REVISTA DE ANALIZĂ NUMERICĂ ŞI TEORIA APROXIMAŢIEI
Volumul 2, Fascicola 2, 1973, pp. 159–166

RELAŢII DE INCLUZIUNE INTRE FAMILII DE LIMBAJE MARGINITE

de

C. JALOBEANU

(Cluj)

În lucrarea [1] M. K. YNTEMA defineşte în mod recursiv familii de limbaje acontextuale şi studiază relaţiile de incluziune care se pot stabili între ele. Pornind de la aceste rezultate, în lucrarea de faţă este studiat cazul familiilor de limbaje acontextuale supuse condiţiei de mărginire. Sînt definite familiile de limbaje acontextuale şi mărginite stabilindu-se relaţiile de incluziune strictă pe care acestea le verifică.

§ 1. Noţiuni introductive

Considerăm o mulţime finită T de simboluri pe care o numim alfabet. Şirurile de elemente din T se numesc cuvinte. Dacă u şi v sînt şiruri de elemente din T, uv este tot un şir obţinut prin concatenare. Şirul vid se va nota cu e. Notăm cu T^* mulţimea tuturor şirurilor de elemente din T. O submulţime din T^* se numeşte limbaj peste alfabetul T.

În mulţimea limbajelor peste T se definesc următoarele operaţii: dacă

$$L_1, L_2 \subseteq T^*,$$

atunci

$$L_1 \cup L_2 = \{u | u \in L_1 \text{ sau } u \in L_2\}$$

$$L_1 L_2 = \{uv | u \in L_1 \text{ şi } v \in L_2\}$$

$$L^* = \bigcup_{n=0}^{\infty} L^n, \text{ unde } L^n = LL^{n-1},\ L^\circ = \{e\}.$$

Definiţia 1.1. *Se numeşte gramatică acontextuală un sistem* $G = (V, T, P, \sigma)$, *unde*
V *este un alfabet*
T *este un alfabet*
P *submulţime finită din* $V \times (V \cup T)^*$
σ *element din* V.

Definiţia 1.2. *Se numeşte gramatică regulară o gramatică aco[ntex]tuală în care* $P \subseteq V \times T^*V \cup T^*$

Vom nota cu $\underset{G}{\Rightarrow}$ o relaţie binară în $(V \cup T)^*$ astfel încît $u \underset{G}{\Rightarrow}$ dacă şi numai dacă $u = u_1 p u_2$, $w = u_1 r u_2$, $u_1, u_2 \in (V \cup T)^*$, $(p, r) \in P$.

Notăm $u \overset{*}{\underset{G}{\Rightarrow}} w$, dacă şi numai dacă există un şir $v_1, v_2, \ldots, v_m$, $v_i \in (V \cup T)^*$ cu proprietatea $v_1 = u$, $v_m = w$ şi $v_i \underset{G}{\Rightarrow} v_{i+1}$, $i = 1, 2, \ldots, m - 1$.

Definiţia 1.3. *Spunem că un limbaj* $L(\sigma)$ *este generat de o gramatică* $G = (V, T, P, \sigma)$ *dacă* $L(\sigma) = \{v \in T^* | \sigma \overset{*}{\underset{G}{\Rightarrow}} v\}$.

Dacă G este o gramatică acontextuală, limbajul se numeşte acontextual; dacă G este o gramatică regulară, limbajul se numeşte regular.

Familiile de limbaje acontextuale studiate în [1] sînt definite recursiv cu ajutorul unei operaţii speciale numită de alegere. Pentru a defini operaţia de alegere sînt necesare următoarele operaţii cu mulţimi de perechi de cuvinte.

Dacă S_1 şi S_2 sînt două mulţimi de perechi de cuvinte,

$$S_1 \cup S_2 = \{(x, y) | (x, y) \in S_1 \text{ sau } (x, y) \in S_2\}$$

$$S_1 S_2 = \{(x_2 x_1, y_1 y_2) | (x_1, y_1) \in S_1 \text{ şi } (x_2, y_2) \in S_2\}$$

$$S_1^* = \{(x_m \ldots x_1, y_1 \ldots y_m) | (x_i, y_i) \in S_1, \quad i = 1, 2, \ldots, m, \quad m \geq 0\}$$

Evident, dacă $S_1 = (x, y)$, atunci $(x, y)^* = \{(x^n, y^n) | n \geq 0\}$.

Fie S o mulţime de perechi de cuvinte şi C o mulţime oarecare de cuvinte, se defineşte următoarea operaţie numită operaţie de alegere

$$S \circ C = \{xcy | (x, y) \in S, \ c \in C\}$$

Mulţimea C şi mulţimile A, B cu proprietate $S = A \times B$ se numesc mulţimi de bază pentru operaţia de alegere.

Dacă de exemplu, $S = (x, y)$ şi $C = \{e\}$,

$$(x, y)^* \circ e = \{x^n y^n | n \geq 0\}.$$

Definiţia 1.4. *Se numeşte limbaj regular de alegere un limbaj care poate fi obţinut prin operaţia* $S \circ C$ *în care mulţimile de bază sînt limbaje regulare.*

Vom nota în continuare cu S_0 familia limbajelor regulare şi cu S_1 familia limbajelor regulare de alegere.

Definiţia 1.5. *Se numeşte limbaj metaliniar un limbaj obţinut ca reuniune finită de concatenări finite de limbaje regulare de alegere. Familia limbajelor metaliniare va fi notată* S_2.

Definiţia 1.6. *Un limbaj din familia* S_0 *se numeşte limbaj standard de rang* 0.

Un limbaj standard de rang n se obţine prin operaţia de alegere în care mulţimile de bază sînt de rang cel mult $n - 1$, una cel puţin fiind chiar de rang $n - 1$. Notăm această familie cu S_3.

Dacă cu S_4 notăm familia limbajelor acontextuale, rezultatul lui YNTEMA [1] poate fi enunţat sub următoarea formă.

TEOREMA 1.1. *Între familiile de limbaje acontextuale definite există următoarele relaţii de incluziune*

$$S_0 \subsetneqq S_1 \subsetneqq S_2 \subsetneqq S_3 \subsetneqq S_4.$$

Definiţia 1.7. *Mulţimea* $M \subset T^*$ *se numeşte mărginită dacă există cuvintele* $w_1, w_2, \ldots, w_n$, *definite peste alfabetul* T *cu proprietatea* $M \subseteq w_1^* w_2^* \ldots w_n^*$

Problema care se pune este de a stabili ierarhia dintre clasele de limbaje atunci cînd ele sînt supuse condiţiei de mărginire.

În continuare ne vom ocupa numai de limbaje acontextuale pe care le vom numi simplu limbaje.

§ 2. Limbaje mărginite

În acest paragraf prezentăm cîteva dintre rezultatele privind limbajele mărginite datorate lui S. GINSBURG şi E. H. SPANIER [2].

Propoziţia 2.1. *Dacă* $T = \{a_1, a_2\}$, *orice limbaj* $M \subseteq a_1^* a_2^*$ *este o reuniune finită de mulţimi de forma*

$$[(x_m, y_m), \ldots, (x_1, y_1)]^* \circ Z, \tag{1}$$

unde $x_i \in a_1^*$, $y_i \in a_2^*$, $Z \in a_1^* a_2^*$; *orice reuniune finită de mulţimi de forma* (1) *este un limbaj* $M \subseteq a_1^* a_2^*$.

Propoziţia 2.2. *Dacă* $T = \{a_i \mid 1 \leq i \leq n\}$, $n \geq 3$, *orice limbaj* $M \subseteq a_1^* \dots a_n^*$ *este reuniune finită de mulţimi de forma*

$$\{a_1^i \, xy \, a_n^j \mid a_1^i a_n^j \in D, \quad x \in E, \ y \in F\},$$

unde mulţimile

$$D \subseteq a_1^* a_n^*, \ E \subseteq a_1^* \dots a_q^*, \ F \subseteq a_q^* \dots a_n^*, \ 1 < q < n,$$

sînt limbaje.

Invers, orice reuniune finită de mulţimi de această formă este un limbaj $M \subseteq a_1^* \dots a_n^*$.

TEOREMA 2.1. *Familia limbajelor mărginite este cea mai mică familie de mulţimi care conţine mulţimile finite şi este închisă în raport cu reuniunea finită, concatenarea finită şi operaţia de alegere* $(x, y)^* \circ Z$, *unde* x *şi* y *sînt cuvinte iar* Z *este un limbaj mărginit.*

Demonstraţie. Dacă $A_1, \dots, A_r$ sînt limbaje mărginite, atunci este evident că reuniunea finită şi concatenările finite sînt tot limbaje mărginite. Dacă Z este limbaj mărginit şi x, y sînt cuvinte din T^*, atunci

$$(x, y)^* \circ Z = \bigcup_{n \geq 0} x^n Z y^n$$

este tot mărginit.

Înseamnă că aplicînd de un număr finit de ori aceste operaţii asupra unor mulţimi finite, care evident sînt mărginite, se obţin tot limbaje mărginite.

Pentru a arăta că familia este cea mai mică cu proprietatea enunţată, se demonstrează mai întîi că orice limbaj mărginit

$$M \subseteq a_1^* \dots a_n^*, \ a_1, \dots, a_n \in T,$$

se obţine din mulţimi finite prin aplicarea de un număr finit de ori a operaţiilor indicate. Demonstraţia se face prin inducţie asupra numărului de litere din alfabetul T. Propoziţia este evidentă pentru $n = 1$, deoarece $M \subseteq a_1^*$ este un limbaj regular, care după cum se ştie are proprietatea cerută.

Presupunem proprietatea adevărată pentru $k \geq 2$. Fie

$$M \subseteq a_1^* \dots a_{k+1}^*$$

un limbaj mărginit. Conform propoziţiei 2.2., M este o reuniune finită de forma

$$\bigcup_{a_1^i a_{k+1}^j \in D} a_1^i EF a_{k+1}^j,$$

unde

$$D \subseteq a_1^* a_{k+1}^*$$

şi pentru

$$1 < q < k + 1,$$

$$E \subseteq a_1^* \dots a_q^*, \ F \subseteq a_q^* \dots a_{k+1}^*$$

sînt limbaje. Dar atunci M este o reuniune finită de mulţimi de forma

$$[(x_m, y_m), \dots, (x_1, y_1)]^* \circ a_1^s EF a_{k+1}^t$$

$$x_i \in a_1^*, \quad y_i \in a_{k+1}^*.$$

Pe baza ipotezei E şi F sînt obţinute prin operaţiile considerate aplicate de un număr finit de ori asupra unor mulţimi finite, deci şi M se obţine în acelaşi mod.

Din cele de mai sus rezultă imediat că afirmaţia are loc şi pentru $M \subseteq w_1^* \dots w_n^*$, unde $w_1, \dots, w_n$ sînt cuvinte din T^*.

Definiţia 2.1. *O mulţime* X *de cuvinte se numeşte comutativă dacă oricare ar fi două cuvinte* $u \in X$ *şi* $v \in X$, *cuvintele obţinute prin concatenare coincid,* $uv = vu$.

TEOREMA 2.2. *Condiţia necesară şi suficientă ca limbajul* $L(G) \neq \emptyset$, *generat de gramatica redusă* $G = (V, T, P, \sigma)$, *să fie mărginit este ca mulţimile*

$$X_\xi = \{x \mid x \in T^*, \ \ \xi \overset{*}{\Rightarrow} x \xi y, \ \ y \in T^*\}$$

$$Y_\xi = \{y \mid y \in T^*, \ \ \xi \overset{*}{\Rightarrow} x \xi y, \ \ x \in T^*\}$$

să fie comutative, oricare ar fi $\xi \in V$.

Demonstraţie 1. Fie $L(G)$ limbaj mărginit. Notăm $W_\xi = \{w \in T^* \mid \xi \overset{*}{\Rightarrow} w\}$ Evident $W_\xi \neq \emptyset$ şi în plus pentru fiecare $\xi \in V$, există $u_\xi, v_\xi \in T^*$ astfel încît $u_\xi W_\xi v_\xi \subseteq L(G)$. Deci W_ξ este o mulţime mărginită.

Dacă presupunem că $u_1, u_2 \in X_\xi$ şi că $u_1 u_2 \neq u_2 u_1$, atunci există v_1, v_2 astfel încît $\xi \overset{*}{\Rightarrow} u_1 \xi v_1$, $\xi \overset{*}{\Rightarrow} u_2 \xi v_2$. Înseamnă că pentru orice $w \in \{u_1, u_2\}^*$ există w' din T^* astfel încît $\xi \overset{*}{\Rightarrow} w \, \xi \, w' \Rightarrow w w_\xi w'$, unde $w_\xi \in W_\xi$ şi deci $\{u_1, u_2\}^* - \{e\} \subseteq X_\xi$. Atunci fiecare cuvînt din $\{u_1, u_2\}^*$ este un subcuvînt din W_ξ şi deci W_ξ nu este o mulţime mărginită, în contradicţie cu presupunerea făcută.

2. Fie X_ξ și Y_ξ mulțimi comutative pentru fiecare variabilă ξ. S demonstrează că $L(G)$ este mărginită prin inducție asupra numărului d variabile.

§ 3. Familii de limbaje mărginite

TEOREMA 3.1. *Condiția necesară și suficientă ca un limbaj regular generat de gramatica $G = (V, T, P, \sigma)$ să fie mărginit este ca oricare ar fi un simbol $\xi \in V$ să există în P cel mult o producție de forma*

$$\xi \to x\xi, \quad x \in T^*, \quad x \neq e.$$

Demonstrație. Fie $\xi \to x\xi$ singura producție de acest tip, pentru un $\xi \in V$, atunci $X_\xi = \{x^+\}$, $Y_\xi = \{e\}$, unde $x^+ = x^* - \{e\}$. Evident, mulțimile X_ξ și Y_ξ sînt comutative, atunci pe baza teoremei 2.2. mulțimea regulară este mărginită.

Invers, fie $L(G)$ mulțime regulară mărginită. Să presupunem că în P ar exista două producții de forma $\xi \to a\xi$ și $\xi \to b\xi$, $\xi \in V$, $a, b \in T$. Atunci $X_\xi = \{a^+, b^+\}$, $Y = \{e\}$. Dar X_ξ nu este comutativă, deci $L(G)$ nu poate fi mărginită, în contradicție cu ipoteza făcută.

Consecință. Familia limbajelor regulare care sînt și mărginite este nevidă și există limbaje regulare care nu sînt mărginite.

Într-adevăr, limbajul generat de gramatica $G = (V, T, P, \xi)$, $V = \{\xi\}$,

$$T = \{a, b\}, \; P = \{\xi \to a\xi, \; \xi \to b\xi, \; \xi \to a, \; \xi \to b\}$$

nu este mărginit.

Notăm cu B_0 familia limbajelor regulare și mărginite.

D e f i n i ț i a 3.1. *Se numește limbaj regular de alegere mărginit peste alfabetul T o mulțime care se obține prin aplicarea operației $(x, y)^* \circ C$ asupra unei mulțimi regulare mărginite C, cînd x și y sînt cuvinte din T^*.*

Notăm cu B_1 familia limbajelor regulare de alegere și mărginite. Observăm că $B_1 \subsetneq S_1$. Într-adevăr, din definiție rezultă $B_1 \subset S_1$. În plus există limbaje în S_1 care nu sînt mărginite, obținute de exemplu prin operația de alegere $(x, y)^* \circ C$ aplicată unei mulțimi regulare C, nemărginite.

TEOREMA 3.2. *Orice limbaj regular mărginit este limbaj regular de alegere mărginit și există limbaje de alegere mărginite care nu sînt limbaje regulare mărginite. Adică $B_0 \subsetneq B_1$.*

Demonstrație. 1. Oricare ar fi R un limbaj regular și mărginit, poate fi pus sub forma: $R = (e, e)^* \circ R$. În acest fel este un limbaj regular de alegere și mărginit.

2. Fie $A = \{a^n b^n \mid n \geqslant 0\}$ un limbaj mărginit care, după cum se știe, nu este limbaj regular. Dar, $A = (a, b)^* \circ e$ este o mulțime regulară de alegere și mărginită.

D e f i n i ț i a 3.2. *Se numește limbaj metaliniar mărginit un limbaj care poate fi obținut ca reuniune finită de produse finite de limbaje regulare de alegere mărginite.*

Dacă B_2 este clasa limbajelor metaliniare mărginite, atunci $B_2 \subsetneq S_2$.

TEOREMA 3.3. *Orice limbaj regular de alegere mărginit este limbaj metaliniar mărginit. Există însă limbaje metaliniare mărginite care nu sînt regulare de alegere mărginite; $B_1 \subsetneq B_2$.*

Demonstrație. 1. Dacă $A \in B_1$, atunci A este metaliniar și mărginit obținut prin aplicarea de 0 ori a operațiilor indicate.

2. Fie $A = \{a^n b^n a^m b^m \mid n \geqslant 0, \; m \geqslant 0\}$ limbaj mărginit. După cum se știe A nu este limbaj regular de alegere. Se vede însă că

$$A = [(a, b)^* \circ e][(a, b)^* \circ e]$$

se obține prin concatenarea a două limbaje regulare de alegere, mărginite, este deci metaliniar și mărginit fără a fi regular de alegere.

TEOREMĂ 3.4. *Orice limbaj mărginit este un limbaj metaliniar.*

Demonstrație. Pe baza definiției 1.5. familia S_2 a limbajelor metaliniare este închisă în raport cu operațiile de reuniune și produs. Atunci și $B_2 \subset S_2$ are această proprietate.

Pe de altă parte, dacă $Z \in B_2$, atunci pentru x, y cuvinte, mulțimea $(x, y)^* \circ Z$ este tot metaliniară. Într-adevăr, $(x, y)^* \circ Z = \bigcup_{n \geqslant 0} x^n Z y^n$.

Dacă notăm cu

$$A = \{x^n \mid n \geqslant 0\}, \quad B = \{y^n \mid n \geqslant 0\},$$

cele două mulțimi sînt evident mulțimi regulare, atunci AZB este o mulțime metaliniară. Dar $(x, y)^* \circ Z \subset AZB$, deci este tot o mulțime metaliniară.

Teorema 2.1. ne asigură că familia limbajelor mărginite este cea mai mică familie închisă față de operațiile de reuniune, produs și $(x, y)^* \circ Z$ aplicate de un număr finit de ori. Rezultă că familia limbajelor mărginite este inclusă în familia limbajelor metaliniare.

INCLUSION RELATIONS AMONG BOUNDED LANGUAGES

SUMMARY

The paper contains a study of the inclusion relations among some families of context free bounded languages. The definition of bounded languages is that given by S. GINSBURG and E. H. SPANIER [2]. Using the operations with sets of pairs of words, as in M. K. YNTEMA [1], the families of regular bounded matching choice sets B_1 and the metaliniar bounded languages B_2 are defined. If B_0 is the family of regular bounded sets it is shown that the following inclusions hold:

$$B_0 \subsetneqq B_1 \subsetneqq B_2.$$

BIBLIOGAFIE

[1] Yntema, M. K., *Inclusion relations among families of context free languages*, Inf. and Control, **10**, 6, 572–597 (1967).

[2] Ginsburg, S. Spanier, E. H. *Bounded ALGOL-like languages*, Trans. Amer. Math. Soc., **113**, 2, 333–367, (1964)

[3] Ginsburg, S. *The mathematical theory of context free languages*, Mc.Grow-Hill, New-York, 1966.

Institutul de calcul din Cluj al Academiei Republicii Socialiste România

Primit la 1. XII 1972.