best illustrated by visualizing languages as trees where it expresses the possibility of replacing a subtree of L_1 by the corresponding subtree of L_2 starting from the same point. In the natural interpretation the replacement property corresponds to the possibility of "switching" from one plan to another.

We define the *R-closure* of a family X: $R(X) = \bigcup_{n \in N} R^n(X)$ where $R^0(X) = X$ and $R^{n+1}(X)$ is the set of all languages of the form $R_u(L_1, \partial_u L_2)$ where $L_1, L_2 \in R^n(X)$ and $u \in \text{Pref}(L_1) \cap \text{Pref}(L_2)$. It is easy to show that $R(X)$ is the smallest family containing X and having the replacement property.

The following notion of compatibility and compatible closure is obtained by considering in certain sense unbounded application of the replacement operator.

1.7. Definition. Let $L \in \mathscr{L}(\Sigma)$ and $X \subseteq \mathscr{L}(\Sigma)$. L is *compatible with* X if for every $u \in \Sigma^*$ there exists $L_u \in X$ such that

$$\text{Fst}_A(\partial_u L) = \text{Fst}_A(\partial_u L_u).$$

In [1] it is proved (Lemma 4.1) that a language L is compatible with X iff above condition holds for every $u \in \text{Pref}(L)$.

1.8. Definition. *C-closure* of a family X is the family

$$C(X) = \{L;\ L \text{ is compatible with } X\}.$$

We say that a family X is *self-compatible* if $C(X) = X$.

The following facts are true about R-closure, compatibility and finiteness of $\mathscr{D}(X)$ (cf. [1]).

Assertions. a) If X is finitely derivable then also $R(X)$ and $C(X)$ are finitely derivable.

b) If X is a self-compatible (resp. R-closed) family then for every $u \in \Sigma^*$, $\partial_u X$ is a self-compatible (R-closed) family.

c) For every family X it holds $R(X) \subseteq C(X)$. The equality does not hold, in general.

d) However, a finite family X is R-closed iff it is self-compatible.

Characterization Theorem. A family $X \subseteq \mathscr{L}(\Sigma)$ is recognizable iff X is self-compatible and finitely derivable.

At the conclusion of this section we give some examples of recognizable families (and their notations) which shall be useful later.

Example 1. The following families are obviously recognizable:

a) The trivial families, i.e. $\emptyset$ and $\mathscr{L}(\Sigma)$.

b) $\{L;\ \emptyset \neq L \subseteq \Sigma_A\}$. This family is denoted Z and its elements Γ (occasionally Γ_i).

c) The family of all complete languages. A language L is complete iff $\mathrm{Pref}(L) = \Sigma^*$.

Example 2. Every singleton $\{L\}$ is obviously self-compatible and the concept of derivation reduces to the classical case. Thus according to the characterization theorem a singleton $\{L\}$ is recognizable iff L is a regular language. Analogically a family $\{\{u\};\ u \in L\}$ is recognizable iff L is a regular language.

1.9. Definition. $X \neq \mathscr{L}(\Sigma)$ is *strong* if $\mathscr{L}(\Sigma)$ is the only recognizable family containing X as a subfamily.

Example 3. We denote $Fin = \{L;\ L \text{ is a finite language}\}$. Fin is not recognizable and moreover it is a strong family (the proof will be given in Section 3).

2. SET-THEORETICAL OPERATIONS ON FAMILIES OF LANGUAGES

The set-theoretical operations on families, as intersection $(X \cap Y)$, union $(X \cup Y)$ and complement $(\bar{X})$, are defined in obvious way, e.g.,

$$X \cap Y = \{L;\ L \in X \ \&\ L \in Y\}$$

and offer analogical natural interpretation as in the case of the classical theory. If for example X (resp. Y) is a family of all plans for a goal A (resp. B) then $X \cap Y$ is just the family of plans realizing at the same time both goals A and B, $X \cup Y$ is the family realizing at least one of the goals A or B (as for the complement, there was discussed in the introduction and independent Section 5 will be dedicated to it).

While in the classical automata theory the set-theoretical operations preserve recognizability of languages, for finite branching automata we have different results as summarized below (cf. [1], Section 5):

2.1. Facts. a) Finite intersection of recognizable (resp. self-compatible) families is recognizable (resp. self-compatible).

b) Neither union nor complement of recognizable (self-compatible) families is recognizable (self-compatible), in general.

c) Finite union of finitely derivable families is finitely derivable.

Since most of our work will be based on the characterization theorem mentioned above we shall first complete the preceding facts by examining whether the set-theoretical operations preserve the second condition of recognizability, i.e. finite derivability of families.

2.2. Proposition. Complement of a finitely derivable family may not be finitely derivable.

Proof. Let L_0 be any fixed language. We show that the family $X = \mathscr{L}(\Sigma) - \{L_0\}$ is always finitely derivable by showing that, in fact, for every $u \in \Sigma^+$, $\partial_u X = \mathscr{L}(\Sigma)$.

Clearly, either $\{L;\ \Lambda \in L\} \subseteq X$ or $\{L;\ \Lambda \notin L\} \subseteq X$ and since for any $u \in \Sigma^+$, $\partial_u\{L;\ \Lambda \in L\} = \partial_u\{L;\ \Lambda \notin L\} = \mathscr{L}(\Sigma)$, it follows that $\partial_u X = \mathscr{L}(\Sigma)$.

Now it suffices to choose as L_0 an arbitrary non-regular language; according to the above, $X = \mathscr{L}(\Sigma) - \{L_0\}$ is finitely derivable but $\overline{X} = \{L_0\}$ is not finitely derivable. Q.e.d.

2.3. Corollary. Every nontrivial family is a subset of a nontrivial finitely derivable family (i.e. there are no "strong" families in the sense of finite derivability).

Proof. Any nontrivial family is contained in a family with a singleton complement, which, by the proof of Proposition 2.2, is finitely derivable.

2.4. Corollary. Any family with a finite complement is finitely derivable.

Proof. Let us consider X such that $\overline{X} = \{L_i;\ i \in K\}$ where K is a finite set. For each language L_i, we choose some $u_i \in \Sigma^*$ such that $u_i \neq u_j$ for $i \neq j$. For $i \in K$ we put

$$L_i' = (L_i \cup \{u_j;\ u_j \notin L_j \,\&\, j \in K\}) \setminus \{u_j;\ u_j \in L_j \,\&\, j \in K\}\ .$$

Clearly, for every L_i' following holds (if L_i' is non-empty):

1) $L_i' \in X$,

2) if $v \in \Sigma^*$, $\lg(v) > \max_{j \in K} \lg(u_j)$ then $\partial_v L_i = \partial_v L_i'$.

Thus $\partial_v X \neq \mathscr{L}(\Sigma)$ only for finite number of $v \in \Sigma^*$. Hence X is finitely derivable. Q.e.d.

2.5. Proposition. Intersection of finitely derivable families may not be finitely derivable.

Proof. Let us choose a non-regular language L_0 and put

$$X = \{L;\ \Lambda \in L\} \cup \{L_0\}\ ,$$

$$Y = \{L;\ \Lambda \notin L\} \cup \{L_0\}\ .$$

X and Y are clearly finitely derivable because for every $u \in \Sigma^+$,

$$\partial_u X = \partial_u Y = \mathscr{L}(\Sigma)\ .$$

But $X \cap Y = \{L_0\}$ and $\{L_0\}$ is not finitely derivable due to our choice of L_0. Q.e.d.

The last result is somewhat surprising because we know that for recognizable families the intersection preserves the recognizability and hence finite derivability.

This suggests that self-compatibility and finite derivability are not quite independent properties.

Other important operation with a natural interpretation is the concatenation of languages. There are two obvious ways how to extend this operation to the case of families (see Definition 5.1 in [1]).

2.6. Definition. Let $X_1 \subseteq \mathscr{L}(\Sigma_1)$, $X_2 \subseteq \mathscr{L}(\Sigma_2)$ be two families. A strong concatenation of X_1 and X_2 is

$$X_1 \,.\, X_2 = \{L_1 L_2;\ L_1 \in X_1 \ \&\ L_2 \in X_2\}\,.$$

(We shall write only $X_1 X_2$.) A weak concatenation of X_1 and X_2 is

$$X_1 \circ X_2 = \{ \bigcup_{u \in L_1} u F(u);\ \text{where } L_1 \in X_1 \text{ and } F \text{ is a total function } L_1 \to X_2\}\,.$$

As it is usual in automata theory we used the simplified notation $u\ F(u)$ instead of $\{u\}\,.\,F(u)$. In general, we shall write uL instead of $\{u\}\,.\,L$ (which is a language) and uX instead of $\{\{u\}\}\,.\,X$ (which is a family of languages). Since for this case the weak and strong concatenations coincide, uX denoted also the family $\{\{u\}\} \circ X$.

In [1] it was shown (Theorem 5.3) that strong concatenation does not preserve recognizability. For weak concatenation this question was stated as an open problem. The following proposition gives the negative answer.

2.7. Proposition. Weak concatenation of recognizable families may not be recognizable.

Proof. Let $X_1 = \{\{u\};\ u \in \Sigma^*\}$ and $X_2 = Z$ be two recognizable families from examples in Section 1. We fix an arbitrary language $L \in \mathscr{L}(\Sigma)$ and show that $L \in C(X_1 \circ X_2)$.

Define a function $F : \mathrm{Pref}\,(L) \to X_2$ by

$$F(u) = \mathrm{Fst}_A\,(\partial_u L)$$

and for Definition 1.7 choose $L_u = u\ F(u)$ in $X_1 \circ X_2$. Then for every $u \in \mathrm{Pref}\,(L)$ it holds $\mathrm{Fst}_A\,(\partial_u L) = F(u) = \mathrm{Fst}_A\,(\partial_u L_u)$, and hence $L \in C(X_1 \circ X_2)$. Thus $C(X_1 \circ X_2) = \mathscr{L}(\Sigma)$. But $X_1 \circ X_2 \subseteq Fin \subset \mathscr{L}(\Sigma)$ and hence $C(X_1 \circ X_2) \neq X_1 \circ X_2$. We conclude $X_1 \circ X_2$ is not self-compatible and thus not recognizable. Q.e.d.

Remark. The family $X_1 \circ X_2$ from the proof of the preceding theorem and thus also the family Fin are examples of strong families of languages (because it was shown that $C(X_1 \circ X_2) = \mathscr{L}(\Sigma)$ and $X_1 \circ X_2 \subseteq Fin$). Besides, these families are clearly countable, finitely derivable and, moreover, the family $X_1 \circ X_2$ contains only languages with a bounded number of strings (for every $L \in X_1 \circ X_2$, card $(L) \leqq$ card (Σ_A)).

In the first part of this section we present some results concerning the strong families. Knowledge about this type of families provides us with better understanding of the structure of the collection of all recognizable families.

Our first theorem gives the negative answer to the question of existence of finite strong families posed in paper [1]. The result was claimed in [4] but with an error in the proof.

3.1. Theorem. There is no finite strong family of languages.

Proof. Let us fix an arbitrary string $u \in \Sigma^*$. We shall show that the family $X_u = \{L;\ u \notin \text{Max}(L)\}$ is recognizable. Clearly, for any $v \in \Sigma^*$ it holds: if $v \leqq u$, i.e. $u = vw$ for some $w \in \Sigma^*$, then $\partial_v X_u = \{L;\ w \notin \text{Max}(L)\}$, in all other cases $\partial_v X_u = \mathscr{L}(\Sigma)$. At the same time we have:

$$L' \notin X_u \equiv u \in \text{Max}(L') \equiv \text{Fst}_\Lambda(\partial_u L') = \{\Lambda\}\,,$$
$$L \in X_u \equiv u \notin \text{Max}(L) \equiv \text{Fst}_\Lambda(\partial_u L) \neq \{\Lambda\}\,.$$

$\mathscr{D}(X_u)$ is thus finite, $C(X_u) = X_u$ and by the characterization theorem X_u is a recognizable family. Take now any $a \in \Sigma$ and $L \in \mathscr{L}(\Sigma)$. There exists at most one $n \in N$ such that $a^n \in \text{Max}(L)$. Indeed, if there were $\{a^n, a^{n+k}\} \subseteq \text{Max}(L)$ for $k > 0$, then $a^k \in \partial_{a^n} L$ which would contradict to the definition of the set Max (L). This leads to the conclusion that if X is any finite family, then by the preceding there exists at most finitely many $n \in N$ such that for some $L \in X$, $a^n \in \text{Max}(L)$. In other words there exists $m \in N$ such that $a^m \notin \text{Max}(L)$ for any $L \in X$. Thus $X \subseteq \{L;\ a^m \notin \text{Max}(L)\} = X_{a^m}$, but X_{a^m} is obviously nontrivial (e.g. $\{a^m\} \notin X_{a^m}$) and by the first part of the proof X_{a^m} is recognizable. Thus no finite family can be strong because it is contained in a nontrivial recognizable family. Q.e.d.

On the other hand the last Remark of the preceding section shows that there exist strong families, which are "only" countable. This enables us to determine more closely whole "population" of strong families.

3.2. Proposition. card $\{X;\ X$ is a strong family$\}$ = card $\{X;\ X$ is not a strong family$\} = \aleph_2$.

Proof. Every nontrivial family with a strong subfamily is by definition strong. For a countable strong family, e.g. *Fin*, clearly card $\{X;\ Fin \subseteq X\} = \aleph_2$.

On the other hand $\mathscr{L}(\Sigma) - \{\Lambda\}$ is recognizable nontrivial family thus no its subfamily can be strong; there are $\aleph_2$ such subfamilies. Q.e.d.

In the next part of this section we shall first study somewhat deeper the relations between C and R in order to prepare for the subsequent attempt to characterize all recognizable families.

As we have seen, $R(X) \subseteq C(X)$. A simple example that the inverse inclusion does not, in general, hold (not even in the case of finitely derivable X) is the set Fin for which $R(Fin) = Fin$ and $C(Fin) = \mathscr{L}(\Sigma)$ (see proof of Proposition 2.7). In [5] an additional requirement on X was formulated (namely, that X is a closed subset of a properly defined metric space of languages) which jointly with the replacement property is equivalent to self-compatibility.

The situation is much simpler in the case of finite families: as it was shown in [1] (Theorem 4.6), they are self-compatible iff they have the replacement property. As a consequence, if $R(X)$ is finite then $R(X) = C(X)$. Let us ask whether in the latter statement the assumption of finiteness of $R(X)$ can be weakened, say, by requiring only the finiteness of X.

Problem. Is it true that for any finite X, $R(X) = C(X)$? Or, more strongly, is $R(X)$ finite for every finite X?

3.3. Proposition. Let X be a finite family of finite languages, $X \subseteq Fin$. Then $R(X)$ is finite family and thus $R(X) = C(X)$ and $R(X)$ is recognizable.

Proof. From the definition of R closure it is clear that $\bigcup R(X) = \bigcup X$ (by a replacement we never obtain a string which is not already in one of the original languages). Since X is a finite family of finite languages, its union is finite and therefore $R(X)$ is necessarily finite. So by the mentioned theorem from [1] we have $R(X) = C(X)$. The recognizability of $R(X)$ follows because $R(X)$ is clearly finitely derivable. Q.e.d.

3.4. Lemma. C-closure may not preserve cardinalities, neither finite nor infinite.

Proof. For finite cardinalities it is obvious. For infinite cardinalities as we know

$$\aleph_0 = \text{card}\,(Fin) < \text{card}\,(C(Fin)) = \text{card}\,(\mathscr{L}(\Sigma)) = \aleph_1\,.$$

Q.e.d.

3.5. Lemma. $\text{card}\,(R(X)) \leqq \aleph_0 \,.\, \text{card}\,(X)$. In particular, if X is infinite then $\text{card}\,(R(X)) = \text{card}\,(X)$.

Proof. If $X = \emptyset$ or $\text{card}\,(X) = \aleph_1$ the lemma holds trivially. Let $0 \neq \text{card}\,(X) = \text{card}\,(R^0(X)) \leqq \aleph_0$. If $\text{card}\,(R^n(X)) \leqq \aleph_0$ then by the definition of $R^{n+1}(X)$ in Section 1 we can easily see the following unequality:

$$\text{card}\,(R^{n+1}(X)) \leqq \text{card}\,(\Sigma^*)\,.\,\text{card}\,(R^n(X) \times R^n(X)) \leqq \aleph_0^3 = \aleph_0\,.$$

By induction we have shown that for $n \in N$, $\text{card}\,(R^n(X)) \leqq \aleph_0$ and thus $\text{card}\,(R(X)) = \text{card}\,(\bigcup_{n \in N} R^n(X)) < \aleph_1$ because it is a countable union of at most countable sets ($\aleph_1$ is a regular cardinal). Q.e.d.

3.6. Proposition. There exists a finite family of languages such that $R(X) \neq C(X)$ and moreover card $(C(X)) = \aleph_1$.

Proof. Take a family Y of all complete languages (see e.g. [6], p. 47). In our notation we can define Y by

$$L \in Y \equiv (\forall u \in \Sigma^*) [\Sigma \subseteq \mathrm{Fst}_\Lambda (\partial_u L)] \equiv$$
$$\equiv (\forall u \in \Sigma^*) [\mathrm{Fst}_\Lambda (\partial_u L) = \Sigma \vee \mathrm{Fst}_\Lambda (\partial_u L) = \Sigma_\Lambda] .$$

Choose $L_0 \in Y$ such that $\bar{L}_0 \in Y$ (such L_0 exists, for example $L_0 = \{u; u \in \Sigma^* \,\&\, \lg (u)$ is odd). Let $X = \{L_0, \bar{L}_0\}$ and let us show that $C(X) = Y$ for any L_0 satisfying the above condition. According to the choice of L_0, for every $u \in \Sigma^*$,

$$\mathrm{Fst} (\partial_u L_0) = \mathrm{Fst} (\partial_u \bar{L}_0) = \Sigma ,$$

and besides

$$\mathrm{Fst}_\Lambda (\partial_u L_0) \neq \mathrm{Fst}_\Lambda (\partial_u \bar{L}_0) .$$

We have

$$L \in C(X) \equiv (\forall u \in \Sigma^*) [\mathrm{Fst}_\Lambda (\partial_u L) = \Sigma \vee \mathrm{Fst}_\Lambda (\partial_u L) = \Sigma_\Lambda] \equiv L \in Y .$$

Thus we have $C(X) = Y$; clearly card $(Y) = \aleph_1$ and since X is finite, card $(R(X)) \leqq \aleph_0$ by the previous lemma. It is shown that $R(X) \neq C(X)$ and therefore card $(R(X)) = \aleph_0$, because otherwise we would have $R(X) = C(X)$. Q.e.d.

Although this assertion gives negative answer to the question posed above it is interesting in itself. It has turned out that there exist pairs of languages (in fact, as one can see from the proof, these languages can be chosen to be regular) so that the smallest recognizable (a family of complete languages is obviously recognizable by one-state automaton) family which contains them is very "large", even uncountable. Thus two languages are sufficient to characterize this recognizable family. A question arises whether it is possible to characterize every recognizable family by a "small" number of languages, respectively a question whether there exists for every family of languages a smallest recognizable superfamily. Proposition 3.1 informs us that a recognizable family cannot be, in general, written as a C-closure of finite number of elements ($\mathscr{L}(\Sigma)$ itself is not characterizable in this way). The second question remains open for the case when the family X is not finitely derivable (otherwise $C(X)$ is obviously the smallest recognizable superfamily). In what follows we shall, after several preparatory definitions, state a theorem which gives a new characterization of recognizable families and we shall outline some open possibilities of its further refinement.

3.7. Definition. Let $\mathscr{R}$ be the power set of the cartesian product $\Sigma^* \times Z = \{\langle u, \Gamma \rangle; u \in \Sigma^* \,\&\, \Gamma \in Z\}$, i.e. $\mathscr{R} = 2^{\Sigma^* \times Z}$. We shall denote the elements of $\mathscr{R}$ by R. We say that $R \in \mathscr{R}$ is a regular graph if for every $\Gamma_i \in Z$ the language $L_i =$

$= \{u;\ \langle u, \Gamma_i\rangle \in \mathscr{R}\}$ is regular (here it is not necessary that $L_i \in \mathscr{L}(\Sigma)$ because L_i can be empty).

3.8. Definition. We define a pair of (total) functions: $G: 2^{\mathscr{L}(\Sigma)} \to \mathscr{R}$ and $\tilde{G}: \mathscr{R} \to \to 2^{\mathscr{L}(\Sigma)}$ by

$$G(X) = \{\langle u, \Gamma_i\rangle;\ (\exists L \in X)\, [\mathrm{Fst}_A(\partial_u L) = \Gamma_i]\}\,,$$
$$\tilde{G}(R) = \{L;\ (\forall u \in \mathrm{Pref}(L))\, [\langle u, \mathrm{Fst}_A(\partial_u L)\rangle \in R]\}\,.$$

3.9. Theorem. Let X be an arbitrary family of languages and $G, \tilde{G}$ just defined functions. Then the following conditions are equivalent:

1) X is recognizable.

2) $\tilde{G}(G(X)) = X$ and $G(X)$ is a regular graph.

3) There exists $R \in \mathscr{R}$ such that $X = \tilde{G}(R)$ and $G(X)$ is a regular graph.

In particular, the following conditions are also equivalent:

1′) X is self-compatible.

2′) $\tilde{G}(G(X)) = X$.

3′) There exists $R \in \mathscr{R}$ such that $X = \tilde{G}(R)$.

Proof. First we prove the equivalence of conditions 1′), 2′) and 3′). Clearly 2′) implies 3′) because $G(X) \in \mathscr{R}$ and thus certainly such R exists. Now for every $R \in \in \mathscr{R}$, $\tilde{G}(R)$ is self-compatible family because

$$L \in C(\tilde{G}(R)) \equiv (\forall u \in \mathrm{Pref}(L))\,(\exists L_u \in \tilde{G}(R))\,[\mathrm{Fst}_A(\partial_u L) = \mathrm{Fst}_A(\partial_u L_u)] \equiv$$
$$\equiv (\forall u \in \mathrm{Pref}(L))\,[\langle u, \mathrm{Fst}_A(\partial_u L)\rangle \in R] \equiv L \in \tilde{G}(R)\,.$$

Thus 3′) implies 1′). Certainly $X \subseteq \tilde{G}(G(X))$ and $\tilde{G}(G(X)) \subseteq C(X)$ because

$$L \notin C(X) \equiv (\exists u \in \mathrm{Pref}(L))\,(\forall L' \in X)\,[\mathrm{Fst}_A(\partial_u L) \neq \mathrm{Fst}_A(\partial_u L')] \equiv$$
$$\equiv \langle u, \mathrm{Fst}_A(\partial_u L)\rangle \notin G(X) \Rightarrow L \notin \tilde{G}(G(X))\,.$$

Thus also 1′) implies 2′) which proves the second part of the theorem.

Now it remains to show that for a self-compatible family X, $\mathscr{D}(X)$ is finite if and only if $G(X)$ is a regular graph. First let us observe that for any $v \in \Sigma^*$, any L_i defined for a family $G(X)$ in Definition 3.7 and $\Gamma_i \in Z$,

$$G(\partial_v X) = \{\langle u, \Gamma_i\rangle;\ (\exists L \in \partial_v X)\,[\mathrm{Fst}_A(\partial_u L) = \Gamma_i]\} =$$
$$= \{\langle u, \Gamma_i\rangle;\ (\exists L' \in X)\,[\mathrm{Fst}_A(\partial_{vu} L') = \Gamma_i]\} =$$
$$= \{\langle u, \Gamma_i\rangle;\ vu \in L_i\} = \{\langle u, \Gamma_i\rangle;\ u \in \partial_v L_i\}\,.$$

Furthermore, a function G partialized on self-compatible families is injective since it was proved that for self-compatible X and X' $(X \neq X')$, $\tilde{G}(G(X)) = X \neq X' = \tilde{G}(G(X'))$ and thus necessarily $G(X) \neq G(X')$. Using just obtained equations and the fact that the derivative preserves the self-compatibility of families we have for a self-compatible family X:

$$\partial_v X = \partial_w X \equiv G(\partial_v X) = G(\partial_w X) \equiv (\forall_i)\ [\partial_v L_i = \partial_w L_i]\ .$$

Since i are indices of a finite set $(\Gamma_i \in Z$ where Z is finite$)$ X is finitely derivable iff $G(X)$ is a regular graph, which concludes the proof. Q.e.d.

The above theorem contains in fact several results which are interesting in varying degree. As the most interesting we consider the fact that every recognizable family of languages can be characterized with the help of a finite – and even bounded since it depends only on the cardinality of the alphabet – number of regular languages L_i (determined by $G(X)$) together with corresponding Γ_i. Recognizable family may be thus expressed either as a behaviour of a branching automaton or as $\tilde{G}(R)$ where R is a regular graph. But since the function G is not surjective the inverse assertion does not hold: the system R defined by a finite number of regular languages determines self-compatible, but not necessarily recognizable family (otherwise the recognizability by means of the branching automata could be easily reduced to the recognizability by means of the classical automata). As a possible topic for further investigation one can consider the search for additional conditions which must be fulfilled by the regular languages (with corresponding Γ_i) in order that for the system R defined by them the relation $R = G(\tilde{G}(R))$ would hold, i.e. they would canonically define recognizable family (necessary but not sufficient condition is that the language $\bigcup L_i$ contains all its prefixes).

We shall now present two simple corollaries and conclude the section by proposition concerning finitely derivable families of languages.

3.10. Corollary. For any family of languages X the following equations hold:

a) $C(X) = \bigcup\{X';\ G(X) = G(X')\}$,

b) $G(X) = \bigcap\{R;\ \tilde{G}(R) = C(X)\}$, where $G, \tilde{G}$ are functions defined in 3.8.

Hint. It is enough to realize that functions G, $\tilde{G}$ preserve the set-theoretical inclusion and to use the facts that G partialized on $\{X;\ C(X) = X\}$ and $\tilde{G}$ partialized on the range of G are one-one functions ($\tilde{G}$ is here inverse function to G).

3.11. Corollary. card $(\{X;\ C(X) = X\}) = \aleph_1$.

Proof. By the proof of Theorem 3.9, G is injective mapping of $\{X;\ C(X) = X\}$ into $\mathscr{R}$, thus necessarily

$$\text{card}\,(\{X;\ C(X) = X\}) \leqq \text{card}\,(\mathscr{R}) = \aleph_1\ .$$

The equality follows from the fact that every singleton is self-compatible and that card $(\mathscr{L}(\Sigma)) = \aleph_1$.

3.12. Proposition. card $(\{X;\ X$ is finitely derivable$\}) = \aleph_2$.

Proof. If we denote $Y = \{L;\ \Lambda \notin L\}$ then clearly card $(\{X;\ X \subseteq Y\}) = \aleph_2$. For every $X \subseteq Y$ we define $X' = X \cup (\mathscr{L}(\Sigma) - Y)$. If $X_1, X_2 \subseteq Y$ then immediately

$$X_1 \neq X_2 \Rightarrow X_1' \neq X_2'$$

thus also card $\{X';\ X \subseteq Y\} = \aleph_2$. Moreover, every family X' is finitely derivable because $\partial_u X' = \mathscr{L}(\Sigma)$ for $u \in \Sigma^+$. Q.e.d.

Surprisingly enough, a relatively "small" (countable) system of recognizable families is obtained by intersecting the set of all self-compatible families and that of all finitely derivable families, which are rather "large" of cardinalities, $\aleph_1$ and $\aleph_2$ respectively.

4. ALGEBRAIC DECOMPOSITION OF RECOGNIZABLE FAMILIES OF LANGUAGES

While in the foregoing section we have aimed at characterizing recognizable families in terms of regular languages of classical automata theory, we shall now try to characterize them according to their "algebraic" structure. Since the results of this section will be used in a considerable degree later for a special type of recognizable families, the theorem about decomposition is given here in two alternative forms. Correspondingly, some concepts introduced here are justified only by needs of the following Section 5.

4.1. Definition. If X' and X'' are families of languages then the family $X' + X''$ defined as

$$X' + X'' = \{L' \cup L'';\ L' \in X' \;\&\; L'' \in X''\}$$

is called *inner union* of families X' and X''.

Since this operation is clearly associative the definition can be extended as follows:

$$\sum_{i=1}^{n} X_i = X_1 + \ldots + X_n \ \left(\text{in particular, } \sum_{i=1}^{1} X_i = X_1\right).$$

4.2. Remark. Later we shall see that this operation does not preserve recognizability.

Convention. For formal reasons we shall occasionally treat the empty string as one of the letters of an alphabet. Thus we shall in the sequel use lower-case Greek letters

$\alpha, \beta, \ldots$ as variables for elements of Σ_Λ, while letters $a, b, \ldots$ of Latin alphabet will be used only for elements of Σ.

4.3. Definition. For $u \in \Sigma^+$ the family $\pi_u X$ defined as

$$\pi_u X = u\,\partial_u X = \{u\,\partial_u L;\ L \in X\} \setminus \{\emptyset\}$$

is called the *projection of X w.r.t. u.* Again for formal reasons it is useful to define projection for $u = \Lambda$ in a different way:

$$\begin{aligned} \pi_\Lambda X &= \{\{\Lambda\}\} \quad \text{if} \quad \Lambda \in \bigcup X\,, \\ &= \emptyset \qquad\quad \text{otherwise}\,. \end{aligned}$$

4.4. Definition. For $\Gamma \in Z(= \{L;\ \emptyset \neq L \subseteq \Sigma_\Lambda\})$ we define

$$X_\Gamma = \{L;\ \mathrm{Fst}_\Lambda(L) = \Gamma\}\,.$$

4.5. Facts. a) $X_{\{\Lambda\}} = \{\{\Lambda\}\}$;

b) $X_{\{a\}} = a\,\mathscr{L}(\Sigma)$;

c) X_Γ is uncountable for $\Gamma \neq \{\Lambda\}$;

d) $\pi_\alpha X_\Gamma = X_{\{\alpha\}}$ if $\alpha \in \Gamma$;

$\pi_\alpha X_\Gamma = \emptyset$ if $\alpha \notin \Gamma$;

e) $X_\Gamma = \sum_{\alpha_i \in \Gamma} X_{\{\alpha_i\}}$;

f) in a more general way:

$$\text{if} \quad \bigcup_{i=1}^{n} \Gamma_i = \Gamma \quad \text{then} \quad X_\Gamma = \sum_{i=1}^{n} X_{\Gamma_i}\,;$$

g) for a singleton: $\sum_{\alpha \in \mathrm{Fst}_\Lambda(L)} \pi_\alpha\{L\} = \{L\}$.

4.6. Lemma. For every $\Gamma \in Z$, X_Γ is a recognizable family.

Proof. X_Γ is finitely derivable because, by Facts 4.5 b), d):

$$\begin{aligned} \partial_{au} X_\Gamma &= \partial_u(\partial_a X_{\{a\}}) = \partial_u \mathscr{L}(\Sigma) = \mathscr{L}(\Sigma) \quad \text{if} \quad a \in \Gamma\,, \\ &= \partial_u(\partial_a X_\Gamma) = \emptyset \quad \text{if} \quad a \notin \Gamma\,. \end{aligned}$$

Besides, $C(X_\Gamma) = X_\Gamma$ because

$$L \in C(X_\Gamma) \equiv (\exists L_\Lambda \in X_\Gamma)\,[\mathrm{Fst}_\Lambda(L) = \mathrm{Fst}_\Lambda(L_\Lambda) = \Gamma)] \equiv L \in X_\Gamma\,.$$

Thus by the characterization theorem, X_Γ is recognizable. Q.e.d

4.7. Lemma. Let X be a recognizable family of languages. Then for an arbitrary $\alpha \in \Sigma_\Lambda$ and $\Gamma \in Z$, the following holds:

$$\pi_\alpha(X \cap X_\Gamma) \neq \emptyset \Rightarrow \pi_\alpha(X \cap X_\Gamma) = \pi_\alpha X .$$

Proof. Certainly $\pi_\alpha(X \cap X_\Gamma) \subseteq \pi_\alpha X \subseteq \pi_\alpha(\mathscr{L}(\Sigma))$. Thus for $\alpha = \Lambda$ we have:

$$\pi_\alpha(X \cap X_\Gamma) \neq \emptyset \Rightarrow \pi_\alpha(X \cap X_\Gamma) = \pi_\alpha X = \pi_\alpha(\mathscr{L}(\Sigma)) = \{\{\Lambda\}\} .$$

For $\alpha = a$ let us assume, for the contradiction, that there exists $aL_1 \in \pi_a X \setminus \setminus \pi_a(X \cap X_\Gamma)$, where $\pi_a(X \cap X_\Gamma) \neq \emptyset$. By the definition of a projection this implies that there exist $L \in X$ and $L' \in (X \cap X_\Gamma)$ of the form: $L = aL_1 \cup L_2$, $L' = aL'_1 \cup L'_2$ where $a \notin \text{Pref}(L_2)$ and $a \notin \text{Pref}(L'_2)$. Since X is by the assumption recognizable it has the replacement property and thus the language $L_0 = R_a(L', \partial_a L) = aL_1 \cup L'_2$ belongs to X and also to $X \cap X_\Gamma$ because $\text{Fst}_\Lambda(aL_1 \cup L'_2) = \text{Fst}_\Lambda L'$. At the same time $\partial_a L_0 = L_1$, i.e. $aL_1 \in \pi_a(X \cap X_\Gamma)$ which contradicts the assumption. Q.e.d.

4.8. Lemma. For any family X and any $u \in \Sigma^*$, X is recognizable iff uX is recognizable.

Proof. If uX is recognizable then $\partial_u(uX) = X$ is also recognizable — as we have already mentioned in Section 1. It remains to show the opposite implication. Without a loss of generality it is enough to show that for $a \in \Sigma$ and a recognizable X, aX is also recognizable. The finite derivability of aX is clear, the self-compatibility follows from the following equivalences (here we use the equation $C(\partial_u X) = \partial_u C(X)$ from Lemma 5.1 of [1] and the assumption that $C(X) = X$):

$$L \in C(aX) \equiv \text{Fst}_\Lambda(\partial_\Lambda L) = a \,\&\, \partial_a L \in \partial_a C(aX) \equiv$$
$$\equiv \text{Fst}_\Lambda(L) = a \,\&\, \partial_a L \in X \equiv L = aL' \,\&\, L' \in X \equiv L \in aX .$$

Q.e.d.

4.9. Theorem. A family of languages X is recognizable iff the following two conditions are fulfilled:

1) For every $a \in \Sigma$, $\pi_a X$ is recognizable.

2) For every $\Gamma \in Z$, if $X \cap X_\Gamma \neq \emptyset$ then $X \cap X_\Gamma = \sum_{\alpha_i \in \Gamma} \pi_{\alpha_i} X$.

Proof. If X is recognizable then by Assertions in Section 1 and by Lemma 4.8 also $\pi_a X = a\, \partial_a X$ is recognizable; thus the first condition is fulfilled.

Now we prove condition 2): Supposing $X \cap X_\Gamma \neq \emptyset$ then by Lemma 4.7

$$\sum_{\alpha_i \in \Gamma} \pi_{\alpha_i} X = \sum_{\alpha_i \in \Gamma} \pi_{\alpha_i}(X \cap X_\Gamma) .$$

It remains to show $X \cap X_\Gamma = \sum_{\alpha_i \in \Gamma} \pi_{\alpha_i}(X \cap X_\Gamma)$. Let $L \in \mathscr{L}(\Sigma)$, clearly we can restrict ourselves only to L with $\mathrm{Fst}_\Lambda(L) = \Gamma$. If $L \notin \sum_{\alpha_i \in \Gamma} \pi_{\alpha_i}(X \cap X_\Gamma)$ then there exists $a \in \Gamma$ such that

$$\partial_a L \notin \partial_a \sum_{\alpha_i \in \Gamma} \pi_{\alpha_i}(X \cap X_\Gamma) = \partial_a \pi_a(X \cap X_\Gamma) = \partial_a(X \cap X_\Gamma)$$

and thus certainly $L \notin X \cap X_\Gamma$. If on the other hand $L \in \sum_{\alpha_i \in \Gamma} \pi_{\alpha_i}(X \cap X_\Gamma)$, then for every $a \in \Gamma$ we have $\partial_a L = \partial_a(L_a)$ for some $L_a \in X \cap X_\Gamma$. Let now $u \in \mathrm{Pref}(L)$. Then u is of the form av for some $a \in \Gamma$ (the case $u = \Lambda$ was already treated) and we have: $\mathrm{Fst}_\Lambda(\partial_{av}L) = \mathrm{Fst}_\Lambda(\partial_v \partial_a L) = \mathrm{Fst}_\Lambda(\partial_v \partial_a L_a) = \mathrm{Fst}_\Lambda(\partial_{av} L_a)$. Thus $L \in C(X \cap X_\Gamma)$, and since X is recognizable, X_Γ is recognizable by Lemma 4.6 and thus also the intersection $X \cap X_\Gamma$ is recognizable (Fact 2.1a)). Thus $C(X \cap X_\Gamma) = X \cap X_\Gamma$. This proves the first half of the theorem.

Assume now that conditions 1) and 2) are satisfied; we shall prove that X is recognizable. From the definition of projection it is clear that condition 1) implies finite derivability of X. It remains to show that $C(X) = X$.

Let $L \in C(X)$. Then for every $a \in \mathrm{Pref}(L)$ we have $\partial_a L \in \partial_a C(X)$. But by Lemma 5.1 of [1], $\partial_a C(X) = C(\partial_a X)$. By assumption $\pi_a X = a\, \partial_a X$ is recognizable, thus by Lemma 4.8, $\partial_a X$ is also recognizable and hence $\partial_a L \in \partial_a X$. Let $\mathrm{Fst}_\Lambda(L) = \Gamma_0$. By the assumption that $L \in C(X)$ there exists $L_\Lambda \in X$ such that also $\mathrm{Fst}_\Lambda(L_\Lambda) = \Gamma_0$, in other words $X \cap X_{\Gamma_0} \neq \emptyset$. Now we can use condition 2) to obtain:

$$X \cap X_{\Gamma_0} = \sum_{\alpha_i \in \Gamma_0} \pi_{\alpha_i} X = \begin{cases} \sum_{a \in \Gamma_0} a\, \partial a X & \text{if} \quad \Lambda \notin \Gamma_0 \\ \{\{\Lambda\}\} + \sum_{a \in \Gamma_0} a\, \partial_a X & \text{if} \quad \Lambda \in \Gamma_0 . \end{cases}$$

Since we have $\partial_a L \in \partial_a X$, by Fact 4.5g) we have

$$\{L\} = \sum_{a \in \Gamma_0} a\, \partial_a \{L\} \subseteq \sum_{a \in \Gamma_0} a\, \partial_a X \quad \text{if} \quad \Lambda \notin \Gamma_0 ;$$

$$\{L\} = \{\{\Lambda\}\} + \sum_{a \in \Gamma_0} a\, \partial_a \{L\} \subseteq \{\{\Lambda\}\} + \sum_{a \in \Gamma_0} a\, \partial_a X \quad \text{if} \quad \Lambda \in \Gamma_0 .$$

Therefore $L \in X \cap X_{\Gamma_0}$ and thus $C(X) = X$ which completes the whole proof.

Q.e.d.

The following decomposition theorem for recognizable families is a reformulation of Theorem 4.9 and immediately follows from it.

4.10. Theorem. (The decomposition of recognizable families of languages.) Let Σ be a finite alphabet, $\mathrm{card}(\Sigma) = n$. Put $\alpha_0 = \Lambda$, $Y_0 = \{\{\Lambda\}\}$, $\alpha_i \neq \alpha_j$ for $i \neq j$. Then $X \subseteq \mathscr{L}(\Sigma)$ is recognizable if and only if there exist recognizable families $Y_1, \ldots, Y_n$ and a family $K \subseteq Z$ such that

$$X = \bigcup_{\Gamma \in K} \sum_{\alpha_i \in \Gamma} \alpha_i Y_i .$$

The decomposition theorem is quite important: it enables to construct new recognizable family of languages from previously given families and among others it implies that the union preserves recognizability only in exceptional cases. The theorem is also interesting since it provides a characterization of families according to their decomposition with respect to derivatives, which is analogous to the well known decomposition of regular languages in the classical automata theory (see, e.g. [8], Section 4.5). Since Y_i in Theorem 4.10 are recognizable and card $(\mathscr{D}(Y_i)) \leqq$ card $(\mathscr{D}(X))$, it is evident that by recurrent application of the theorem we could describe every recognizable family by a finite system of characteristic equations (similarly as the regular language in the classical theory). This possibility will not, however, be exploited further and in the next section attention will be devoted to a certain special case.

5. WELL-RECOGNIZABLE FAMILIES OF LANGUAGES

In Section 1 and in [1] it has been demonstrated that unlike to the classical automata theory a system of recognizable families is not closed under the basic set-theoretical operations (except for the intersection). An important natural interpretation of these operations has already been pointed out in the Introduction and in Section 1. In this section we shall thus confine ourselves to a study of a smaller collection of families, namely the well-recognizable families (recognizable families with recognizable complement). We shall investigate some of its properties and consider the question whether there exists an important subcollection of the recognizable families closed under the Boolean operations (intersection, union, complement). Our main objective, however, will be to characterize the well-recognizable families, analogously as in the decomposition theorem in the foregoing section. We shall see that this in a certain sense large collection is interesting, among others, by the fact that the characterization can be made much more explicit than in the general case. This will appear to be a consequence of the fact that in the restricted case of well-recognizable families Theorem 4.9 obtains a stronger form. Since the relation $(X \cap X_\Gamma \neq \emptyset) \vee (\overline{X} \cap X_\Gamma \neq \emptyset)$ certainly holds, then either the family or its complement must fulfil the consequent of the implication from condition 2).

5.1. Definition. Family of languages X is *well-recognizable* if both X and $\overline{X}$ are recognizable families.

The following are examples of well-recognizable families.

Example 1. Trivial families ($\emptyset$ and $\mathscr{L}(\Sigma)$) are well-recognizable as we already know.

Example 2. For any $K \subseteq Z$ the family $\bigcup_{\Gamma_i \in K} X_{\Gamma_i}$ is well-recognizable.

Proof. Clearly $\overline{X} = \bigcup_{\Gamma_j \in (Z-K)} X_{\Gamma_j}$. By Lemma 4.6 for every Γ, the family X_Γ is
recognizable. Since Z is finite, X and $\overline{X}$ are by Fact 2.1 c) finitely derivable. Also $C(X) = X$ because clearly $L \in C(X) \equiv \mathrm{Fst}_\Lambda(L) \in K \equiv L \in X$. Analogously $C(\overline{X}) = \overline{X}$.

Q.e.d.

Results of this section will show that it is rather difficult to present other simple examples of well-recognizable families. The following example suggests that the class of well-recognizable families is countable.

Example 3. The family $\{L; u \in \mathrm{Max}(L)\}$ (and thus also the family $\{L; u \notin \mathrm{Max}(L)\}$) is well-recognizable for every $u \in \Sigma^*$ (the proof will be given later).

5.2. Lemma. Let X and $\overline{X}$ be self-compatible. Then for every $u \in \Sigma^*$, $\partial_u X \cap \partial_u \overline{X}$ is a trivial family. In fact, the same conclusion holds already if both X and $\overline{X}$ have the replacement property.

Proof. For $u = \Lambda$ the assertion is immediate.

Let us assume, for contradiction, that there exists a string $u \in \Sigma^+$ such that $\partial_u X \cap \partial_u \overline{X}$ is nontrivial. The intersection is nonempty and thus there exist L_0, L_1, L_2 such that*): $uL_0 \cup L_1 \in X$, $uL_0 \cup L_2 \in \overline{X}$ and $u \notin \mathrm{Pref}(L_1)$, $u \notin \mathrm{Pref}(L_2)$. The intersection is different from $\mathscr{L}(\Sigma)$, thus there exists L' such that $L' \notin \partial_u X \cap \partial_u \overline{X}$; by symmetry we may assume, say, $L' \notin \partial_u \overline{X}$. Thus $uL' \cup L_2 \in X$ because $\partial_u(uL' \cup L_2) = L' \notin \partial_u \overline{X}$. But X has by the assumption the replacement property and therefore $R_u((uL' \cup L_2), \partial_u(uL_0 \cup L_1)) = uL_0 \cup L_2 \in X$ which yields a contradiction.

Q.e.d.

5.3. Proposition. Let X be a well-recognizable family ,$u \in \Sigma^*$. Then $\partial_u X$ is well-recognizable.

Proof. If $\partial_u X = \mathscr{L}(\Sigma)$ there is nothing to prove.

Let $\partial_u X \neq \mathscr{L}(\Sigma)$; then by Lemma 5.2, $\partial_u X \cap \partial_u \overline{X} = \emptyset$. So $\overline{\partial_u X} = \partial_u \overline{X}$ and thus $\partial_u X$ is well-recognizable because the derivative preserves recognizability of families.

Q.e.d.

5.4. Theorem. X is well recognizable iff both X and $\overline{X}$ are self-compatible.

Proof. We need to show that X and $\overline{X}$ are finitely derivable. From Lemma 5.2 it immediately follows that $\partial_u X$ is nontrivial iff $\partial_u \overline{X}$ is nontrivial and that $\mathrm{card}(\mathscr{D}(X)) = \mathrm{card}(\mathscr{D}(\overline{X}))$. It is enough to show finite derivability of X. Assume the contrary. Then $\tilde{L} = \{u;\ \partial_u X \text{ is nontrivial}\}$ is an infinite language. For any $v, w \in \Sigma^*$, if $v < w$

*) In this and some subsequent constructions we allow ourselves a slight deviation from rigorous presentation by admitting one of the languages L_1 and L_2 to be empty (and thus not in $\mathscr{L}(\Sigma)$). This, however, cannot yield an error in proofs since we shall deal always with languages of the form $uL_0 \cup L_1$ and $uL_0 \cup L_2$ which are non-empty due to assumption that $L_0 \in \mathscr{L}(\Sigma)$.

and $\partial_v X$ is trivial then necessarily $\partial_w X$ is also trivial. Thus $\text{Pref}(\tilde{L}) = \tilde{L}$ and we can interpret $\tilde{L}$ as a tree. By the König theorem any infinite tree with finite amount of branches at any node must contain an infinite path, thus there exist $L \subseteq \tilde{L}$ of the form: $L = \{u_i\}_{i=0}^{\infty}$ where $\lg(u_i) = i$ and $u_i < u_j \equiv i < j$. For $u \in L$, $\partial_u X$ (and also $\partial_u \bar{X}$) is nontrivial and thus we can, for every $n \in N$, choose $L_n \in \partial_{u_n} X$, $L'_n \in \partial_{u_n} \bar{X}$. Since by Lemma 5.2 for every $n \in N$, $\partial_{u_n} X \cap \partial_{u_n} \bar{X} = \emptyset$ clearly $u_n L_n \cup \{u_{n-1}\} \in X$, $u_n L'_n \cup \{u_{n-1}\} \in \bar{X}$. But for every $u_i \in \text{Pref}(L) = L$ we have

$$\text{Fst}_A(\partial_{u_i} L) = \text{Fst}_A(\partial_{u_i}(u_{i+1} L_{i+1} \cup \{u_i\})) = \text{Fst}_A(\partial_{u_i}(u_{i+1} L'_{i+1} \cup \{u_i\})) .$$

Thus $L \in C(X)$ as well as $L \in C(\bar{X})$, in contradiction to the self-compatibility of X and $\bar{X}$. Q.e.d.

In fact, we have proved a somewhat stronger result:

5.5. Corollary. Let X be well-recognizable family of languages. Then $\{u; \partial_u X$ is nontrivial$\}$ is finite. (Cf. proof of the last theorem.)

Theorem 5.4 gives a possible characterization of well-recognizable families and, indirectly, throws a new light on the relationship between self-compatibility and finite derivability. The next theorem shows that under certain circumstances we can further weaken the assumption of Theorem 5.4.

5.6. Theorem. Let X be self-compatible and $\bigcup \bar{X} \neq \Sigma^*$. Then X is well-recognizable.

Proof. It is enough to show that $C(\bar{X}) = \bar{X}$ and then use Theorem 5.4. To obtain a contradiction let us assume that there exists $L \in X$ such that $L \in C(\bar{X})$. Take v such that $v \notin \bigcup \bar{X}$ and put $w = \text{Max}(\text{Pref}(\{v\}) \cap \text{Pref}(L))$ (w is well-defined because the intersection is finite, non-empty and linearly ordered). Now, $w \in \text{Pref}(L)$ and thus there exists $L_w \in \bar{X}$ such that $\text{Fst}_A(\partial_w L_w) = \text{Fst}_A(\partial_w L)$. However, at the same time we have, for any $u \nleq v$ or $u < w$:

$$\text{Fst}_A(\partial_u L_w) = \text{Fst}_A(\partial_u(L_w \cup \{v\})) .$$

Here $(L_w \cup \{v\}) \in X$ (because $v \notin \bigcup \bar{X}$) and we have chosen w such that if $w < u \leqq v$ then $u \notin \text{Pref}(L)$. Thus $L \in X$ implies $L_w \in C(X) = X$ which is a contradiction. Q.e.d.

5.7. Notation. Let us denote by $\mathscr{M}(\Sigma)$ the class of families obtained as a Boolean closure of families $\{X_\Gamma; \Gamma \in Z\}$. Again we shall mostly write only $\mathscr{M}$ and we shall use M for denoting elements of $\mathscr{M}$.

5.8. Remark. Since the generators of $\mathscr{M}$ are clearly disjoint, every $M \in \mathscr{M}$ is of the form $M = \bigcup_{\Gamma_i \in K} X_{\Gamma_i}$ where K is any subset of Z (possibly empty). It is easy to show that

$$\text{card}(\mathscr{M}(\Sigma)) = 2^{2^{n+1}-1}$$

where $n = \text{card}(\Sigma)$. By Facts 4.5, $\mathscr{M}$ is closed also under the operation of inner union $(+)$ from Definition 4.1.

5.9. Theorem. Every $M \in \mathscr{M}$ is well-recognizable.

Proof was given in Example 2 at the beginning of this section.

Thus we have found a (finite) class $\mathscr{M}$ of recognizable families of languages, closed under all the Boolean operations (union, intersection and complement). One can ask whether there exists any "larger" class with the same properties (for reasons to be discussed at the end of this section we can restrict ourselves to classes containing $\mathscr{M}$). The following theorem provides a negative answer and, incidentally, again demonstrates the "intolerance" of union with respect to recognizability.

5.10. Proposition. No class of recognizable families of languages containing $\mathscr{M}$ as a proper subclass is closed under union. In other words: For every recognizable family X which is not in $\mathscr{M}$ there exists $M \in \mathscr{M}$ such that $X \cup M$ is not recognizable.

Proof. Let X be recognizable and $X \notin \mathscr{M}$, i.e. there exists $\Gamma \in Z$ such that $\emptyset \neq X \cap X_\Gamma \neq X_\Gamma$. By Theorem 4.10 and Facts 4.5 the following holds:

$$X \cap X_\Gamma = \sum_{\alpha_i \in \Gamma} \pi_{\alpha_i} X \neq \sum_{\alpha_i \in \Gamma} X_{\{\alpha_i\}} = X_\Gamma$$

and thus there exists $a \in \Gamma$ such that $\pi_a X \neq X_{\{a\}}$. It is enough to choose any $\Gamma' \neq \Gamma$ for which $a \in \Gamma'$ and to put $M = X_{\Gamma'}$. Then $\pi_a(X \cup X_{\Gamma'}) = X_{\{a\}}$ and at the same time

$$\emptyset \neq \pi_a((X \cup X_{\Gamma'}) \cap X_\Gamma) = \pi_a(X \cap X_\Gamma) = \pi_a X .$$

Thus by Lemma 4.7, $X \cup X_{\Gamma'}$ is not recognizable. Q.e.d.

Problem. For every finite alphabet Σ we constructed a finite class $\mathscr{M}(\Sigma)$ of well-recognizable families of languages. One can now ask whether there exist some other well-recognizable families and whether we are able to specify them in a uniform manner. In Example 3 at the beginning of this section we have answered the first part of the question positively. Concerning the second part of the question the next theorem gives a partial answer, i.e. necessary condition on the form of well-recognizable families.

5.11. Theorem. Let X be a well-recognizable family, $X \notin \mathscr{M}(\Sigma)$. Then there exists exactly one $a \in \Sigma$ such that:

1) $X \cap X_{\{a\}} \neq \emptyset \;\&\; \overline{X} \cap X_{\{a\}} \neq \emptyset$.

And moreover, for every $\Gamma \in Z$,

2) $(X \cap X_\Gamma \neq \emptyset \;\&\; \overline{X} \cap X_\Gamma \neq \emptyset) \equiv a \in \Gamma$.

Proof. a) *Existence.* In the proof of the preceding proposition we have shown the existence of $a \in \Sigma$ for which $\emptyset \neq \pi_a X \neq X_{\{a\}}$. Using the definition of a projection (Definition 4.3) this relation can be written in the form: $\emptyset \neq a\, \partial_a X \neq a\, \mathscr{L}(\Sigma)$. In other words, $\emptyset \neq \partial_a X \neq \mathscr{L}(\Sigma)$ and thus also $\emptyset \neq \partial_a \overline{X} \neq \mathscr{L}(\Sigma)$ because by Lemma 5.2, $\partial_a X \cap \partial_a \overline{X} = \emptyset$. Since $\partial_a X_{\{a\}} = \mathscr{L}(\Sigma)$ neither $X_{\{a\}} \subseteq X$ nor $X_{\{a\}} \subseteq \overline{X}$ which we wanted to prove.

b) *Unicity.* Let there exist $a, b \in \Sigma$, $a \neq b$, satisfying condition 1) of the theorem. In the first part of this proof we obtained: $\emptyset \neq \partial_a X \neq \mathscr{L}(\Sigma)$, $\emptyset \neq \partial_b \overline{X} \neq \mathscr{L}(\Sigma)$. Take any $L_1 \in \partial_a X$, $L_2 \in \partial_b \overline{X}$ and put $L = aL_1 \cup bL_2$. If $L \in X$ then $L_2 \in \partial_b X$ and if $L \in \overline{X}$ then $L_1 \in \partial_a \overline{X}$. However, both cases yield a contradiction because by Lemma 5.2,

$$\partial_a X \cap \partial_a \overline{X} = \partial_b X \cap \partial_b \overline{X} = \emptyset$$

(clearly different from $\mathscr{L}(\Sigma)$).

c) *Proof of condition* 2). Let a be a letter satisfying condition 1). Let there exist Γ such that, e.g. (by symmetry) $X_\Gamma \subseteq X$. Then by Fact 4.5d), for every $\alpha \in \Gamma$, $\pi_\alpha X = X_{\{\alpha\}}$ and hence (see part a) of this proof) $a \notin \Gamma$.

On the contrary, let $X \cap X_\Gamma \neq \emptyset$ and $\overline{X} \cap X_\Gamma \neq \emptyset$. Then by part a) of this proof there exists $b \in \Gamma$ satisfying condition 1). By part b) there is at most one such letter, thus $a = b$, $a \in \Gamma$. Q.e.d.

The following theorem leads us already to the algebraic characterization of all well-recognizable families. First we shall need two easy lemmas.

5.12. Lemma. Let X be a well-recognizable family, $X \notin \mathscr{M}$. Then there is at most one $a \in \Sigma$ for which $\partial_a X \neq \mathscr{L}(\Sigma)$.

Proof. By Theorem 5.11 there exists only one a such that $X \cap X_{\{a\}} \neq \emptyset$ and $\overline{X} \cap X_{\{a\}} \neq \emptyset$. We prove that for $b \neq a$, $\partial_b X = \mathscr{L}(\Sigma)$. Condition 2) of Theorem 5.11 guarantees $X \cap X_{\{a,b\}} \neq \emptyset$. Thus $\partial_b X \neq \emptyset$ and either $X_{\{b\}} \subseteq X$ and then $\partial_b X = \mathscr{L}(\Sigma)$, or $X_{\{b\}} \subseteq \overline{X}$ and then $\partial_b X \cap \partial_b \overline{X} \neq \emptyset$ and again $\partial_b X = \mathscr{L}(\Sigma)$ (Lemma 5.2). Q.e.d.

5.13. Lemma. Let X be a well-recognizable family and $v, w \in \Sigma^*$ such that $v \parallel w$. Then at least one of the families $\partial_v X$ and $\partial_w X$ is trivial.

Proof. $v \parallel w$ iff there exist $a, b \in \Sigma$, $a \neq b$, and $u, v', w' \in \Sigma^*$ such that $v = uav'$ and $w = ubw'$. By Theorem 5.3, $\partial_u X$ is well-recognizable. If $\partial_u X \in \mathscr{M}$ then certainly $\partial_v X$ and also $\partial_w X$ are trivial families. If $\partial_u X \notin \mathscr{M}$ then by Lemma 5.12, $\partial_a \partial_u X = \mathscr{L}(\Sigma)$ or $\partial_b \partial_u X = \mathscr{L}(\Sigma)$ and thus also $\partial_v X = \mathscr{L}(\Sigma)$ or $\partial_w X = \mathscr{L}(\Sigma)$. Q.e.d.

5.14. Theorem. Let X be a nontrivial well-recognizable family. Then there exists exactly one $u \in \Sigma^*$ such that for all $v \in \Sigma^*$, $\partial_v X$ is nontrivial iff $v \leqq u$.

Proof. We use an induction on number n of such v, for which $\partial_v X$ is nontrivial. By Corollary 5.5 there are only finitely many of such v. By the assumption that X is nontrivial, $\partial_\Lambda X = X$ is nontrivial and thus $n \in N$.

Basis. Let $n = 1$. Then clearly $u = \Lambda$ because $\partial_\Lambda X = X$ is nontrivial and thus $\partial_v X$ is nontrivial iff $v \leqq \Lambda$.

Induction Step. Assume that the theorem is true for $n \geqq 1$ and let X have a nontrivial derivative by $n + 1$ strings. Then surely $X \notin \mathscr{M}$ because the only string by which $M \in \mathscr{M}$ can have a nontrivial derivative is Λ. Thus by the proof of Theorem 5.11 there exists $a \in \Sigma$ such that $\partial_a X$ is nontrivial and by Lemma 5.13, $\partial_v X$ is trivial for every $v \parallel a$. Thus $\partial_a X$ has a nontrivial derivative by n strings (it is well-recognizable by Proposition 5.3) and by the inductive assumption there exists u' such that $\partial_v(\partial_a X)$ is nontrivial iff $v \leqq u'$. Thus clearly $u = au'$ satisfies the requirement of the theorem, which concludes the proof. Q.e.d.

5.15. Definition. Let X be a nontrivial well-recognizable family. The unique $u \in \Sigma^*$ whose existence was established by Theorem 5.14 will be called the characteristic string of X and will be denoted by u_X. Its length $\lg(u_X)$ will be called the degree of complexity of X.

5.16. Corollary. $u_X = u_{\overline{X}}$.

Proof. By Lemma 5.2, $\partial_v X$ is nontrivial iff $\partial_v \overline{X}$ is nontrivial.

5.17. Corollary. Let X be a nontrivial well-recognizable family, $v \leqq u_X$. Then $\lg(u_{\partial_v X}) = \lg(u_X) - \lg(v)$.

Proof. By construction in Theorem 5.14, $u_X = v u_{\partial_v X}$.

5.18. Corollary. Let X be a nontrivial well-recognizable family. Then a) $X \in \mathscr{M}$ iff $u_X = \Lambda$;

b) $\partial_v X \notin \mathscr{M}$ iff $v < u_X$;

c) if $v < w \leqq u_X$ then $\partial_v X \neq \partial_w X$.

Proof. *Ad a)* Note that $\lg(u_X) + 1$ is equal to the number of nontrivial derivatives and use proof of Theorem 5.14.

Ad b) For $v \nleqq u_X$, $\partial_v X$ is trivial and thus certainly $\partial_v X \in \mathscr{M}$. For $v \leqq u_X$, we have by a) that $\partial_v X \in \mathscr{M}$ iff the degree of complexity of $\partial_v X$ is zero, which is, by Corollary 5.17, iff $v = u_X$.

Ad c) By 5.17, $\partial_v X$ and $\partial_w X$ have different degree of complexity and thus by Theorem 5.14, $\partial_v X \neq \partial_w X$. Q.e.d.

5.19. Corollary. Let X be a nontrivial well-recognizable family over Σ where $\text{card}(\Sigma) \geqq 2$ and let $X \neq \{\{\Lambda\}\}$. Then X is uncountable.

Proof. For $X \in \mathscr{M}$ see Fact 4.5c); for $X \notin \mathscr{M}$ the result immediately follows from Lemma 5.12 (if there exist $a, b \in \Sigma$, $a \neq b$).

5.20. Corollary. Let X be a nontrivial well-recognizable family. Then $\lg(u_X) + 2 \leqq \text{card}(\mathscr{D}(X)) \leqq \lg(u_X) + 3$.

Proof. Clear from Theorem 5.14 and Corollary 5.18c).

Since the cardinality of $\mathscr{D}(X)$ gives the lower bound on the number of states of a branching automaton recognizing X, the last results show the close relationship between the degree of complexity of a well-recognizable family as defined and the complexity of a corresponding branching automaton.

5.21. Definition. For $a \in \Sigma$, $M \in \mathscr{M}$ such that $\pi_a M = \emptyset$ we define a unary operation $\langle a, M\rangle$ on families of languages by:

$$\langle a, M\rangle X = aX \cup \bigcup\{aX + X_\Gamma;\ \Gamma \in Z \,\&\, a \notin \Gamma\} \cup M\,.$$

We shall write:

$$\langle a_1, M_1\rangle \langle a_2, M_2\rangle X = \langle a_1, M_1\rangle (\langle a_2, M_2\rangle X)\,.$$

Note that there is only a finite number of distinct operations $\langle a, M\rangle$. Now we proceed to the main theorem of this section according to which the class of all nontrivial well-recognizable families of languages is the smallest class containing nontrivial elements of $\mathscr{M}$ and closed under just defined operations and, moreover, nontrivial elements of $\mathscr{M}$ form a basis of this class.

5.22. Theorem. (The decomposition of well-recognizable families of languages.) A family X is a nontrivial well-recognizable family iff it is of the form

$$X = \langle a_1, M_1\rangle \langle a_2, M_2\rangle \dots \langle a_n, M_n\rangle M_0$$

where M_0 is a nontrivial element of $\mathscr{M}$, $\langle a_i, M_i\rangle$ are some of the operations defined above. Furthermore, the characteristic string of X can be uniquely written as $u_X = a_1 \dots a_n$.

Proof. Denote by $\mathscr{W}$ the class of all nontrivial well-recognizable families, by $\mathscr{M}_0$ the class of all nontrivial elements of $\mathscr{M}$ and by $\hat{\mathscr{M}}_0$ the class generated from $\mathscr{M}_0$ by all operations of the form $\langle a, M\rangle$.

First we prove $\hat{\mathscr{M}}_0 \subseteq \mathscr{W}$. Surely $\mathscr{M}_0 \subseteq \mathscr{W}$ and it is enough to show that for $X \in \mathscr{W}$, $Y = \langle a, M\rangle X$ is $Y \in \mathscr{W}$. By definition

$$Y = aX \cup \bigcup\{aX + X_\Gamma;\ a \notin \Gamma\} \cup M\,.$$

$\partial_a Y = X$ and X is nontrivial thus necessarily Y is also nontrivial. Now we show that Y is recognizable. Clearly,

$$\pi_b Y = X_{\{b\}} \quad \text{for} \quad b \neq a \quad \text{and}$$
$$\pi_a Y = aX .$$

$X_{\{b\}}$ and aX are recognizable by Lemma 4.6 and 4.7; thus Y satisfies condition 1) of Theorem 4.9. Next it holds: If $a \in \Gamma$ then $Y \cap X_\Gamma = \sum_{\alpha \in \Gamma} \pi_\alpha Y$. If $a \notin \Gamma$ then $Y \cap \cap X_\Gamma \neq \emptyset \equiv X_\Gamma \subseteq M \equiv Y \cap X_\Gamma = \sum_{\alpha \in \Gamma} \pi_\alpha Y$. Thus Y satisfies also condition 2) and by Theorem 5.9 is recognizable.

It remains to show that $\bar{Y}$ is recognizable. Put

$$M' = \bigcup\{X_\Gamma;\ \pi_a X_\Gamma = \emptyset\} \setminus M .$$

It is easy to see that

$$\bar{Y} = a\bar{X} \cup \bigcup\{a\bar{X} + X_\Gamma;\ a \notin \Gamma\} \cup M' = \langle a, M' \rangle\, \bar{X} .$$

Thus by the first part of this proof $\bar{Y}$ is recognizable so Y is well-recognizable and we have $\hat{\mathscr{M}}_0 \subseteq \mathscr{W}$.

Now by induction on the degree of complexity of X we prove that $X \in \mathscr{W}$ implies $X \in \hat{\mathscr{M}}_0$.

Basis. If $\lg(u_X) = 0$ then by Corollary 5.18, $X \in \hat{\mathscr{M}}_0$.

Inductive Step. Assume the implication proved for families of degree of complexity n, where $n \geqq 0$ and let $\lg(u_X) = n + 1$. Then u_X is of the form av and by Corollary 5.17, $\lg(u_{\partial_a X}) = n$. Thus by the inductive assumption $\partial_a X \in \hat{\mathscr{M}}_0$. Put

$$M = \bigcup\{X_\Gamma;\ a \notin \Gamma\} \cap X .$$

Clearly $\langle a, M \rangle$ is well-defined operation and

$$\langle a, M \rangle\, \partial_a X = a\, \partial_a X \cup \bigcup\{a\, \partial_a X + X_\Gamma;\ a \notin \Gamma\} \cup M .$$

The union on the right-hand side is clearly disjoint, for $a \in \Gamma$ there is by Theorem 5.11, $X \cap X_\Gamma \neq \emptyset$ and thus we have:

1) If $\Gamma = \{a\}$ then $\langle a, M \rangle (\partial_a X \cap X_\Gamma) = a\, \partial_a X = \pi_a X = X \cap X_\Gamma$.

2) If $a \in \Gamma$ and $\{a\} \neq \Gamma$ then $\langle a, M \rangle (\partial_a X \cap X_\Gamma) = a\, \partial_a X + \sum_{\alpha \in (\Gamma - \{a\})} X_{\{\alpha\}} = X \cap X_\Gamma$.

3) If $a \notin \Gamma$ then $\langle a, M \rangle (\partial_a X \cap X_\Gamma) = X_\Gamma \cap M = X_\Gamma \cap X$.

So $X = \langle a, M \rangle\, \partial_a X$, i.e. $X \in \hat{\mathscr{M}}_0$ which proves $\mathscr{W} = \hat{\mathscr{M}}_0$. At the same time it is clear from the above construction that the expression for X is unique. Q.e.d.

Remark 1. If we don't require M_0 to be nontrivial we would get again all well-recognizable families, but their decomposition would not be unique (in the sense of M in operations). If we changed the definition of the operation $\langle a, M\rangle X$ by the requirement that $\pi_a M = \emptyset$ only if $X \neq \emptyset$, then all well-recognizable families would form a carrier of the algebra generated by empty set and operations of the type $\langle a, M\rangle$. Again the decomposition would not be unique (the empty set is not a basis of this algebra).

Remark 2. With the help of the preceding theorem it is easy to see that the family $X = \{L; u \in \text{Max}(L)\}$ from Example 3 on the beginning of this section is well-recognizable. If $u = a_1 \dots a_n$ then $X = \langle a_1, M\rangle \dots \langle a_n, M\rangle M_0$, where we put $M_0 = \{\{\Lambda\}\}$ and $M = \emptyset$. We can prove it in the other way too: we know (see proof of Theorem 3.1) that its complement $\bar{X} = \{L; u \notin \text{Max}(L)\}$ is recognizable. Since for any $a \in \Sigma$, $ua \notin \bigcup X$, we have $\bigcup X \neq \Sigma^*$ and thus by Theorem 5.6, X is well-recognizable.

We have paid a great deal of attention to well-recognizable families of languages. It is our opinion that they form a rather important and appealing subclass of the class of families recognizable by finite branching automata. We can support our opinion by several reasons.

First, from the point of view of the original motivation, one may be interested in problem-solving domains addmitting a "dual" world: where the successful plans will be exactly those that are "nonplans" on the original world. Such a domain might be represented by a finite branching automaton whose behavior is a well-recognizable family.

Second, as we have seen, well-recognizable families can be generated by a finite number of elements using a finite number of operations – a property which was not yet demonstrated for the general case. There is an interesting relationship between their complexity defined structurally and the complexity of corresponding automata. (Incidentally, most of the results of the last section are of a constructive nature.)

Third, the study of these families reveals typical aspects in which the theory of finite branching automata essentially differs from the classical automata theory. This is, in a certain sense, also a justification for developing a novel approach. It appears that there are also some purely mathematical reasons why the well-recognizable families are interesting. This aspect is the subject of another paper [3].

(Received February 21, 1977.)

REFERENCES

[1] I. M. Havel: Finite Branching Automata. Kybernetika *10* (1974), 281–302.

[2] V. Benda, K. Bendová: On specific features of recognizable families of languages. In: Mathematical Foundations of Computer Science (A. Mazurkiewicz, Ed.). Lecture Notes in Computer Science 45, Springer-Verlag, Berlin 1976, 187–193.

[3] V. Benda, K. Bendová: Recognizable Filters and Ideals. CMUC *17*, 2 (1976), 251–259.

[4] I. M. Havel: Finite branching automata: automata theory motivated by problem solving. In: Mathematical Foundations of Computer Science (A. Blikle, Ed.). Lecture Notes in Computer Science 28, Springer-Verlag, Berlin 1975, 53–61.

[5] I. M. Havel: On the branching structure of languages. In: Mathematical Foundations of Computer Science (A. Mazurkiewicz, Ed.). Lecture Notes in Computer Science 45, Springer-Verlag, Berlin 1976, 81–98.

[6] S. Eilenberg: Automata, Languages and Machines. Volume A. Academic Press, New York 1974.

[7] P. M. Cohn: Universal Algebra, Harper and Rower, New York 1965.

[8] A. Ginzburg: Algebraic Theory of Automata. Academic Press, New York–London 1968.

PhDr. Václav Benda, Výzkumný ústav matematických strojů (Research Institute for Mathematical Machines), Loretánské nam. 3, 110 00 Praha 1. Czechoslovakia.
RNDr. Kamila Bendová, Matematický ústav ČSAV (Mathematical Institute — Czechoslovak Academy of Sciences), Žitná 25, 115 67 Praha 1. Czechoslovakia.