MTA Számitástechnikai és Automatizálási Kutató Intézete, **Közlemények** 19/1978.

THE LATTICE OF SYMMETRIC LANGUAGES INVARIANT UNDER INNER LINEAR TRANSFORMATIONS *

Bagyinszki János – Demetrovics János
MTA KFKI – MTA SzTAKI

Let $V = \{1,2,\ldots,k-1\}$ for $k \geqslant 2$, $V_0 = V \cup \{0\}$ a finite alphabet and L a language over $V_0 : L \subseteq V_0^*$. Notations: $V_0^* = V_0^+ \cup \{\varepsilon\}$, $V_0^+ = \bigcup_{n=1}^{\infty} V_0^n$, $V_1 \cdot V_2 = \{\nu_1\nu_2 \mid \nu_1 \epsilon V_1, \nu_2 \epsilon V_2\}$, and "$\varepsilon$" is the empty sentence. A language L is termed <u>symmetric</u>, if it contains the word $a_0 a_{\pi(1)} \ldots a_{\pi(n)}$ for any $a_0 a_1 \ldots a_n \epsilon L \subseteq V_0^+$ and for any permutation $\pi \epsilon S_n$ over the index-set $N = \{1,2,\ldots,n\}$. The class of symmetric languages (S-languages) over the alphabet V_0:

$$\mathscr{I} = \{L \mid (a_0 a_1 \ldots a_n \epsilon L \subseteq V_0^+) \text{ iff } (\forall \pi \epsilon S_n)(a_0 a_{\pi(1)} \ldots a_{\pi(n)} \epsilon L)\}.$$

A language $L \subseteq V_0 V^*$ is said to be <u>invariant under inner linear transformations</u> (IL-language) if it is closed under the following two operations 0_1 and 0_2 $(a_0 a_1 \ldots a_n, b_0 b_1 \ldots b_m \epsilon L)$:

1.) $0_1(a_0) = a_0$, $0_1(a_0 a_1) = a_0 a_1$,

$$0_1(a_0 a_1 \ldots a_n) = a_0 \ldots a_{n-2} a' \text{ for } n \geqslant 2,\ a' = \begin{cases} a_{n-1} + a_n, & \text{if } a_{n-1} + a_n \neq 0; \\ \varepsilon, & \text{if } a_{n-1} + a_n = 0. \end{cases}$$

2.) $0_2(a_0, b_0 b_1 \ldots b_m) = a_0$, $0_2(a_0 a_1, b_0 b_1 \ldots b_m) = (a_0 + c_{10}) c_{11} c_{12} \ldots c_{1m}$,

$$0_2(a_0 a_1 \ldots a_n, b_0 b_1 \ldots b_m) = (a_0 + c_{n0}) a_1 \ldots a_{n-1} c_{n1} \ldots c_{nm}, \quad \text{for } n \geqslant 2,$$

$$c_{ij} = \begin{cases} a_i \cdot b_j, & \text{if } j = 0 \text{ or } a_i \cdot b_j \neq 0; \\ \varepsilon, & \text{if } j \neq 0 \text{ and } a_i \cdot b_j = 0. \end{cases}$$

(the addition "+" and multiplication "·" are carried out mod k in this lecture.)

The class of IL-languages over the alphabet V_0 is

$$\mathscr{J} = \{L \mid \text{if } \underline{a}, \underline{b} \ \epsilon L \subseteq V_0 V^*, \text{ then } 0_1(\underline{a}),\ 0_2(\underline{a}, \underline{b}) \epsilon L\}.$$

The main purpose of this lecture is to investigate the class of symmetric languages invariant under inner linear transformation (SIL-languages):

$$\mathscr{L} = \mathscr{S} \cap \mathscr{J}$$

* Presented at the Internat. Colloquium on Automata and Formal Languages, Szeged, (Hungary), Aug 30 – Sept. 2, 1977.

Results:

a.) The complete lattice-structure of $\mathscr{L}$ and therefore the exact (finite) cardinality of $\mathscr{L}$ are presented for $k = p$ prime number.

b.) The base and the rank of each languages $L \in \mathscr{L}$ are given.

c.) The elements of $\mathscr{L}$ are generable by regular grammars.

d.) The correspondence between $\mathscr{L}$ and the class of linear functions on the set V_0 is presented.

Remark: A. Salomaa presented in [6] very impressive results concerning closed sets of sequences over the set V_0. Still, he defined the closedness of a set in a way according to Malcev-algebras and thus a little different from ours. Besides, his essential results concern the case of infinite cardinalities. Moreover, according to his definition, it is not languages what he deals with. In the case $k = p$ with p being a prime number he presents the sets corresponding to the sets $L, L_\alpha, L_\Delta, L_{\Delta 0}, L^{(1)}$ with the remark that because of the finiteness of $L^{(1)}$ the cardinality in question must be finite as well.
A word $a_0 a_1 \ldots a_n$ can be interpreted as a linear polynominal over $GF(p)$; by the correspondence $a_0 a_1 \ldots a_n \longleftrightarrow a_0 + a_1 x_1 + \ldots + a_n x_n$ a connection between many-valued logics and our results is presented. Some significant results on many-valued logics are tabelled as follows:

Structure of 2-valued logics (Post-lattice)	– E. Post, 1921. [4]
All precomplete subsets in P_3	– Sz. V. Jablonszkij 1953. [7]
Closed and precomplete subsets in P_k	– Sz. V. Jablonszkij, 1958. [7]
All precomplete subsets in P_k	– I. Rosenberg, 1965. [5]
Closed subsets { infinitely generated in P_k $(k \geqslant 3)$ / without bases	– Ju. I. Janov, A.A. Muchnik, 1959. [8]
Maximal and precomplete sets in $L(k)$	– A. Salomaa, 1964. [6]
Lattice of $SIL(p)$-languages	– J. Bagyinszki, J. Demetrovics, 1976. [1], [2].

Table 1.

It can be checked, that the following sets are SIL-languages (notations: $a_0 \epsilon V_0$, $a_i \epsilon V$ for $i \geqslant 1$, $\sum_{i=1}^{n} a_i = a$, $\alpha \epsilon V_0$):

$$L(k) = \{ a_0 a_1 \ldots a_n \mid n = 0,1,2, \ldots \} = V_0 \cdot V^*$$

$$L_\Delta = \{ a_0 a_1 \ldots a_n \mid a = 1 \}$$

$$L_\alpha = \{ a_0 a_1 \ldots a_n \mid a_0 = \alpha(1 - a),\ n \geqslant 1 \} \cup \{ \alpha \}$$

$$L^{(1)} = V_0 \cup V_0 V$$

$$L^{(0)} = V_0$$

$$L^{(1)} \setminus L^{(0)} (= V_0 V)$$

$$L_{\Delta\alpha} = L_\Delta \cap L_\alpha = L_{\Delta 0}$$

$$L_\Delta^{(1)} = L_\Delta \cap L^{(1)}$$

$$L_\alpha^{(1)} = L_\alpha \cap L^{(1)} = \{ a_0 a_1 \mid a_0 = \alpha(1 - a_1) \} \cup \{ \alpha \}$$

$$L_\alpha^{(1)} \setminus \{ \alpha \} (= L_\alpha \cap (L^{(1)} \setminus L^{(0)}))$$

$$L_\alpha^{(0)} = L_\alpha \cap L^{(0)} = \{ \alpha \}$$

$$L_\alpha^{(1)} \cup L^{(0)}.$$

Theorem 1: *If* $k = p$ *(prime), then* $< \mathscr{L} \cup \{ \emptyset \}, \subseteq >$ *is a* _finite_ *lattice, with the unit element* $L(p)$ *and zero element* $\emptyset$ *(empty set).*

The linear sublattice of the known Post-lattice ($k = 2$) is given for an example on Fig. 1.

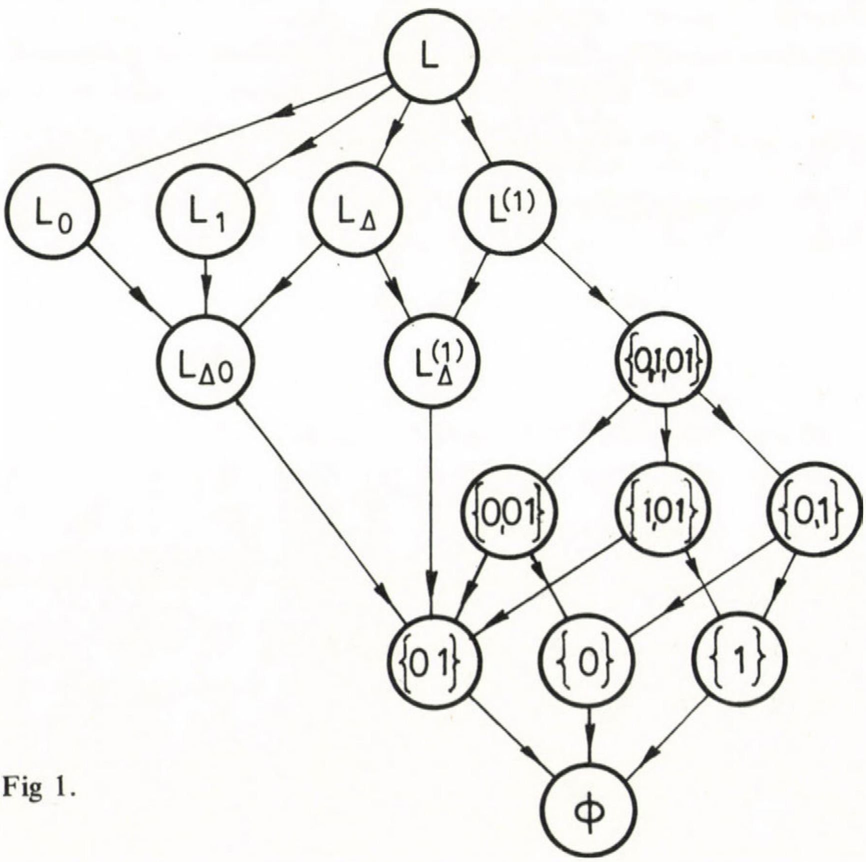

Fig 1.

In the rest of the paper k is supposed to be a prime number $p \geqslant 3$.

Let $[A] \in \mathcal{L}$ be the least language in $\mathcal{L}$ containing A (thus, $A \subseteq \bar{A} \subseteq [A]$, $\bar{A}, [A] \in \mathcal{L}$ implies $\bar{A} = [A]$.). The set $B \subseteq L' \in \mathcal{L}$ is a base of L', if $[B] = L'$ and $B' \subseteq B$, $[B'] = L'$ implies $B' = B$.
"A" is called to be complete in $L' \in \mathcal{L}$, if $[A] = L'$.

Let $L', L'' \in \mathcal{L}$. L'' will be called precomplete in L', if $L'' \subset A \subseteq L'$ implies $[A] = L'$.

Let L' be a *SIL*-language. To prove that all the precomplete *SIL*-languages in L' are given, we need bases of the *SIL*-languages. Bases of $L(p)$, precomplete languages in $L(p)$ and their bases are given in theorems 2 – 6. (without proofs). This is the first level in the lattice $< \mathcal{L} \cup \{\emptyset\}, \subseteq >$.

Theorem 2. *The following sets are bases in* $L(p)$:

(a) $\{ 011, 11 \}$.

(b) $\{ a_0 a_1 a_2, b_0, c_0 \}$; $a = 1$, $a_0 = 0$, $b_0 \neq c_0$.

(c) $\{ a_0 a_1 a_2, b_0 \}$; $a = 1$, $a_0 \neq 0$.

(d) $\{ a_0 a_1 a_2, b_0 \}$; $a \neq 1$, $b_0 \neq (p - a_0)(a-1)^{p-2}$.

Theorem 3.

(1) *Languages* L_α *are precomplete in* $L(p)$, $\alpha = 0, 1, \ldots, p-1$.

(2) *The language* L_Δ *is a precomplete in* $L(p)$.

(3) *The language* $L^{(1)}$ *is a precomplete in* $L(p)$.

Theorem 4.

(a) *The set of base-functions (bases with one element each) in the language* L_α *is:* $L_\alpha \setminus (L_\Delta \cup L^{(1)})$.

(b) *The set of base functions in the set* L_Δ *is*: $L_\Delta \setminus (L_{\Delta 0} \cup L^{(1)})$.

(c) *The set of base functions in the set* $L_{\Delta 0}$ *is*: $L_{\Delta 0} \setminus L^{(1)}$.

Let $c_0 \in V_0$, $B = \{ a_{10} a_{11}, a_{20} a_{21}, \ldots, a_{s0} a_{s1} \}$, *and* $r_i = (a_{i1})$ *be the multiplicative order of* $a_{i1} \in V$.

Theorem 5.

A.) *The following statements are equivalent*:

(1) *The set* B *is a base in the language* $L^{(1)} \setminus L^{(0)}$.

(2) *The set* $B_0 = B \cup \{c_0\}$ *is a base in the language* $L^{(1)}$.

(3) *For elements of the set* B *are valid*:

(a) $l.c.m. \{r_1, \ldots, r_s\} = p-1$.

(b) $B \setminus L_\alpha^{(1)} \neq \emptyset \quad \alpha = 0,1,\ldots,p-1$.

(c) *if* $B' \subset B$, $B' \neq B$, *then* (a) *and* (b) *cannot hold for* B' *at the same time.*

B.) *If* B *is a base of* $L^{(1)} \setminus L^{(0)}$, *then* $|B| \geqslant 2$, $|B_0| \geqslant 3$.

Theorem 6.

Every language $L \in \mathscr{L}$ *different from* $L(p)$ *is a subset at least in one of the precopmlete languages* $L_0, L_1, \ldots, L_{p-1}, L_\Delta, L^{(1)}$.

Languages of the next level can be determined in a similar way. Results are presented only in a more compact form on Fig. 2. giving the structure of the lattice $< \mathscr{L} \cup \{\emptyset\}, \subset >$. If the language L'' is precomplete in $L' \in \mathscr{L}$, then L'' is of the next level and there is an edge connecting it with L'.

It can be seen that the set $L^{(1)} \setminus L^{(0)}$ constitutes a group of order $p(p-1)$ with respect to the operation 0_2. Let $p-1 = q_1^{\kappa_1} \ldots q_u^{\kappa_u}$ be the canonical decomposition of $p-1$, where $2 = q_1 < q_2 < \ldots < q_u$ are prime numbers, $\kappa_i \geqslant 1$, $p_i = \frac{p-1}{q_i}$ and $L^{(1,i)} = \{a_0 a | r(a) \text{ divides } p_i\}_{i=1,2,\ldots u}$ ($r(a)$ is the order of "a" in the group $< V, \cdot >$).

To complete the structure of $L^{(1)}$ it needs, for example, the following statements:

(1) The group $L_\Delta^{(1)}$ is contained in a subgroup G of the group $L^{(1)} \setminus L^{(0)}$, if and only if the order of G, $|G| \geqslant p$.

(2) The subgroup $G \subseteq L^{(1)} \setminus L^{(0)}$ of order $|G| \leqslant p-1$ is cyclic, G is a subgroup of $L_\alpha^{(1)} \setminus \{\alpha\}$ for some suitable α.

The next theorem involves the result on cardinality cited in theorem 1.

Theorem 7.

(1) $\mathscr{L}$ has the cardinality

$$|\mathscr{L}| = p + 2 - (p-2)2^{p} + 2d(p-1) + 2p \cdot \sum_{e|p-1} 2^e.$$

(2) $\mathscr{L}$ has maximal and minimal chains of lenght $p + 2 + \sum_{i=1}^{u} \kappa_i$ and 3, respectively.

At last, we shall show that $\mathscr{L}$ is a subclass in the class of regular languages . Thus we shall describe the regular grammars which generate the elements of $\mathscr{L}$, and finite accepting automata. The grammars G for languages L, L_Δ, $L_{\Delta 0}$ and L_α are given as follows. Let $G = (K, V_0, P, A_0)$ be with non-terminals K, terminals V_0, productions P.

$$L: K_L = \{A_0, A_1\} \quad , P_L = \bigcup_{\substack{i \in V \\ j \in V_0}} \{A_0 \to j, A_0 \to jA_1, A_1 \to i, A_1 \to iA_1\}.$$

$$L_\Delta - L_{\Delta 0}: K_0 = \{A_0, A, A_1, \ldots, A_{p-1}\},$$

$$P_\Delta = \bigcup_{\substack{i,j \in V \\ j_0 \in V_0}} \{A_0 \to j_0 A, A \to 1, A \to iA_i, \quad A_i \to p - i + 1, A_i \to jA_{i+j}\}.$$

$$P_{\Delta 0} = P_\Delta \setminus \{A_0 \to jA \mid j \in V\}.$$

$$L_\alpha: K = \{A_0, A_1, \ldots, A_{p-1}, \quad B_0, B_1, \ldots, B_{p-1}\}$$

$$P_\alpha = \bigcup_{\substack{i,j,m \in V \\ j_0 \in V_0}} \{A_0 \to \alpha, A_0 \to jB_{j \cdot \beta}, B_m \to p - m + 1, B_m \to i \cdot A_{m+i}, A_i \to p - i + 1, A_i \to jA$$

$$\beta = \alpha^{p-2}.$$

It is clear, that for finite languages there are accepting finite automata.

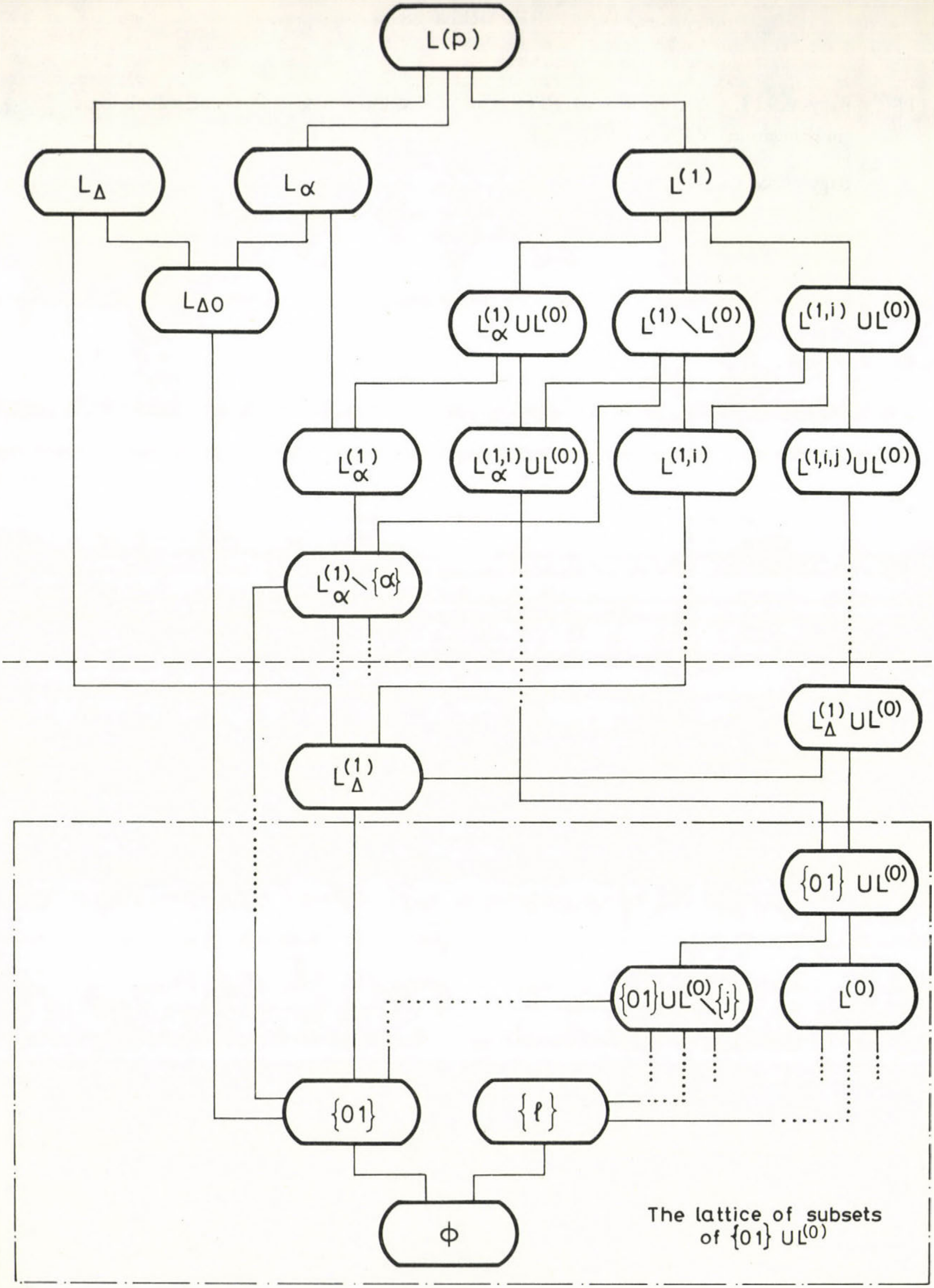
L(p)
L_Δ
L_α
L^(1)
L_Δ0
L_α^(1) U L^(0)
L^(1) \ L^(0)
L^(1,i) U L^(0)
L_α^(1)
L_α^(1,i) U L^(0)
L^(1,i)
L^(1,i,j) U L^(0)
L_α^(1) \ {α}
L_Δ^(1) U L^(0)
L_Δ^(1)
{01} U L^(0)
{01} U L^(0) \ {j}
L^(0)
{01}
{ℓ}
φ
The lattice of subsets
of {01} U L^(0)

Fig. 2.

REFERENCES

[1] Bagyinszki J. – Demetrovics J. : The structure of linear classes in prime valued logics (in Hungarian) MTA SzTAKI Közlemények, 16/1976, 25-52.

[2] Bagyinszki J. – Demetrovics J. : The structure of the class of symmetric languages invariant for inner linear transformations, in Proc.s of second Hung. Comp. Sci. Conf. 100-130. (1977, Budapest)

[3] Hopcroft J.B. – Ullman J.D. : Formal languages and their relation to automata, 1969. Addison – Wesley

[4] Post, E. : The two-valued iterative systems of mathematical logic, Annals of Math. Studes 5 (1941).

[5] Rosenberg, I. : La structure des fonktions de plusieurs variables sur un ensemle fini. C.R. Acad. Sci. Paris Ser A.B 260 (1965) 3817-19.

[6] Salomaa, A.; On infinitely generated sets of operations in finite algebras. Ann. Univ. Turkuensis, ser. A.I. 74. (1964) 1-13.

[7] Jablonszkij, S.V. : Functional constructions in *k*-valued logics (Russian) Trudy Mat. - Inst. Steklov 51 (1958) 5-142.

[8] Janov, Ju.I., Muchnik, A.A.: Existence of *k*-valued closed classes having no finite basis (Russian), Dokl.Akad.Nauk. SSSR. 127 (1959) 44-46.

Összefoglaló

A belső lineáris transzformációra invariáns szimmetrikus nyelvek hálója

Bagyinszki János – Demetrovics János

A V_0 alaphalmazon definiált L nyelvet szimmetrikusnak nevezzük, ha $a_0 a_1 a_2 \ldots a_n \epsilon L$ esetén $a_0 a_{\pi(1)} a_{\pi(2)} \ldots a_{\pi(n)} \epsilon L$ is igaz minden $\pi(x)$ permutációra. Egy nyelv invariáns a belső lineáris transzformációra nézve, ha zárt a dolgozatban definiált 0_1 és 0_2 operációkra. A szimmetrikus és a belső lineáris transzformációra invariáns nyelveket SIL-nyelveknek nevezzük. A jelen dolgozatban azokat SIL nyelveket tanulmányozzuk, amelyekben a V_0 alaphalmaz számossága primszám.

Dolgozatunkban leirjuk a SIL(p) nyelvek osztályának teljes szerkezetét, megadjuk a tartalmazási reláció által indukált hálót, a minimális bázisokat és a pontos elemszámot. Megmutatjuk, hogy a SIL(p)- nyelvek mindegyike reguláris és utalunk a többértékü logika lineáris függvény--osztályaival való kapcsolatra.

Р Е З Ю М Е

Структура симметрических языков, инвариантных по отношению к внутренним линейным трансформациям.

Янош Бадински - Янош Деметрович

Назовем язык L симметрическим над алфавитом $V_0=\{0,1,2,\ldots,\ldots,k-1\}$, если для каждой перестановки $\P(x)$ из $a_0a_1a_2\ldots a_n \in L$ следует, что $a_0a_{\P(1)}\ldots a_{\P(n)} \in L$. Язык L является инвариантным поотношению к внутренним линейным трансформациям, если он замкнут относительно операций 0_1, 0_2 определенных в данной работе. Язык L называется SIL-языком, если он симметрический и инвариантен по отношению к внутренним линейным трансформациям. В настоящей работе мы исследовали SIL-язык, в случае когда мощность алфавита V_0 равна простому числу.

В работе описана полная структура SIL-языков относительно операции включения и охарактерезован каждый класс /SIL-языков/ в этой структуре. Кроме того, для каждого SIL-языка задан минимальный базис и приведена точная арифметическая формула для мощности SIL-языков /элементов в структуре/. Доказывается, что каждый SIL-язык является регулярным и, кроме того, указывается связь между линейными замкнутыми классами k-замкнутой логики и SIL-языками.