

Karel Čulík

Some notes on finite state languages and events represented by finite automata using labelled graphs

*Časopis pro pěstování matematiky*, Vol. 86 (1961), No. 1, 43--55

Persistent URL: <http://dml.cz/dmlcz/117361>

## Terms of use:

© Institute of Mathematics AS CR, 1961

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

SOME NOTES ON FINITE STATE LANGUAGES AND EVENTS  
REPRESENTED BY FINITE AUTOMATA USING LABELLED GRAPHS

KAREL ČULÍK, Brno

(Received October 15, 1959)

The theory of labelled graphs is used to demonstrate the equivalence between finite state languages (N. CHOMSKY-G. A. MILLER [1]) and events represented by finite automata (S. C. KLEENE [2] and JU. T. MEDVEDĚV [3]), and to show some possibilities of generalising or specialising the notions of finite state grammar and of finite automata (the notion of a finite indeterminate automaton is introduced).<sup>1)</sup>

**1. Introduction.** Let  $O = \{o_1, \dots, o_n\}$ ,  $n \geq 1$  be a finite set of *labels* (or symbols, words, letters, inputs) and let there be associated to every finite sequence  $o_{i_1}, o_{i_2}, \dots, o_{i_k}$  ( $k > 0$ ) of labels, a so-called *string*  $\alpha = o_{i_1}o_{i_2} \dots o_{i_k}$  of length  $k$ . By N. CHOMSKY and G. A. MILLER [1] a set of strings is called a *language* over vocabulary  $O$ , whilst by JU. T. MEDVEDĚV [3] such a set is called an *event* upon the symbols (or inputs) of  $O$ . These languages or events will be denoted by capitals letter  $\mathbf{M}, \mathbf{N}, \mathbf{P}, \dots$ , and the strings by lower-case Greek letters  $\alpha, \beta, \gamma, \dots$ . The set  $\mathbf{U}$  of all strings is said to be an *universal event* (or language); on  $\mathbf{U}$  there can be defined the well-known binary operation, *concatenation* of strings, as follows:

$$(1) \quad \alpha = o_{i_1} \dots o_{i_k}, \beta = o_{j_1} \dots o_{j_n} \Rightarrow \alpha\beta = o_{i_1} \dots o_{i_k}o_{j_1} \dots o_{j_n}.$$

This (free) semigroup of strings was investigated and axiomatised e. g. by P. C. ROSENBLOOM [4] and may be completed by a new so-called *null-string*  $\omega$  ( $\omega \notin \mathbf{U}$ ) having all the properties of unit

$$(2) \quad \omega\alpha = \alpha\omega = \alpha \quad \text{for every } \alpha \in \mathbf{U} \cap \{\omega\} = \mathbf{U}_0$$

(i. e.,  $\omega$  has the length zero). In another way one may take a new label  $o_0$  ( $o_0 \notin O$ ), start with a vocabulary  $O_0 = O \cup \{o_0\}$  and obtain, under the conditions (1) and (2) ( $o_0 = \omega$ ), the same semigroup with the unit  $\omega$ .

<sup>1)</sup> First after the finishing this paper I have found further important results of N. CHOMSKY: On certain formal properties of grammars (Inf. a. Control 2, 137—167, 1959) and of M. O. RABIN and D. SCOTT: Finite automata and their decision problems (IBM Jour. of research and development 3, 114—125, 1959), where, among others, the notion of indetermined automaton was introduced.

The concatenation on  $U_0$  induces another binary operation – also to be called *concatenation* – on the Boolean algebra of all subsets of  $U_0$  as follows

$$(3) \quad \emptyset \neq M, N \subset U_0 \Rightarrow MN = \{\alpha\beta \mid \alpha \in M, \beta \in N\}$$

either  $M = \emptyset$  or  $N = \emptyset \Rightarrow MN \Rightarrow \emptyset$ .

This semigroup also has a unit  $O_0 = \{o_0\}$ .

Now, if we replace every row of the  $k \times l$ -matrix ( $k$  fixed,  $2^k \leq n$ ) studied by S. C. KLENNE [2] by our label  $o_i$  ( $1 \leq i \leq n$ ), we obtain a string of length  $l$ . Thus to every event in sense of Kleene (i. e. a set of such matrices) there corresponds an our event  $M \subset U$ . In the algebra of events the Kleene *disjunction* “ $\vee$ ” corresponds to our set-theoretical sum, his *product* (the symbols are identical) to our concatenation, and his third operation of *iteration*  $E * F$  corresponds (using  $O_0$ , which is not introduced in [2]) to the *infinite power* of an event, defined as follows

$$(4) \quad M^{\infty_0} = \bigcup_{i=0}^{\infty} M^i, \quad \text{where } M^0 = O_0, M^1 = M, M^2 = MM, \dots,$$

since we may write (using the correspondence)

$$(5) \quad E * F = F \vee EF \vee E^2F \vee \dots = O_0F \cup EF \cup E^2F \cup \dots = E^{\infty_0}F.$$

Several identities may be established, e. g.

$$(6) \quad M(\bigcup_{i \in I} N_i) = \bigcup_{i \in I} (MN_i), \quad (\bigcup_{i \in I} M_i)N = \bigcup_{i \in I} (M_iN),$$

$$(7) \quad M \subset M^{\infty_0}, \quad (M^{\infty_0})^{\infty_0} = M^{\infty_0}, \quad M \subset N \Rightarrow M^{\infty_0} \subset N^{\infty_0}$$

(these show that infinite power is a *closure operation*),

$$(8) \quad (M \cup N)^{\infty_0} = (M^{\infty_0} \cup N^{\infty_0})^{\infty_0},$$

$$(M \cup N)^{\infty_0} = M^{\infty_0} \cup N^{\infty_0} \cup (M^{\infty_0}N^{\infty_0})^{\infty_0} \cup (N^{\infty_0}M^{\infty_0})^{\infty_0},$$

$$M^{\infty_0} = M^k M^{\infty_0} \cup \bigcup_{i=0}^k M^i, \text{ etc.}$$

Many other identities may be found in [2].

If we denote by  $O_i$  ( $1 \leq i \leq n$ ) an event containing only a single string  $o_i$  and if we introduce another infinite power

$$(9) \quad M^{\infty} = \bigcup_{i=1}^{\infty} M^i \quad (\text{thus } M^{\infty_0} = M^{\infty} \cup O_0),$$

we may define the set  $\Omega$  resp.  $\Omega_0$  as the smallest set of events containing  $O_i$  for all  $i = 1, 2, \dots, n$  and the *zero-event*  $\emptyset$ , which is closed under the operations of set-theoretical sum, of concatenation and of infinite power  $\infty$  resp.  $\infty_0$ . In regard to the notion of a regular event of [2] there holds.

**1.1. Theorem.** *For an event  $M$  the following conditions are equivalent:*

- (a)  $M$  is a regular event in the sense of Kleene,
- (b)  $M \in \Omega_0$  and  $O_0 \notin M$ ,
- (c)  $M \in \Omega$ .

Proof. (a)  $\Rightarrow$  (b) follows by (5) and by definition of  $\Omega$ , (b)  $\Rightarrow$  (c) follows by (9) and by definitions of  $\Omega$  and  $\Omega_0$  and finally (c)  $\Rightarrow$  (a) follows by (5), because evidently  $\mathbf{M}^\infty = \mathbf{M} * \mathbf{M}$  (always in the correspondence established above).

Note that the infinite powers  $\mathbf{M}^{\infty_0}$  resp.  $\mathbf{M}^\infty$  are free semigroups with resp. without an unit element ( $\mathbf{M}$  need not be the set of independent generators) and it is possible to write  $\mathbf{U}_0 = \left(\bigcup_{i=1}^n \mathbf{O}_i\right)^{\infty_0}$  and  $\mathbf{U} = \left(\bigcup_{i=1}^n \mathbf{O}_i\right)^\infty$ .

From Theorem 1.1 it follows that if we want to use  $\infty_0$ -powers to describe our regular events then this can be performed only in some "context", namely each  $\infty_0$ -power must be concatenated (or multiplied in [2]) by some event which does not contain  $\mathbf{O}_0$  (this is sense of (5)); the precise formulation of this condition is given in the definition of the so-called Chomsky-Miller notation (see section 3 of this paper).

**2. Labelled graphs.** In [1] to every finite state grammar there is associated a labelled finite graph, the so-called *state diagram* of this grammar. Now we introduce the most general notion of a finite, directed and labelled graph over vocabulary  $O$  (see the previous section) as a quadruple  $(V, E, I, L)$ , where  $V$  resp.  $E$  is the finite set of *vertices* resp. *edges* and  $I$  resp.  $L$  is a mapping of  $E$  into  $V \times V$  (i. e. *incidence*) resp. of  $E$  onto  $O$  (i. e. *labeling*). Two graphs  $(V, E, I, L)$  and  $(V', E', I', L')$  are *isomorphic* if there exist two one-to-one mappings  $f$  of  $V$  onto  $V'$  and  $g$  of  $E$  onto  $E'$  such that

$$(10) \quad e \in E, \quad I(e) = [x, y] \Rightarrow I'(g(e)) = [f(x), f(y)] \quad \text{and} \quad L'(g(e)) = L(e).$$

To every edge  $e_i \in E$  there corresponds a symbol  $e_i[x_i, y_i] o_{a_i} = L(e_i)$ , the so-called *labelled edge*, where  $[x_i, y_i] = I(e_i)$  and  $O_{a_i} = L(e_i)$ . If  $\mathfrak{E}$  is the set of all labelled edges, we denote our graph by  $\Gamma = \langle V, \mathfrak{E} \rangle$ . The set of vertices resp. of labelled edges of a graph  $\Gamma_i, \Gamma'$  etc., will be denoted by  $V_i, V'$  resp.  $\mathfrak{E}_i, \mathfrak{E}'$ . In figures there we use the usual representation of graphs in the plane.

A finite sequence of labelled edges  $\{e_i[x_{i-1}, x_i] o_{a_i}\}_{i=1}^k$  ( $k \geq 1$ ) is called a *labelled chain* in  $\langle V, \mathfrak{E} \rangle$  with first resp. last vertex  $x_0$  resp.  $x_k$ ; we shall say that this labelled chain *generates* the string  $\alpha = o_{a_1} o_{a_2} \dots o_{a_k}$ .

Let there be given an ordered pair  $(X, Y)$ , where  $X \subset V$  resp.  $Y \subset V$  is said to be the *set of first resp. last vertices*. Denote by  $\Gamma(X, Y)$  an event containing all the strings which are generated by all the labelled chains in  $\Gamma$  with first vertices in  $X$  and last vertices in  $Y$ .

Let  $\Omega_G$  be the set of all events  $\mathbf{M}$  such that there exist a labelled graph  $\Gamma = \langle V, \mathfrak{E} \rangle$  (over vocabulary  $\mathbf{O}$ ) and sets of first and last vertices  $X, Y \subset V$  satisfying  $\Gamma(X, Y) = \mathbf{M}$ . In this section we shall prove

**2.1. Theorem.**  $\Omega \subset \Omega_G$ .

First of all we prove some lemmas and introduce some new notions.

**2.2. Lemma.** *If  $\mathbf{M} = \Gamma(X, Y)$ , then there exists a subgraph  $\Gamma'$  in  $\Gamma$  such that:  $\mathbf{M} = \Gamma'(X', Y')$ , where  $X' \subset X, Y' \subset Y$ , every  $x \in X'$  resp.  $y \in Y'$  is first resp. last vertex*

of at least one labelled chain with first resp. last vertex in  $X'$  resp. in  $Y'$ , every  $v \in V'$  and every  $e[u, v] o_k \in \mathfrak{E}'$  is contained in some labelled chain with first resp. last vertex in  $X'$  resp. in  $Y'$ , and finally  $\Gamma'$  is minimal in regard to these properties. Thus

$$(I) \quad e[u, v] o_k, \quad e'[u, v] o_k \in \mathfrak{E}' \Rightarrow e = e'.$$

Proof.  $\Gamma$  is finite and therefore it suffices to remove suitable labelled edges and possibly some vertices also.

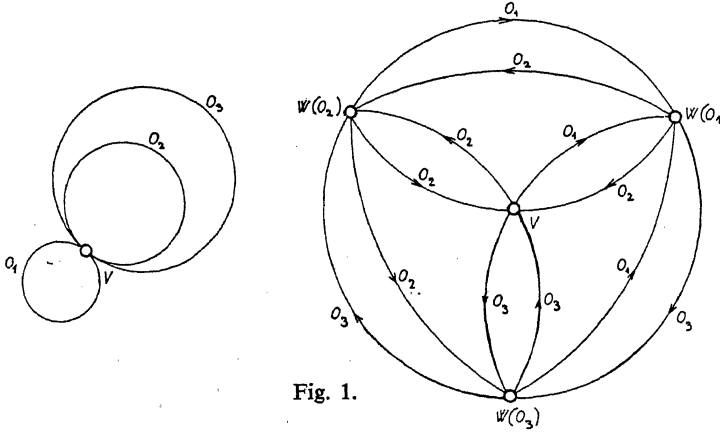


Fig. 1.

The subgraph  $\Gamma'$  in 2.2 is called *minimal in regard to the event*  $\Gamma(X, Y)$ . When dealing with labelled graph satisfying (I) we always denote its labelled edges only by  $[u, v] o_k$ , since in this case

$$(11) \quad [u_1, v_1] o_k = [u_2, v_2] o_h \Leftrightarrow u_1 = u_2, \quad v_1 = v_2, \quad o_k = o_h.$$

**2.3. Lemma.** Let  $\mathbf{M} = \Gamma(v, v)$ , where  $\Gamma$  is minimal in regard to  $\mathbf{M}$ ,

$$V = \{v\}, \quad \mathfrak{E} = \{[v, v] o_{a_1}, \dots, [v, v] o_{a_p}\}$$

and let  $\Gamma'$  be defined as follows:

$$V' = \{v\} \cup W, \quad v \text{ non} \in W, \quad W = \{w(o_{a_1}), \dots, w(o_{a_p})\},$$

where  $w(o_{a_j})$  denotes a new vertex, and

$$\mathfrak{E}' = \{[v, w(o_{a_j})] o_{a_j} \mid 1 \leq j \leq p\} \cup \{[w(o_{a_j}), v] o_{a_j} \mid 1 \leq j \leq p\} \cup \\ \cup \{[w(o_{a_j}), w(o_{a_h})] o_{a_h} \mid j \neq h, 1 \leq j, h \leq p\}.$$

Then (i)  $\Gamma'(v', v') = \mathbf{M}$  for all  $v' \in V'$ , (ii)  $\Gamma'$  is minimal in regard to  $\mathbf{M}$  and (iii)  $\Gamma'$  satisfies the following condition

$$(II) \quad e[u, v] o_k \in \mathfrak{E}' \Rightarrow u \neq v \quad (\text{i. e. } \Gamma' \text{ does not contain slings}).$$

Proof. (i)–(iii) follow from the definition of  $\Gamma'$ , because in every  $v' \in V'$  starts just one edge labelled by  $o_{a_j}$  for all  $j = 1, 2, \dots, p$ . Note only that  $\mathbf{M} = (O_{a_1} \cup \dots \cup O_{a_p})^\infty$  (a simple example is in Fig. 1).

**2.4. Lemma.** *If  $\mathbf{M} = \Gamma(x, X)$ , then there exist  $\Gamma'$  and  $y \in V'$ ,  $Y \subset V'$  such that  $\Gamma'$  satisfies (I), (II) and  $\Gamma'(y, Y) = \mathbf{M}$ .*

*Proof.* By 2.2 we may suppose that  $\Gamma$  is minimal in regard to  $\mathbf{M}$ . Let now  $[v_i, v_i] o_{a_j}$  for  $j = 1, 2, \dots, p_i$  be all the slings in  $\Gamma$  incident to the vertex  $v_i$ . We construct graphs  $\Gamma'_i$  as in 2.3 for graphs  $\Gamma_i$ , when  $V_i = \{v_i\}$  and

$$\mathfrak{C}_i = \{[v_i, v_i] o_{a_1}, \dots, [v_i, v_i] o_{a_{p_i}}\},$$

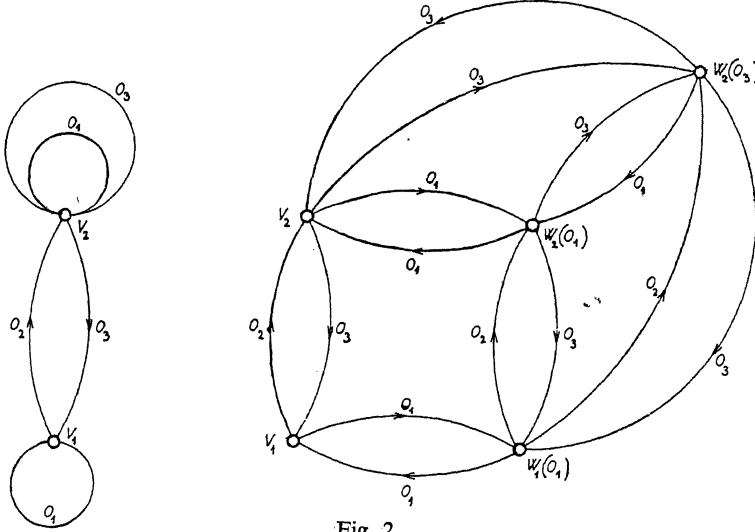


Fig. 2.

for all  $i = 1, 2, \dots, m$  (where  $V = \{v_1, \dots, v_m\}$ ) and we demand  $V'_i \cap V'_k = \emptyset$  if  $i \neq k$ ,  $1 \leq i, k \leq m$ . Now define  $\Gamma'$  as follows:

$$V' = \bigcup_{i=1}^m V'_i, \quad \mathfrak{C}' = \bigcup_{i=1}^m \mathfrak{C}_i \cup \{[w_i(o_{a_r}), w_k(o_{a_s})] o_{a_j} \mid i \neq k, 1 \leq i, k \leq m, \\ [v_i, v_k] o_{a_j} \in \mathfrak{C}, 1 \leq r \leq p_i, 1 \leq s \leq p_k\}$$

(a simple example is in Fig. 2). Finally let

$$y = x \quad \text{and} \quad Y = X \cup \bigcup_{i=1}^m W_i.$$

It is easily shown that  $\Gamma'$  satisfies (I) and (II), and by 2.3 also that  $\Gamma'(y, Y) = \mathbf{M}$ .

**2.5. Lemma.** *If  $\mathbf{M} = \Gamma(x, X)$ , then there exist  $\Gamma'$ ,  $y \in V'$  and  $Y \subset V'$  such that the following conditions are satisfied:  $\Gamma'(y, Y) = \mathbf{M}$ , (I), (II) and*

(III) *there is no labelled chain in  $\Gamma'$  the first and last vertex of which is  $y$ .*

*Proof.* By 2.2 and 2.4 we may suppose that  $\Gamma$  satisfies (I), (II) and that it is minimal in regard to  $\mathbf{M}$ . We construct  $\Gamma'$  as follows:  $V' = V \cup \{y\}$ , where  $y \notin V$ ,  $\mathfrak{C}' =$

$= \mathfrak{E} \cup \{[y, u] o_j \mid [x, u] o_j \in \mathfrak{E}, u \in V, 1 \leq j \leq n\}$ ,  $Y = X$ . By this construction it is easy to see that  $\Gamma'$  is the required graph.<sup>2)</sup>

**2.6. Lemma.** *If  $\mathbf{M} = \Gamma(X, Y)$ , then there exist  $\Gamma'$ ,  $p \in V'$  and  $Q \subset V'$  such that  $\Gamma'(p, Q) = \mathbf{M}$ .*

*Proof.* By 2.2 we may suppose that  $\Gamma$  satisfies (I). If  $X = \{x_1, \dots, x_m\}$ , let  $\Gamma_i$  be a minimal subgraph of  $\Gamma$  in regard to  $\Gamma(x_i, Y)$  for  $i = 1, 2, \dots, m$ , i. e.  $\Gamma_i(x_i, Y_i) = \Gamma(x_i, Y)$ , where  $Y_i \subset Y$ . Let  $\Gamma'_i$  be isomorphic to  $\Gamma$  and let  $p_i \in V'_i$ ,  $x_i \in V$  resp.  $Q_i \subset V'_i$ ,  $Y_i \subset V$  be corresponding elements resp. sets of elements in this isomorphism, i. e.  $\Gamma'_i(p_i, Q_i) = \Gamma_i(x_i, Y_i)$ . By 2.5 we may suppose that  $\Gamma'_i$  satisfies (III). If we put  $p = p_i$  and require  $V'_i \cap V'_j = \{p\}$  for all  $i \neq j$ ,  $1 \leq i, j \leq m$ , we may define  $\Gamma'$  as follows:  $V' = \bigcup_{i=1}^m V'_i$ ,  $\mathfrak{E}' = \bigcup_{i=1}^m \mathfrak{E}'_i$ ,  $Q = \bigcup_{i=1}^m Q_i$ . Evidently  $\Gamma'(p, Q) = \bigcup_{i=1}^m \Gamma'_i(p_i, Q_i) = \mathbf{M}$ .

**2.7. Lemma.** *If  $\mathbf{M} = \Gamma(x, y)$  and  $\Gamma$ ,  $x$  and  $y$  satisfy (I)–(III), then there exist  $\Gamma'$ ,  $x' \in V'$  and  $Y' \subset V'$  such that  $\Gamma'(x', Y') = \mathbf{M}$ , and that  $\Gamma'$ ,  $x'$  and  $Y'$  satisfy conditions (I)–(III) and*

(IV) *there is no labelled chain in  $\Gamma'$  the first and last vertex of which is some vertex  $y' \in Y'$ .*

*Proof.* If  $\Gamma(y, y) = \emptyset$ , put  $\Gamma' = \Gamma$ ,  $x' = x$  and  $Y' = \{y\}$ . Further let  $\Gamma(y, y) = \mathbf{N} \neq \emptyset$  and therefore by (III)  $x \neq y$ . Let  $\Gamma^*$  be isomorphic to a minimal subgraph of  $\Gamma$  in regard to  $\Gamma(y, y)$  and let  $y^* \in V^*$  and  $y \in V$  be corresponding vertices in this isomorphism. Now we “split” the vertex  $y^*$  into parts  $y_0^*, y_1^*, \dots, y_k^*$  such that all labelled edges starting in  $y^*$  will start in  $y_0^*$  and all labelled edges finishing in  $y^*$  will finish in just one vertex  $y_i^*$ ,  $1 \leq i \leq k$ . This new graph we denote  $\Gamma_0$  (an example of this and following constructions is in Fig. 3). It is easy to see that if  $\Gamma_0(y_0^*, Y_0) = \mathbf{Q}$ , where  $Y_0 = \{y_1^*, \dots, y_k^*\}$ , then  $\mathbf{Q}^\infty = \mathbf{N}$ .

Let  $\mathbf{P}$  be the set of strings generated by all labelled chains

$$\{[w_{i-1}, w_i] o_{a_i}\}_{i=1}^k$$

in  $\Gamma$  such that  $w_0 = x$ ,  $w_k = y$  and  $w_i \neq y$  for all  $1 \leq i < k$ . Let  $\Gamma^{**}$  be the minimal subgraph of  $\Gamma$  containing all vertices and labelled edges of these generating labelled chains. Let  $\Gamma_1$  be isomorphic to  $\Gamma^{**}$  and let  $x_1 \in V_1$ ,  $x \in V^*$  and  $y_1 \in V_1$ ,  $y \in V^{**}$  be corresponding vertices in this isomorphism, i. e.  $\Gamma_1(x_1, y_1) = \mathbf{P}$ .

$\Gamma'$  is constructed as follows:  $V' = V \cup V_0 \cup V_1$ ,  $\mathfrak{E}' = \mathfrak{E} \cup \mathfrak{E}_0 \cup \mathfrak{E}_1$ , where we identify  $x' = x = x_1$ ,  $y = y_0^*$ ,  $Y' = \{y_1\} \cup Y_0$  and require  $V \cap V_0 = \{y\}$ ,  $V \cap V_1 = \{x\}$  and  $V_0 \cap V_1 = \emptyset$ . Now  $\mathbf{M} = \mathbf{P} \cup \mathbf{PN} = \mathbf{P}(\mathbf{O}_0 \cup \mathbf{N})$  and  $\Gamma'(x', Y') = \mathbf{P} \cup \mathbf{MQ} = \mathbf{P} \cup \mathbf{P}(\mathbf{O}_0 \cup \mathbf{N})\mathbf{Q} = \mathbf{P} \cup \mathbf{P}(\mathbf{Q} \cup \bigcup_{i=2}^{\infty} \mathbf{Q}^i) = \mathbf{P} \cup \mathbf{PN} = \mathbf{M}$  and by construction of  $\Gamma'$  it follows that  $\Gamma'$ ,  $x'$  and  $Y'$  satisfy (I)–(IV).

<sup>2)</sup> The idea to „double” a first vertex and an essential abbreviation of this proof is due to J. BEČVÁŘ. I wish thank him also for other remarks to this paper.

Before proceeding to the proof of Theorem 2.1 we note that to the set-theoretical sum of events there corresponds, in the algebra of graphs, the cardinal sum, to the

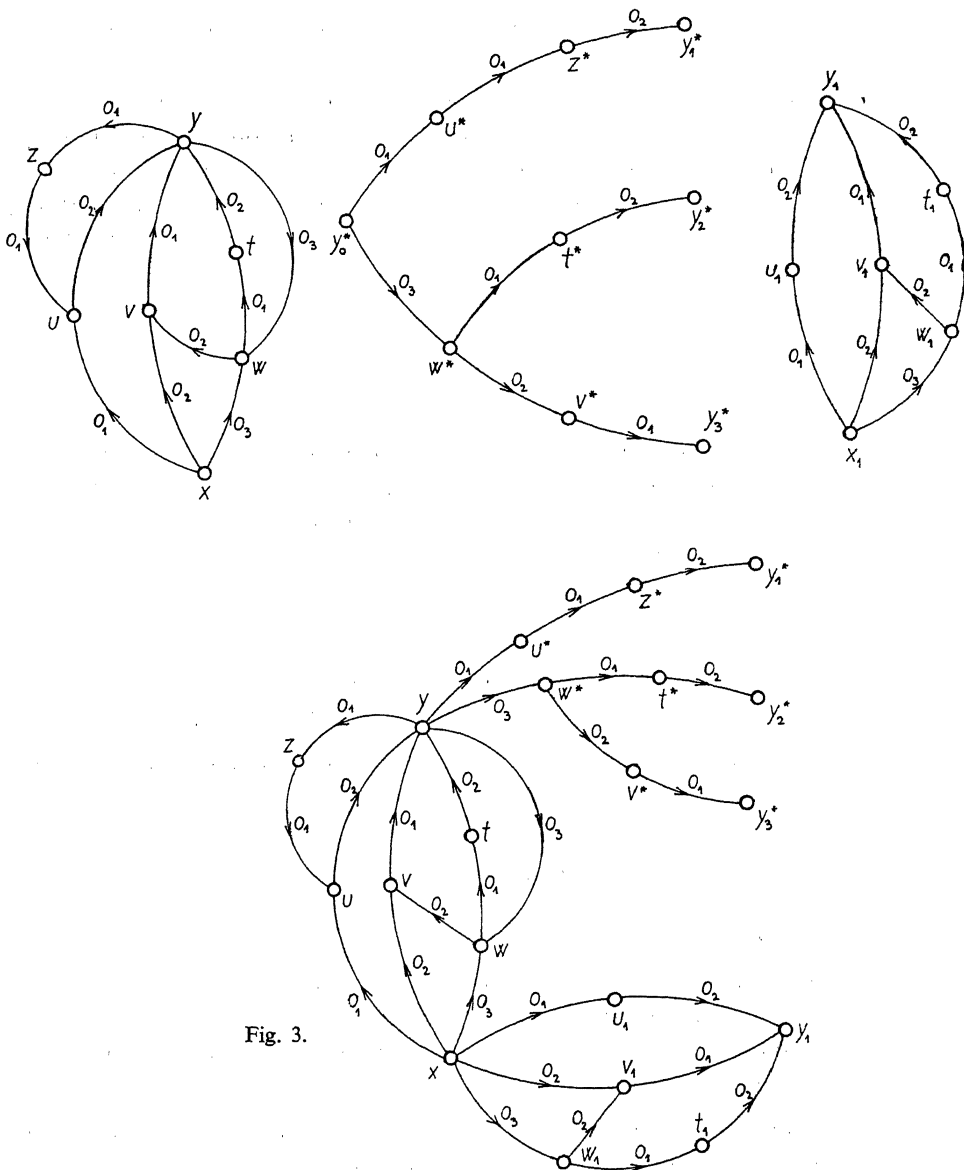


Fig. 3.

concatenation there corresponds (under certain conditions) identifying last vertices of one and first vertices of another graph and to the infinite power there corresponds (under certain conditions) identifying first and last vertices in the same graph (see proof of 2.7).



Proof of Theorem 2.1. If  $\mathbf{M} = \mathbf{O}_i$ ,  $1 \leq i \leq n$  or  $\mathbf{M} = \emptyset$ , then evidently  $\mathbf{M} \in \Omega_G$ . Next we shall prove  $\mathbf{M}, \mathbf{N} \in \Omega_G \Rightarrow \mathbf{M} \cup \mathbf{N}, \mathbf{MN}, \mathbf{M}^\infty \in \Omega_G$ .

1. If  $\mathbf{M}' = \Gamma_1(X_1, Y_1)$ ,  $\mathbf{N} = \Gamma_2(X_2, Y_2)$ , then  $\mathbf{M} \cup \mathbf{N} = \Gamma(X, Y)$ , where  $V = V_1 \cup V_2$ ,  $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2$  (we suppose, of course, that  $V_1 \cap V_2 = \emptyset$ ) and  $X = X_1 \cup X_2$ ,  $Y = Y_1 \cup Y_2$ .

2. Let  $\mathbf{M} = \Gamma_0(X, Y)$ , where  $Y = \{y_1, \dots, y_s\}$  and let  $\mathbf{N} = \Gamma'(P, Q)$ . By 2.6 we may suppose that  $X = \{x\}$ ,  $P = \{p\}$  and by 2.5 that  $x \text{ non } \in Y$ ,  $p \text{ non } \in Q$  (this follows by condition (III)). Let  $\Gamma_i$  be isomorphic to  $\Gamma'$  for  $1 \leq i \leq s$  and let  $p_i \in V_i$ ,  $p \in V'$  resp.  $Q_i \subset V_i$ ,  $Q \subset V'$  be corresponding vertices resp. sets of vertices in this isomorphism. We identify  $p_i = y_i$  and require  $V_i \cap V_j = \emptyset$  for  $i \neq j$  and  $V_0 \cap V_i = \{y_i\}$  for  $1 \leq i, j \leq s$ . Now if we define

$$V = \bigcup_{i=0}^s V_i, \quad \mathcal{E} = \bigcup_{i=0}^s \mathcal{E}_i,$$

then obviously  $\Gamma(x, \bigcup_{i=1}^s Q_i) = \mathbf{MN}$ .

3. Let  $\mathbf{M} = \Gamma(X, Y)$ . By 2.6 and 2.5 we may suppose that  $X = \{x\}$  and that conditions (I)–(III) are satisfied. Let  $Y = \{y_1, \dots, y_s\}$  and let  $\Gamma_i$  be the minimal subgraph in  $\Gamma$  in regard to  $\Gamma(x, y_i)$ . Further let  $\Gamma_i^*$  be isomorphic to  $\Gamma_i$  and let  $x^* \in V_i^*$ ,  $x \in V_i$  and  $y_i^* \in V_i^*$ ,  $y_i \in V_i$  be corresponding vertices in this isomorphism, i. e.

$$\bigcup_{i=1}^s \Gamma_i^*(x^*, y_i^*) = \mathbf{M}.$$

The graphs  $\Gamma_i^*$  satisfy all the conditions (I)–(III) and by 2.7 there exist graphs  $\Gamma'_i$ ,  $x' \in V'_i$  and  $Y'_i \subset V'_i$  satisfying  $\Gamma'_i(x', Y'_i) = \Gamma_i^*(x^*, y_i^*)$  for  $1 \leq i \leq s$  and all the conditions (I)–(IV). We identify  $x' = u$  for all  $u \in \bigcup_{i=1}^s Y'_i$  and require  $V'_i \cap V'_j = \{x'\}$  for  $i \neq j$ . Now, if we define

$$V' = \bigcup_{i=1}^s V'_i, \quad \mathcal{E}' = \bigcup_{i=1}^s \mathcal{E}'_i,$$

then  $\Gamma'(x', x') = \mathbf{M}^\infty$ .

**3. Finite state grammars.** A *finite state language*  $L_G$  is an event *generated* (see [1]) by a (finite state) *grammar*  $G$ , the so-called *state diagram* of which is a finite labelled graph over extended vocabulary  $O_0 = O \cup \{o_0\}$  (cf. section 1). Let  $\Omega_L$  be the set of all  $L_G$  generated by grammars  $G$  such that

(12) *in the state diagram of  $G$  there are no labelled chains starting and finishing in the initial state all edges of which are labelled by  $o_0$ .*

**3.1. Theorem.**  $\Omega_G \subset \Omega_L$ .

**Proof.** If  $\mathbf{M} \in \Omega_G$ , i. e. if there exist  $\Gamma$ ,  $X$  and  $Y$  such that  $\mathbf{M} = \Gamma(X, Y)$  and  $\Gamma$  is a labelled graph with vocabulary  $O$ , then by 2.5 and 2.6 we may suppose that  $X = \{x\}$  and that  $\Gamma$ ,  $x$  and  $Y$  satisfy (I)–(III). We define a labelled graph  $\Gamma_0$  with extended

vocabulary  $O_0$  as follows:  $V_0 = V$ ,  $\mathfrak{C}_0 = \mathfrak{C} \cup \{[y_i, x] o_0 \mid y_i \in Y\}$ . If  $x = v_0$  is the initial state, then  $\Gamma_0$  is the state diagram of a grammer  $G_0$ . But from the conditions (I)–(III) for  $\Gamma$  it follows that  $G_0 \in F_3$  (cf. [1], p. 97) and also  $L_{G_0} \in \Omega_L$  (because from (III) follows  $x \text{ non} \in Y$ ). Thus it is sufficient to prove  $L_{G_0} = \mathbf{M}$ .

If

$$\alpha = o_{a_1} o_{a_2} \dots o_{a_q} \in L_{G_0} \quad (a_i \neq 0 \text{ for } 1 \leq i \leq q),$$

there exist (cf. [1], p. 95) a string  $\beta = o_{b_1} o_{b_2} \dots o_{b_r}$  and a sequence

$$v_{c_1}, \dots, v_{c_{r+1}} \quad (v_{c_i} \in V_0)$$

such that

- (i)  $v_{c_1} = v_{c_{r+1}} = v_0$ , (ii)  $v_{c_i} \neq v_0$  for  $1 < i < r + 1$ ,  
 (iii)  $[v_{c_i}, v_{c_{i+1}}] o_{b_i} \in \mathfrak{C}_0$  for  $1 \leq i \leq r$  and (iv)  $\alpha = \beta$

(in regard to (2)). From

$$[v_{c_r}, v_{c_{r+1}}] o_{b_r} \in \mathfrak{C}_0$$

and from  $G_0 \in F_3$  (i. e.  $[u, v] o_0 \in \mathfrak{C}_0 \Rightarrow v = v_0$  and  $[u, v_0] o_k \in \mathfrak{C}_0 \Rightarrow o_k = o_0$ ) it follows that  $o_{b_r} = o_0$ ,  $v_{b_i} \neq o_0$  for  $1 \leq i < r$  and therefore  $r - 1 = q$  and  $o_{b_i} = o_{a_i}$  for  $1 \leq i \leq q$ . At least  $v_{c_r} \in Y$  and therefore

$$\{[v_{c_i}, v_{c_{i+1}}] o_{a_i}\}_{i=1}^q$$

is a labelled chain in  $\Gamma$  starting in  $x$  and finishing in  $Y$ , generating the string  $\alpha$ , so that  $\alpha \in \mathbf{M} = \Gamma(x, Y)$ .

If on the contrary  $\alpha = o_{a_1} \dots o_{a_q} \in \mathbf{M} = \Gamma(x, Y)$ , then there exists a labelled chain

$$\{[v_{c_i}, v_{c_{i+1}}] o_{a_i}\}_{i=1}^q$$

in  $\Gamma$  such that  $v_{c_1} = x (= v_0)$  and  $v_{c_{q+1}} \in Y$ . Now the sequence  $v_{c_1}, \dots, v_{c_{q+1}}, v_{c_{q+1}} = v_0$  and the string (over vocabulary  $O_0$ )

$$\beta = o_{b_1} \dots o_{b_q} o_{b_{q+1}}, \quad \text{where } o_{b_i} = o_{a_i} \text{ for } 1 \leq i \leq q \text{ and } o_{b_{q+1}} = o_0,$$

satisfy conditions (i), (iii) and (iv) and by (III) for  $\Gamma$  the condition (ii) also. Therefore  $\alpha \in L_{G_0}$ .

In [1], p. 104, there is introduced the so-called *Chomsky-Miller notation*  $X_1(X_2, \dots, \dots, X_t) X_{t+1}$ , where  $X_i$  are strings or, recurrently, Ch. M. notations, and it is proved (Theorem 6) that every  $L_G$  can be represented by finite number of Ch. M. notations.

**3.2. Lemma.** *If  $L_G$  is represented by a single Ch. M. notation, then  $L_G$  is an event which may be expressed in the algebra of events as follows: in the Ch. M. notation there are symbols representing strings, commas and pairs of brackets; if we replace each string by an event containing this single string only, each comma by the symbol of set-theoretical sum  $\cup$  and if we place the  $\infty_0$ -power of every expression in a pair of brackets (neighbouring expressions concatenated of course), we obtain an event equal to  $L_G$ .*

*Proof.* Let  $m$  be the number of pairs of brackets in the given Ch. M. notation. If  $m = 0$ , our lemma is obviously true; we may suppose that it is true for each  $p$ ,

$0 \leq p < m$ . If now our Ch. M. notation contains  $m$  pairs of brackets, it has the following form:  $L_G = \alpha(A_1, \dots, A_a) \beta(B_1, \dots, B_b) \gamma(\dots) \delta(D_1, \dots, D_d) \varepsilon$ , where  $\alpha, \beta, \gamma, \dots, \delta, \varepsilon$  are strings, while  $A_i, B_j, \dots, D_k$  are either strings also or Ch. M. notations containing less than  $m$  pairs of brackets. Therefore by induction  $A_i, B_j, \dots, D_k$  represent some events  $\mathbf{A}_i, \mathbf{B}_j, \dots, \mathbf{D}_k$  respectively, each of which is constructed in the described manner.

We construct (cf. [1], p. 105) a string of the set of strings represented by the Ch. M. notation  $\alpha(A_1, \dots, A_a) \beta$  as follows: we take a sequence  $X_1 = \alpha, X_2, \dots, X_i, X_{i+1} = \beta$  such that  $X_2, \dots, X_i$  is some quite arbitrary sequence of symbols  $A_1, \dots, A_a$ . Further if  $A_i$  is a Ch. M. notation too, we take a new sequence of form similar to  $A_i$  and put this sequence in place of  $A_i$  in the previous sequence, etc. Each string which is result of such a construction (cf. [1], p. 105), may also be obtained by taking suitable strings  $\alpha_i \in \mathbf{A}_i, 1 \leq i \leq a$  and forming the concatenation  $\alpha\alpha_1 \dots \alpha_a\beta$ ; and conversely. Therefore the Ch. M. notation  $\alpha(A_1, \dots, A_a) \beta$  represents the event  $\{\alpha\}(\mathbf{A}_1 \cup \dots \cup \mathbf{A}_a)^{\circ\circ\circ\{\beta\}}$ , and the same holds for the other Ch. M. notations  $\beta(B_1, \dots, B_b) \gamma, \dots, \delta(D_1, \dots, D_d) \varepsilon$ . Also

$$L_G = \{\alpha\}(\mathbf{A}_1 \cup \dots \cup \mathbf{A}_a)^{\circ\circ\circ\{\beta\}}(\mathbf{B}_1 \cup \dots \cup \mathbf{B}_b)^{\circ\circ\circ\{\gamma\}}(\dots)^{\circ\circ\circ\{\delta\}} \cdot (\mathbf{D}_1 \cup \dots \cup \mathbf{D}_d)^{\circ\circ\circ\{\varepsilon\}},$$

which proves our lemma.

**3.3. Corollary.** *The Ch. M. notation  $\alpha(A_1, \dots, A_a) \beta \dots \delta(D_1, \dots, D_d) \varepsilon$  represents an  $L_G \in \Omega_L$  if and only if*

(13) *at least one of strings  $\alpha, \beta, \dots, \delta, \varepsilon$  outside every pair of brackets is not equal to  $o_0$ .*

The proof follows immediately from the definition of  $\Omega_L$  and by (4) and 3.2.

**3.4. Theorem.** *If  $\mathbf{M}$  is an event over the extended vocabulary  $O_0$ , the following conditions are equivalent:*

- (a)  $\mathbf{M}$  is a regular event in the sense of Kleene,
- (d) There exist a labelled graph  $\Gamma = \langle V, \mathfrak{E} \rangle$  over the vocabulary  $O$  and the sets  $X, Y \subset V$  such that  $\mathbf{M} = \Gamma(X, Y)$ , i. e.  $\mathbf{M} \in \Omega_G$ ,
- (e) There exists a grammar  $G$  over the extended vocabulary  $O_0$ , satisfying (12), and such that  $L_G = \mathbf{M}$ , i. e.  $\mathbf{M} \in \Omega_L$ ,
- (f)  $\mathbf{M}$  may be represented by a finite number of Ch. M. notations satisfying (13) over the vocabulary  $O_0$ ,
- (b)  $\mathbf{M} \in \Omega_0$  and  $\mathbf{O}_0 \notin \mathbf{M}$ .

Proof. (a)  $\Rightarrow$  (d) follows by 1.1 and 2.1, (d)  $\Rightarrow$  (e) follows by 3.1, (e)  $\Rightarrow$  (f) follows by [1], Th. 6 and 3.3, (f)  $\Rightarrow$  (b) follows by 3.2 and 1.1, and (b)  $\Rightarrow$  (a) follows by 1.1.

**3.5. Corollary.** *Theorem 6 in [1] may be inverted.*

The proof follows from 3.4 and 3.3.

**3.6. Corollary.** Let  $F_7$  be the class of all grammars the state diagrams of which satisfy (II) and let  $L(F_7)$  be the set of all  $L_G$ , where  $G \in F_7$ . Then  $L(F_7) = L(F_1)$  (cf. [1]).

Proof. It is evident that  $L(F_7) \subset L(F_1)$ . If there were  $L_G \in L(F_1)$ , then either  $L_G = \mathbf{M} \cup \mathbf{O}_0$  or  $L_G = \mathbf{M}$ , where  $\mathbf{M} \in \Omega$  (in virtue of 3.4 and 1.1). By the same construction as in the proof of 3.1 of a grammar  $G_0$  such that  $L_{G_0} = \mathbf{M}$ , we have constructed a state diagram  $\Gamma_0$  satisfying (I)–(III), i. e.  $G_0 \in F_7$ . In the second case  $L_G \in L(F_7)$  and in the first case we add to  $\Gamma_0$  a new vertex  $y$  and two labelled edges  $[v_0, y] o_0$  and  $[y, v_0] o_0$ , obtaining a new state diagram  $\Gamma'_0$  of some grammar  $G'_0$  satisfying (II), of course. Therefore  $L_{G'_0} = L_G$  and  $G'_0 \in F_7$ , i. e.  $L_G \in L(F_7)$ .

**3.7. Corollary.** The set of all finite state languages is not changed if in the definition of generation of strings (or sentences, see [1], p. 95) the condition (ii) is omitted.

The proof follows directly from 3.4.

**3.8. Lemma.** If  $\mathbf{M} = \Gamma(X, Y)$ , then there exist  $\Gamma', p' \in V'$  and  $Q' \subset V'$  such that  $\Gamma'(p', Q') = \mathbf{M}$  and  $\Gamma'$  satisfies conditions (I) and

$$(V) \quad [u, v] o_k, [u, w] o_k \in \mathcal{E}' \Rightarrow v = w.$$

Proof. By 3.4 there exists a grammar  $G_0$  satisfying (12) such that  $\mathbf{M} = L_{G_0}$  and whose state diagram is  $\Gamma_0$  and initial state  $v_0$ . By [1], Theorems 1, 2, 3 we may suppose that  $G_0 \in F_4$  (see [1], p. 98), i. e.  $\Gamma_0$  satisfies (I), (V) and the conditions  $[u, v] o_0 \in \mathcal{E}_0 \Rightarrow v = v_0$  and  $[u, v_0] o_k \in \mathcal{E}_0 \Rightarrow o_k = o_0$ . Now we define  $V' = V_0$ ,  $\mathcal{E}' = \mathcal{E}_0 - \{[x, v_0] o_0 \mid x \in V_0\}$ ,  $p' = v_0$ ,  $Q' = \{x \mid x \in V_0, x \neq v_0, [x, v_0] o_0 \in \mathcal{E}_0\}$ . Then  $\Gamma'$  satisfies (I) and (V),  $\Gamma'$  is a labelled graph over the vocabulary  $O$  and evidently  $\Gamma'(p', Q') = \mathbf{M}$ .

**4. Finite determinated and indeterminated automata.** A finite automaton  $\mathcal{A} = (S; I_1, \dots, I_n)$  in the sense of Medvedev (see [3]) is a finite set  $S$  of states with a finite set  $I_1, \dots, I_n$  of mappings of  $S$  into  $S$  (so-called *inputs*). This automaton may be called *determinated*, because by choosing a state  $x \in S$  and an input  $I_j$  there is uniquely determined next following state  $I_j(x) \in S$ . In [3], to every  $I_j$  there is associated a symbol  $o_j$ ,  $1 \leq j \leq n$ , i. e. to every set of finite sequences of inputs there is a uniquely defined event over the vocabulary  $O$  (cf. section 1). Now, an event  $\mathbf{M}$  over  $O$  is *representable* (by a finite determinated automaton), if there exist an  $\mathcal{A} = (S; I_1, \dots, I_n)$ , an initial state  $x^0 \in S$  and a set of final states  $X^0 \subset S$  such that

$$(14) \quad \alpha = o_{a_1} \dots o_{a_p} \in \mathbf{M} \Leftrightarrow I_{a_p} I_{a_{p-1}} \dots I_{a_1}(x^0) \in X^0.$$

In [3] this is denoted by  $\mathbf{M} = \mathcal{A}(x^0, X^0)$ .

Let  $\Omega_A$  be the set of all events representable by finite determinated automata.

**4.1. Lemma.** Let  $\mathbf{M} = \mathcal{A}(x^0, X^0)$ , where  $\mathcal{A} = (S; I_1, \dots, I_n)$ . If we define  $V = S$ ,  $\mathcal{E} = \{[u, v] o_j \mid u, v \in S, I_j(u) = v, 1 \leq j \leq n\}$ , then  $\Gamma(x^0, X^0) = \mathbf{M}$ , and  $\Gamma$  satisfies conditions (I), (V) and

$$(VI) \quad o_j \in \mathbf{O}, \quad u \in V \Rightarrow \text{there exists a } v \in V \text{ such that } [u, v] o_j \in \mathcal{E}.$$

The proof is obvious.

**4.2. Lemma.** Let  $\Gamma'$  be an labelled graph satisfying (V) over the vocabulary  $O$ . Then there may be defined a labelled graph  $\Gamma$  over the vocabulary  $O$  ( $V' \subset V$ ), satisfying (V) and (VI) and  $\Gamma(x, Y) = \Gamma'(x, Y)$  for all  $x \in V'$ ,  $Y \subset V'$ , as follows: 1) if  $\Gamma'$  satisfies (VI) we put  $\Gamma = \Gamma'$ ; 2) if  $\Gamma'$  does not satisfy (VI), we take a new vertex  $w \text{ non} \in V'$  and define  $V = V' \cup \{w\}$ ,  $\mathfrak{E} = \mathfrak{E}' \cup \{[u, w] o_j \mid [u, v] o_j \text{ non} \in \mathfrak{E}' \text{ for all } v \in V'\}$ .

The proof is obvious.

**4.3. Theorem.** If  $M$  is an event over the vocabulary  $O$ , then the following conditions are equivalent: (c)  $M \in \Omega$ , (g)  $M \in \Omega_A$ .

Proof. (c)  $\Rightarrow$  (g) If  $M = \Gamma(P, Q)$ , by 3.8 we may suppose that  $\Gamma$  satisfies  $P = \{p\}$ , (I) and (V), and by 4.2 also (VI). Now we define  $\mathcal{A} = (S; I_1, \dots, I_n)$  as follows:  $S = V$ ,  $I_j(u) = v$  if  $[u, v] o_j \in \mathfrak{E}$  (obviously by (VI) and (V)  $I_j$  is the required mapping).

(g)  $\Rightarrow$  (c) follows by 4.1.

**4.4. Corollary.** The set  $\Omega_A$  is not changed if either 1) by input  $I_j$  we understand a mapping of some subset  $S_j \subset S$  into  $S$  ( $S_j \neq \emptyset$ ), or 2) we always choose the initial state  $x^0$  and the set of final states  $X^0$  in such a manner that  $x^0 \text{ non} \in X^0$ , or 3) we use a set of initial states instead of a single initial state.

Proof. 1) follows from 4.2 and 4.3, 2) follows from 4.3 and 2.5 and 3) follows from 4.3 and 2.6.

Now, in the notion of determinated automaton  $\mathcal{A} = (S; I_1, \dots, I_n)$  of Medvedev instead of the usual mappings  $I_j$ , choose many-valued mappings  $f_i$ ,  $1 \leq i \leq n$ . These many-valued mappings  $f_i$  might be called *random mappings*, because if we admit repeated use of them (possibly in time), then in different cases  $f_i(x)$ ,  $x \in S$  will denote different elements of  $S$  (of course the notation  $f_i(x)$  is not suitable). Therefore a system  $\mathcal{B} = (S; f_1, \dots, f_n)$  may be called an *indeterminated automaton*. To every many-valued mapping  $f_i$  there may be associated an ordinary mapping  $F_i$  of  $S$  into the set of all subsets of  $S$  in the following manner:  $F_i(x) = S_i(x)$ , where  $S_i(x) \subset S$  is the set of all "possible" values  $f_i(x)$  (by repetition). To every  $f_i$  we associate the symbol  $o_i \in O$ ,  $1 \leq i \leq n$ , and say that  $M$  (an event over the vocabulary  $O$ ) is *representable* by a finite indeterminated automaton if there exist  $\mathcal{B} = (S; f_1, \dots, f_n)$  and  $x^0 \in S$ ,  $X^0 \subset S$  such that

$$(15) \quad \alpha = o_{a_1} \dots o_{a_p} \in M \Leftrightarrow f_{a_1} f_{a_{p-1}} \dots f_{a_1}(x^0) \in X^0.$$

In this case we shall write  $M = \mathcal{B}(x^0, X^0)$ .

Let  $\Omega_B$  be the set of all events which are representable by a finite indeterminated automata.

Now if  $M = \mathcal{B}(x^0, X^0)$ , we define a labelled graph  $\Gamma$  over the vocabulary  $O$  as follows:  $V = S$ ,  $\mathfrak{E} = \{[u, v] o_j \mid v \in F_i(u) \text{ (i. e. it is "possible" that } f_i(u) = v) \text{ for all } u, v \in S \text{ and } 1 \leq i \leq n\}$ . Then obviously  $\Gamma$  satisfies (I) and  $\Gamma(x^0, X^0) = M$ .

If on the contrary we have a labelled graph  $\Gamma$  over the vocabulary  $O$  ( $V = S$ ), satisfying (I), then we define mappings  $F_i$  as follows:  $F_i(u) = \{v \in S \mid [u, v] o_i \in \mathfrak{E}\} = S_i(u)$  for all  $i$ ,  $1 \leq i \leq n$  and all  $u \in S$ . Random mappings  $f_i$  are uniquely determi-

ned by the  $F_i$  and therefore an indeterminated automaton  $\mathcal{B} = (S; f_1, \dots, f_n)$  is defined, such that  $\mathcal{B}(x^0, X^0) = \Gamma(x^0, X^0)$  for all  $x^0 \in S, X^0 \subset S$ . We have thus proved

**4.5. Theorem.** *If  $M$  is an event over the vocabulary  $O$ , the following conditions are equivalent: (c)  $M \in \Omega$  and (h)  $M \in \Omega_B$ .*

#### References

- [1] *N. Chomsky, G. A. Miller: Finite State Languages, Information and Control 1 (1958), 91—112.*
- [2] *S. C. Kleene: Представление событий в нервных сетях и конечных автоматах, Автоматы (Russian Translation), 1956, 15—69.*
- [3] *Ю. Т. Медведев: О классе событий, допускающих представление в конечном автомате, Автоматы, 385—401.*
- [4] *P. C. Rosenbloom: The Elements of Mathematical Logic, 1950.*

#### Výtah

### NĚKOLIK POZNÁMEK S HLEDISKA POJMENOVANÝCH GRAFŮ O JAZYCÍCH S KONEČNĚ MNOHA STAVY A O JEVECH REPRESENTOVATELNÝCH KONEČNÝMI AUTOMATY

KAREL ČULÍK, Brno

Teorie pojmenovaných grafů je využito k důkazu ekvivalence pojmu jazyka s konečně mnoha stavy (N. CHOMSKY-G. A. MILLER [1]) s pojmem jevu representovatelného konečnými automaty (S. C. KLEENE [2] a JU. T. MEDVEDĚV [3]). Dále jsou uvedeny některé možnosti zobecnění nebo specialisace pojmů gramatiky s konečně mnoha stavy a pojmu konečného automatu (zejména je zaveden pojem nedeterminovaného konečného automatu).

#### Резюме

### НЕСКОЛЬКО ИСПОЛЬЗУЮЩИХ ТЕОРИЮ ГРАФОВ ЗАМЕЧАНИЙ О ЯЗЫКАХ С КОНЕЧНЫМ ЧИСЛОМ СОСТОЯНИЙ И О СОБЫТИЯХ, ПРЕДСТАВИМЫХ В КОНЕЧНЫХ АВТОМАТАХ

КАРЕЛ ЧУЛИК (Karel Čulík), Brno

Теория графов использована для доказывательства эквивалентности понятия языка с конечным числом состояний (N. CHOMSKY-G. A. MILLER [1]) с понятием события, допускающего представление конечными автоматами (S. C. KLEENE [2] и Ю. Т. Медведев [3]). Далее указаны некоторые возможности генерализации и специализации понятия грамматики с конечным числом состояний и понятия конечного автомата (именно введено понятие конечного недетерминированного автомата).