ON SPECIFIC FEATURES OF RECOGNIZABLE FAMILIES

OF LANGUAGES

Václav Benda
Railway Research Institute
118 00 Prague 1

and

Kamila Bendová
Mathematical Institute ČSAV
115 67 Prague 1
Czechoslovakia

/Dedicated to Professor Jiří Bečvář
on the occasion of his fiftieth birthday/

A study of families of languages recognizable by recently introduced finite branching automata /cf. [1] / leads to a variety of results reminding analogical results from "classical" theory of automata and formal languages, if the involved concepts have the same or similar meaning for families of languages as they have for languages /sets of strings/. Let us mention for instance the property of having a finite number of distinct derivatives and the related decomposition of families [1], [2]. There are, on the other hand, some old questions with different answers: the class of all recognizable families is not closed under union and complement and it is properly included in the class of nondeterministically recognizable families.

In the present paper we are concerned with some problems that are inherently specific to the new area and which often require using notions and approaches of certain other branches of mathematics, otherwise seldom used in computer science /as, e.g., infinite cardinalities and the notions of a filter and ultrafilter/.

1 We use the following notation and terminology. Σ is a finite alphabet, Σ* the free monoid of strings over Σ /including the empty string Λ/, Ξ(Σ) is the set of all nonempty subsets of Σ*. L ∈ Ξ(Σ) is called a language, X ∈ Ξ(Σ) a family /of languages/. ∂uL = {v ; v ∈ Σ* & uv ∈ L} is the derivative of L with respect to u ∈ Σ*, Pref L is the set of all prefixes of L ; we define

FstL = (Pref L ∩ Σ) ∪ (L ∩ {Λ}) ,
i.e., using the notation of [3], FstL = ΛL(Λ). Analogously to ∂uL we de-
fine the derivative of a family $X$ with respect to $u \in \Sigma^*$ as the family
\[ \partial_u X = \{ \partial_u L ; L \in X \} - \{ \emptyset \}. \]

We define the $C$-closure of a family $X$ as the family
\[ C(X) = \{ L ; (\forall u \in \Sigma^*) (\exists L \in X) [\text{Fst}_A \partial_u L = \text{Fst}_A \partial_u L_u] \}. \]

We say that a family $X$ is finitely derivable if the set $D(X) = \{ \partial_u X ; u \in \Sigma^* \}$ is finite. $X$ is self-compatible if $C(X) = X$.

When defining the recognizability of families we shall bypass the definition of a deterministic finite branching automaton (c.f., e.g. [2]) by using the characterization theorem from [1]:

Characterization Theorem. A family of languages is recognizable by a finite branching automaton iff it is finitely derivable and self-compatible.

2 Our first observation concerns the cardinalities. Even if the set of all recognizable families — being associated with automata, i.e. finitary objects — is countable $\aleph_0$, it is obtained by intersecting two relatively "large" sets of cardinalities $\aleph_1$ and $\aleph_2$; we tacitly use the continuum hypothesis.\]

Theorem 1. The cardinality of the set of all self-compatible families is $\aleph_1$. The cardinality of the set of all finitely derivable families is $\aleph_2$.

Outline of the proof. Every singleton \{L\} is clearly self-compatible and thus the number of all self-compatible families is larger or equal to $\text{card} (\mathcal{L}(\Sigma)) = \aleph_1$. On the other hand, every self-compatible family is fully specified by a countable set of pairs $\langle u, F' \rangle$ where $F' = \text{Fst}_A \partial_u L$ for any $L \in X$. Thus, in fact, the equality holds.

Let $Y = \{ L ; \forall u \partial_u L \}$ and for every $X \in Y$ define $X' = X \cup (\mathcal{L}(\Sigma) - Y)$. Then for any $u \not\in \Lambda$, $\partial_u X' = \mathcal{L}(\Sigma)$; thus the family $X'$ is finitely derivable and the set of such families is clearly uncontrollable with the cardinality $\aleph_2$. Q. e. d.

3 Let us now consider the family $W$ of all complete languages (using the terminology of [4]):

1 Detailed and rigorous proofs of theorems 1-5 as well as some other results will be included in [2].