

ALGEBRA LIMBAJELOR FORMALIZATE

DE

TEODOR RUS

(Cluj)

În lucrare se studiază *CF*-gramaticile din punctul de vedere al posibilității simplificării regulilor lor de producere. Pornindu-se apoi de la noțiunea de *CF*-gramatică normală, care este definită ca și o *CF*-gramatică a cărei reguli de producere au o formă standard, se cercetează posibilitatea organizării limbajului formalizat $L(G)$, generat de o astfel de gramatică ca și o structură algebrică. Se demonstrează astfel că gramatica G induce în $L(G)$ un domeniu de operații algebrice Ω' , astfel că

$$\mathcal{A}_G = \{L(G) \cup \{\varepsilon\}, \Omega'\}$$

este o algebră universală liberă.

Fie V o mulțime nevidă de simboluri, care admite o descompunere de forma $V = V_N \cup V_T$, $V_N \cap V_T = \emptyset$, $V_N \neq \emptyset$, $V_T = \emptyset$.

Elementele lui V_N se vor numi în continuare categorii sintactice și vor fi notate cu litere mari latine, însoțite de indici dacă este nevoie. Elementele mulțimii V_T se vor numi litere ale alfabetului limbajului formalizat și vor fi notate cu litere mici latine, însoțite de indici dacă este nevoie.

Notăm cu $W(V)$ semigrupul liber cu unitate generat de mulțimea V . Prin urmare $W(V)$ este mulțimea tuturor cuvintelor peste alfabetul V , înzestrată cu o operație — juxtapunerea — care de altfel este operația de semigrup. Elementul unitate va fi notat cu ε și este cuvântul vid. Elementele mulțimii $W(V)$ le vom numi cuvinte și le vom nota cu litere mici latine de la sfârșitul alfabetului, însoțite de indici dacă este nevoie. Dacă notăm cu $\circ$ operația de juxtapunere (sau uneori concatenare), atunci avem :

$$(a) \quad V \subset W$$

$$(b) \quad x, y \in W \Rightarrow x \circ y = xy \in W$$

$$(c) \quad x \in W \Rightarrow \varepsilon \circ x = x \circ \varepsilon = x \in W.$$

Fie acum $\mathcal{P}$ o mulțime finită de perechi de forma

$$\mathcal{P} = \{(A_1, w_1), (A_2, w_2), \dots, (A_n, w_n)\} \quad A_i \in V_N, w_i \in W, w_i \neq \varepsilon.$$

Mulțimea $\mathcal{P}$ va fi numită în continuare mulțimea producțiilor sau a regulilor de producere a limbajului formalizat. Dacă perechea $(A_i, w_i) \in \mathcal{P}$ uneori o vom nota cu $A_i \rightarrow w_i$ și o vom numi regulă de producere.

După cum am observat în [1], fiecare regulă de producere induce în W o relație binară [2], [4] pe care o notăm cu R_i și care se definește în felul următor:

$$x, y \in W, x R_i y \Leftrightarrow (x = x_1 A_i x_2 \& y = x_1 w_i x_2 \& (A_i, w_i) \in \mathcal{P})$$

cuvintele x_1, x_2 putînd fi identice cu cuvîntul vid ε . Dacă notăm cu Δ_{R_i} domeniul relației R_i și cu Γ_{R_i} rangul ei, atunci

$$\Delta_{R_i} = \{x \in W \mid x = x_1 A_i x_2\}$$

$$\Gamma_{R_i} = \{y \in W \mid y = x_1 w_i x_2\}.$$

În cele ce urmează $x R_i y$ va fi notat cu $x \vdash_{R_i} y$.

DEFINIȚIA 1. Un cvadruplu ordonat de forma

$$(W, V_N, V_T, \mathcal{P})$$

va fi numit gramatică în semigrupul liber W și va fi notat cu G .

DEFINIȚIA 2. Dacă $x, y \in W$, atunci spunem că y este o derivare a lui x cu ajutorul gramaticii G dacă $x = y$ sau dacă există șirul de cuvinte $x_0, x_1, \dots, x_m$ astfel ca:

$$x_0 = x$$

$$x_i \vdash_{R_{j_i}} x_{i+1}, \quad i = 0, 1, \dots, m-1$$

$$x_m = y.$$

După cum s-a observat în [1], dacă y este o derivare a lui x , atunci există relațiile $R_{j_1}, R_{j_2}, \dots, R_{j_m}$ astfel ca

$$x \prod_{i=1}^m R_{j_i} y.$$

Dacă notăm cu $R_j = \prod_{i=1}^m R_{j_i}$ și cu $x \vdash y$, faptul că y este o derivare a lui x , atunci este justificată notația

$$x \vdash_{R_j} y, \quad R_j = \prod_{i=1}^m R_{j_i}.$$

DEFINIȚIA 3. Numim limbaj formalizat în semigrupul liber W generat de gramatica G cea mai mare submulțime $L \subset W$ cu următoarele proprietăți:

(a) Dacă $x \in L$, atunci $x \in W(V_T)$.

(b) Dacă $x \in L$, atunci există $A_i \in V_N$ astfel ca $A_i \vdash_R x$, unde R este un produs finit de relație binare induse de $\mathcal{P}$ în W .

Notăm limbajul formalizat generat de gramatica G cu $L(G)$.

DEFINIȚIA 4. Două gramatici G și G' în semigrupul liber W se numesc echivalente dacă $L(G) = L(G')$.

În lucrarea [1] am demonstrat că dacă G este o gramatică în W , atunci există un algoritm finit $\mathfrak{A}$ care partiționează mulțimea $\mathcal{P}$ în submulțimile $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_N$, astfel ca

$$(a) \quad \mathcal{P} = \bigcup_{i=0}^N \mathcal{Q}_i.$$

(b) Dacă $(A_i, w_i) \in \mathcal{Q}_0$, atunci $w_i \in W(V_T)$.

(c) $\mathcal{Q}_0 \neq \Phi$ și dacă notăm cu $\mathcal{A}_0 = \{A_{01}, A_{02}, \dots, A_{0n}\}$ mulțimea tuturor categoriilor sintactice care apar în părțile stîngi ale perechilor din $\mathcal{Q}_0$ și dacă $\mathcal{A}_1 = \{A_{10}, A_{11}, \dots, A_{1k}\}$, $j = 1, 2, \dots, N$ este mulțimea tuturor categoriilor sintactice care apar în părțile stîngi ale perechilor mulțimilor $\mathcal{Q}_1, \mathcal{Q}_2, \dots, \mathcal{Q}_N$ atunci

$$\mathcal{A}_0 \subseteq \mathcal{A}_1$$

(d) $\mathcal{Q}_i \cap \mathcal{Q}_j = \emptyset$, $i, j = 0, 1, 2, \dots, N$, $i \neq j$.

(e) Dacă $(A_j, w_j) \in \mathcal{Q}_i$, $i > 0$, atunci cuvîntul w_j conține numai intrările acelor categorii sintactice care apar în părțile stîngi ale perechilor grupelor $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_i$, și pentru fiecare $(A_j, w_j) \in \mathcal{Q}_i$, w_j conține intrări ale cel puțin uneia din categoriile sintactice care este parte stîngă a cel puțin uneia din perechile grupei $\mathcal{Q}_{i-1}$.

Fie acum $\mathcal{Q}_N = \{(A_{N1}, w_{N1}), (A_{N2}, w_{N2}), \dots, (A_{Nk}, w_{Nk})\}$.

DEFINIȚIA 5. Limbajul formalizat $L(G)$ generat de gramatica G cu proprietățile înșirate mai sus se numește CF-limbaj (Context-Free).

O gramatică G care definește un CF-limbaj se numește CF-gramatică.

Fie acum G o CF-gramatică și fie $(A_{Ni}, w_{Ni}) \in \mathcal{Q}_N$. Considerăm toate lanțurile de perechi de forma

$$(A_{Ni}, A_1) \in \mathcal{Q}_N, A_{Ni}, A_1 \in V_N$$

$$(A_1, A_2) \in \mathcal{Q}_{j_1}, A_1, A_2 \in V_N$$

$$\dots$$

$$(A_{p-1}, A_p) \in \mathcal{Q}_{j_{p-1}}, A_p, A_{p-1} \in V_N$$

$$(A_p, w_p) \in \mathcal{Q}_{j_p}, A_p \in V_N, w_p \in W$$

$$0 \leq j_p \leq j_{p-1} \leq \dots \leq j_1 \leq N,$$

unde $w_p = t_1 x_1 t_2 x_2 \dots t_n x_n t_{n+1}$, $t_i \in W(V_T)$, $x_i \in W(V_N)$.

Dacă nu există i astfel ca $t_i \neq \varepsilon$ în reprezentarea lui w_p , atunci se presupune că $l(w_p) > 1$, unde prin $l(x)$ am notat lungimea lui x .

Un lanț de forma de mai sus se va numi în continuare lanț cu proprietatea (s).

Fie acum $\mathbb{M}_{A_{N_i}}$ mulțimea tuturor lanțurilor de perechi cu proprietatea (s) pentru o anumită pereche $(A_{N_i}, w_{N_i}) \in \mathcal{Q}_N$.

DEFINIȚIA 6. O CF-gramatică G se va numi simplificată dacă pentru orice $(A_{N_i}, w_{N_i}) \in \mathcal{Q}_N$, $\mathbb{M}_{A_{N_i}} \neq \emptyset$ și dacă orice lanț obținut pornind de la (A_{N_i}, w_{N_i}) poate fi completat la un lanț cu proprietatea (s).

Are loc următoarea teoremă :

TEOREMA 1. Dacă G este o CF-gramatică și dacă numărul grupelor sale $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_N$ este astfel ca $N > 0$ și cel puțin o pereche din $\mathcal{Q}_i$, $i > 0$ este de forma (A_i, w_i) , $l(w_i) > 1$, $w_i \in W(V_T)$, atunci există o CF-gramatică G' simplificată, echivalentă cu prima, astfel ca dacă $\mathcal{Q}_0, \mathcal{Q}'_1, \dots, \mathcal{Q}'_{N'}$ sînt grupele sale, atunci pentru fiecare $(A_{N'_i}, w_{N'_i}) \in \mathcal{Q}'_{N'}$, $\mathbb{M}_{A_{N'_i}} \neq \emptyset$ și $N' > 0$.

Demonstrație. Fie G o CF-gramatică. Să considerăm $(A_{N_i}, w_{N_i}) \in \mathcal{Q}_N$. Dacă $\mathbb{M}_{A_{N_i}} \neq \emptyset$, atunci avem următoarele două cazuri :

- (1) $\mathbb{M}_{A_{N_i}}$ conține numai lanțuri cu proprietatea (s).
- (2) $\mathbb{M}_{A_{N_i}}$ conține și lanțuri de altă formă.

În primul caz dacă $\mathbb{M}_{A_{N_i}}$ are această proprietate pentru fiecare $(A_{N_i}, w_{N_i}) \in \mathcal{Q}_N$, atunci G este o gramatică simplificată. În cazul al doilea fiecare lanț care nu are proprietatea (s) și nu poate fi prelungit la un lanț cu această proprietate are una din formele :

$$(A_{N_i}, A_1)(A_1, A_2), \dots, (A_p, w_p); (A_{N_i}, A'_1) \in \mathcal{Q}_N, (A_p, w_p) \in \mathcal{P};$$

$$(A_i, A_{i+1}) \in \mathcal{P}; i = 1, 2, \dots, p-1,$$

unde

- (a) $w_p \in W(V_T)$ sau
- (b) $w_p \in W(V'_N)$, $l(w_p) = 1$.

Să considerăm acum toate lanțurile de forma

$$(l_1) (A_{N_i}, A_1), (A_1, A_2), \dots, (A_i, A_{i+1}), \dots, (A_p, w_p)$$

$$(l_2) (A_{N_i}, A_1), (A_1, A_2), \dots, (A_i, B_{i+1}), \dots, (B_k, w_k)$$

în care (l_1) nu posedă proprietatea (s) pe cînd (l_2) posedă această proprietate

Dacă $w_p \in W(V_T)$, atunci din punct de vedere generativ perechile $(A_i, A_{i+1}), \dots, (A_{p-1}, A_p)$ pot fi eliminate din $\mathcal{P}$ deoarece

$$A_{N_i} \vdash_R x, \quad x = x_1 w_p x_2$$

este echivalentă cu

$$A_p \vdash_{R_p} x, \quad x = x_1 w_p x_2.$$

Pentru realizarea acestei eliminări se propune următorul algoritm :

Se consideră un lanț cu proprietatea (s) de tipul (l_2) . Se formează apoi un lanț de tipul (l_1) care nu posedă proprietatea (s) și care coincide cu (l_2) pe o anumită porțiune și diferă de acesta pe o altă porțiune, terminîndu-se cu perechea (A_p, w_p) , unde $w_p \in W(V_T)$. Se elimină apoi din $\mathcal{P}$ toate perechile care nu coincid cu perechile lanțului (l_2) , mai puțin ultima pereche (A_p, w_p) . Se procedează apoi în mod analog cu toate lanțurile de tipul (l_1) care se pot forma în acest fel pornindu-se de la perechea $(A_{N_i}, w_{N_i}) \in \mathcal{Q}_N$.

Dacă are loc cazul (b), adică $w_p \in V_N$, atunci în mod absolut analog se arată că se pot elimina din $\mathcal{P}$ de data aceasta toate perechile lanțurilor care nu coincid cu perechile unui lanț de tipul (l_2) cu proprietatea (s). Aceasta deoarece nu pot exista cuvinte $z \in L(G)$ astfel ca să avem

$$A_{N_i} \vdash_{R_p} y \vdash_{R_p} z.$$

Dacă ar exista astfel de cuvinte, atunci am avea

$$y = y_1 A_p y_2$$

$$z = y_1 w_p y_2$$

și deoarece $w_p \in V_N$, ar urma că $z \in L(G)$, ceea ce reprezintă o contradicție. De aici rezultă că (A_p, w_p) se poate elimina.

După examinarea tuturor lanțurilor mulțimii $\mathbb{M}_{A_{N_i}}$ avem sau $\mathbb{M}_{A_{N_i}} \neq \emptyset$ sau $\mathbb{M}_{A_{N_i}} = \emptyset$. Dacă există cel puțin o pereche (A_{N_i}, w_{N_i}) în $\mathcal{Q}_N$ astfel ca în urma eliminărilor făcute mai sus să avem $\mathbb{M}_{A_{N_i}} \neq \emptyset$, atunci gramatica obținută este gramatica G' .

Dacă asemenea perechi nu există, înseamnă că prin eliminările făcute am eliminat din $\mathcal{P}$ toate perechile mulțimii $\mathcal{Q}_N$. Gramatica astfel obținută o notăm cu G_1 . Cu această gramatică se procedează în mod absolut analog ca și cu gramatica G . În felul acesta după un număr finit de pași se ajunge la o gramatică G_i care posedă fie numai producții din $\mathcal{Q}_0$, fie la o gramatică G_i simplificată și echivalentă cu G . Să presupunem acum că are loc cazul în care gramatica G_i are numai producții din $\mathcal{Q}_0$. Atunci este contrazisă ipoteza teoremei care afirmă că pentru cel puțin o pereche (A_i, w_i) din $\mathcal{Q}_i$, $i > 0$, (A_i, w_i) este astfel ca $l(w_i) > 1$, și $w_i \in W(V_T)$. Această contradicție demonstrează complet teorema.

TEOREMA 2. Dacă G este o CF-gramatică, atunci ea este echivalentă cu o CF-gramatică G' ale cărei reguli de producere au următoarele proprietăți :

Dacă $(A_i, w_i) \in \mathcal{P}'$, $A_i \in V'_N$, $w_i \in W(V')$, atunci w_i este de una din formele

$$(a) w_i \in W(V_T)$$

$$(b) w_i = AB, \quad A, B \in V_N,$$

$$(c) w_i = B, \quad B \in V_N,$$

$$(d) w_i = t_1 B_1 t_2 B_2 \dots t_n B_n t_{n+1}; B_i \in V_N, i = 1, 2, \dots, n$$

$$t_i \in W(V_T), i = 1, 2, \dots, n+1, \text{ și } t_i \neq \varepsilon \text{ pentru } i \neq 1, i \neq n+1.$$

Demonstrație. Vom construi în continuare gramatica G' care satisface condițiile teoremei și este echivalentă cu G . Fie pentru aceasta gramatica G dată și fie descompunerea mulțimii $\mathcal{P}_G$ în grupele $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_N$. Notăm cu $\mathcal{Q}'_0, \mathcal{Q}'_1, \dots, \mathcal{Q}'_N$ grupele gramaticii G' . Aceste grupe se construiesc în felul următor:

$$(1) \mathcal{Q}_0 \subset \mathcal{Q}'_0.$$

(2) Fie $\mathcal{Q}_1 = \{(A_{11}, w_{11}), (A_{12}, w_{12}), \dots, (A_{pp}, w_{pp})\}$; $\mathcal{Q}'_1$ se va construi atunci astfel: Considerăm prima producție $(A_{11}, w_{11}) \in \mathcal{Q}_1$. Dacă w_{11} are una din formele (a)–(d), atunci $(A_{11}, w_{11}) \in \mathcal{G}'_1$. Să presupunem acum că w_{11} nu este de una din formele (a)–(d). Atunci

$$w_{11} = A_1 A_2 \dots A_k, \quad A_j \in V_N, \quad j = 1, 2, \dots, k, \quad k > 1$$

sau

$$w_{11} = t_1 x_1 t_2 x_2 \dots t_n x_n t_{n+1}, \quad t_i \in W(V_T), \quad t_i \neq \varepsilon, \quad i \neq 1, i \neq n+1$$

unde pentru cel puțin un i , $l(x_i) > 1$ și $x_i \in W(V_N)$. Construim atunci un șir de perechi de producere care din punct de vedere generativ este echivalent cu perechea (A_{11}, w_{11}) și înlocuim această pereche cu șirul astfel construit. Pentru aceasta considerăm simbolurile $B^1, B^2, \dots, B^{k-1}$ care nu aparțin mulțimii V . Fie atunci

$$V' = V \cup \{B^1, B^2, \dots, B^{k-1}\}.$$

și luăm

$$V'_N = V_N \cup \{B^1, B^2, \dots, B^{k-1}\}.$$

Considerăm acum șirul de perechi

$$(B^1, A_1 A_2), (B^2, B^1 A_3), \dots, (B^{k-1}, B^{k-2} A_k) (A_{11}, B^{k-1}).$$

Notăm cu $R^1, R^2, \dots, R^k$ relațiile binare induse de aceste perechi în $W(V')$ și cu R_{11} relația binară indusă de (A_{11}, w_{11}) în $W(V)$. Atunci are loc următoarea egalitate:

$$\prod_{i=k}^1 R^i = R_{11}.$$

Într-adevăr, fie $(x, y) \in R_{11}$. Atunci

$$x = x_1 A_{11} x_2, \quad y = x_1 w_{11} x_2.$$

Construim acum șirul de cuvinte $x_0, x_1, \dots, x_k$ astfel ca

$$x_0 = x$$

$$x_i \vdash_{R^i} x_{i+1}, \quad i = 0, 1, \dots, i-1$$

$$x_k = y.$$

Pentru aceasta punem

$$x_0 = x = x_1 A_{11} x_2$$

$$x_1 = x_1 B^{k-1} x_2$$

$$x_2 = x_1 B^{k-1} A_k x_2$$

$$\dots$$

$$x_k = x_1 A_1 A_2 \dots A_k x_2$$

și prin urmare $(x, y) \in \prod_{i=k}^1 R^i = R_k \circ R_{k-1} \circ \dots \circ R_1$. În mod analog se arată

și invers, dacă $(x, y) \in \prod_{i=k}^1 R^i$, atunci $(x, y) \in R_{11}$. Acest lucru justifică înlocuirea lui (A_{11}, w_{11}) cu șirul de perechi construit mai sus. În acest caz punem:

$$(B^1, A_{11}) \in \mathcal{Q}'_1$$

$$(B^i, B^{i-1} A_{i+1}) \in \mathcal{Q}'_i, \quad i = 2, 3, \dots, k-1$$

$$(A_n, B^{k-1}) \in \mathcal{Q}'_k.$$

Considerăm acum cazul $(A_{11}, w_{11}) \in \mathcal{Q}_1$, unde w_{11} are forma $w_{11} = t_1 x_1 t_2 x_2 \dots t_n x_n t_{n+1}$, $t_i \in W(V_T)$, $t_i \neq \varepsilon$, $i \neq 1, i \neq n+1$ și pentru cel puțin un j , $l(x_j) > 1$, $x_j \in W(V_N)$.

Notăm cu $x_{j_1}, x_{j_2}, \dots, x_{j_p}$ toate cuvintele care intră în w_{11} , $x_{j_i} \in W(V_N)$, $i = 1, 2, \dots, p$ și $l(x_{j_i}) > 1$. Să presupunem că acestea sînt de forma

$$x_{j_i} = A_i^1 A_i^2 \dots A_i^{s_i}, \quad s_i > 1, \quad A_i^l \in V_N, \quad l = 1, 2, \dots, s_i.$$

Fie atunci mulțimea

$$V'' = V' \cup \{B_i^1, B_i^2, \dots, B_i^{s_i-1}\}, \quad i = 1, 2, \dots, p,$$

$$B_i^l \in V', \quad l = 1, 2, \dots, s_i - 1$$

și punem

$$V''_N = V'_N \cup \{B_i^1, B_i^2, \dots, B_i^{s_i-1}\}.$$

În continuare alegem perechile

$$(B_i^1, A_i^1 A_i^2), (B_i^2, B_i^1 A_i^3) \dots (B_i^{s_i-1}, B_i^{s_i-2} A_i^{s_i})$$

și notăm cu $R_i^1, R_i^2, \dots, R_i^{s_i-1}$ relațiile binare induse de aceste perechi în semigrupul liber $W(V'')$. Considerăm apoi perechea (A_{11}, w'_{11}) , unde w'_{11} se obține din w_{11} prin înlocuirea lui x_{j_i} cu $B_i^{s_i-1}$, $i = 1, 2, \dots, p$. Atunci se vede imediat că avem $R_i^{s_i-1} \circ \dots \circ R_1^1 \circ R_2^{s_2-1} \circ \dots \circ R_p^{s_p-1} \circ \dots \circ R_p \circ R_{11} = R'_{11}$, unde R_{11} este relația binară indusă de perechea (A_{11}, w_{11}) în $W(V)$.

În felul acesta perechea (A_{11}, w_{11}) se poate înlocui cu perechile

$$\begin{aligned} & (B_1^1, A_1^1 A_1^2), (B_1^2, B_1^1 A_1^3), \dots, (B_1^{s_1-1}, B_1^{s_1-2} A_1^{s_1}) \\ & (B_2^1, A_2^1 A_2^2), (B_2^2, B_2^1 A_2^3), \dots, (B_2^{s_2-1}, B_2^{s_2-2} A_2^{s_2}) \\ & \dots \dots \dots \\ & (B_p^1, A_p^1 A_p^2), (B_p^2, B_p^1 A_p^3), \dots, (B_p^{s_p-1}, B_p^{s_p-2} A_p^{s_p}) \\ & (A_{11}, w_{11}). \end{aligned}$$

În acest caz punem

$$\begin{aligned} & (B_i^1, A_i^1 A_i^2) \in Q'_1, i = 1, 2, \dots, p \\ & (B_i^j, B_i^{j-1} A_i^{j+1}) \in Q'_j, j = 2, 3, \dots, s_i - 1, i = 1, 2, \dots, p. \\ & (A_{11}, w_{11}) \in Q'_l, l = \max_{i=1}^p s_i. \end{aligned}$$

Aplicînd acest procedeu pentru toate regulile de producere din Q_1 , vom construi regulile de producere ale grupei Q'_1 . Procesul se repetă apoi în mod absolut analog pentru grupele $Q_2, Q_3, \dots, Q_N$. Gramatica obținută în urma acestui proces este echivalentă cu G și satisface condițiile cerute în teoremă.

Se poate observa acum că procesul descris mai sus are drept scop o anumită simplificare a regulilor de producere a unei CF-gramatici. Această simplificare aduce după sine o mărire uneori substanțială a numărului acestor reguli.

Vom arăta în continuare că regulile de producere ale unei CF-gramatici pot fi simplificate și mai mult în sensul că fiecare regulă de producere de forma (A_i, w_i) cu $w_i = t_1 B_1 t_2 B_2 \dots t_n B_n t_{n+1}$, $t_i \in W(V_T)$, $t_i \neq \varepsilon$, $i \neq 1$ și $i \neq n+1$, $B_i \in V'_N$, se poate înlocui cu un șir de reguli de forma (A_i, w'_i) , unde w'_i are forma

$$\begin{aligned} & t_1 B_2 t_2, B \in V'_N, t_1, t_2 \neq \varepsilon, t_1, t_2 \in W(V_T) \text{ sau} \\ & A t B, A, B \in V'_N, t \neq \varepsilon, t \in W(V_T). \end{aligned}$$

Pentru aceasta facem următoarea analiză a regulii (A_i, w_i) , unde

$$\begin{aligned} & w_i = t_1 B_1 t_2 B_2 \dots t_n B_n t_{n+1}, t_i \in W(V_T), t_i \neq \varepsilon \text{ pentru } i \neq 1, i \neq n+1 \\ & B_i \in V'_N. \end{aligned}$$

a) Dacă $t_1 = \varepsilon$, atunci considerăm regula de producere

$$(B_i, B_1 t_2 B_2), B_i \in V'_N,$$

unde V'_N este mulțimea categoriilor sintactice ale gramaticii G' . Considerăm apoi regula de producere $(A_i, B_i w'_i)$, unde w'_i se obține din w_i prin ștergerea cuvîntului $B_1 t_2 B_2$.

Dacă $t \neq \varepsilon$, atunci $w_i = t_1 B_1 t_2 w_i$ și noua regulă de producere care este considerată are forma:

$$(B_i, t_1 B_1 t_2), t_1, t_2 \neq \varepsilon, t_1, t_2 \in W(V_T), B_i \in V'_N B_i \in V'_N.$$

După aceea se consideră regulile de producere de forma (A_i, B_i, w_i) cu care se procedează în mod analog ca și mai sus. Aplicînd acest procedeu, eventual alternat cu procedeu descris în demonstrația teoremei 2, se vede ușor că după un număr finit de pași obținem o gramatică G'' echivalentă cu gramatica G ale cărei reguli de producere (A_i, w_i) au proprietatea că w_i are una din formele:

- (a) $w_i \in W(V_T)$
- (b) $w_i = B, B \in V'_N$
- (c) $w_i = AB, A, B \in V'_N$
- (d) $w_i = t_1 B t_2, B \in V'_N, t_1, t_2 \in W(V_T), t_1, t_2 \neq \varepsilon$
- (e) $w_i = A t B, A, B \in V'_N, t \in W(V_T), t \neq \varepsilon$.

DEFINIȚIA 7. O CF-gramatică simplificată ale cărei reguli de producere sînt (A_i, w_i) , $i = 1, 2, \dots, n$, unde w_i are una din formele (a) — (e), va fi numită CF-gramatică normală.

TEOREMA 3. Dacă G este o CF-gramatică în semigrupul liber $W(V)$, $V = V_N \cup V_T$, $V_N \cap V = \emptyset$, $V_N = \emptyset$, $V_T \neq \emptyset$, atunci există o extindere V' a lui V și o CF-gramatică normală G' în semigrupul liber $W(V')$, $V' = V'_N \cup V_T$ echivalentă cu G .

Această teoremă este o consecință a teoremelor 1 și 2 și a procedurii de simplificare a regulilor de producere expus mai sus.

Vom demonstra în continuare cîteva leme care ne vor servi la organizarea unui CF-limbaj ca și o algebră universală liberă indusă de G în $W(V)$.

LEMA 1. Dacă G este o CF-gramatică normală în semigrupul liber W , atunci mulțimea $\mathcal{P}_G$ a regulilor ei de producere induce în W un domeniu Ω de operații algebrice.

Demonstrație. Fie $\mathcal{P}_G$ mulțimea regulilor de producere ale gramaticii G , împărțită în grupele $Q_0, Q_1, \dots, Q_N$. Vom construi pentru fiecare regulă $(A_i, w_i) \in \mathcal{P}_G$ o operație algebrică în W . Pentru aceasta vom proceda pe rînd cu grupele $Q_0, Q_1, \dots, Q_N$. Înainte de aceasta considerăm însă grupa Q_0 astfel că

$$Q_0 = \{(A_{01}, w_{01}), (A_{02}, w_{02}), \dots, (A_{0p}, w_{0p})\},$$

unde $A_{0i} \in V_N$, $w_{0i} \in W(V_T)$, $i = 1, 2, \dots, p$. Vom numi în continuare cuvintele w_{0i} , $i = 1, 2, \dots, p$, constante individuale și atunci vom avea

$$\mathcal{Q} = \{w_{0i} \mid w_{0i} \in W(V_T) \text{ \& } \exists (A_{0i}, w_{0i}) \in Q_0\}.$$

Elementele lui $\mathcal{Q}$ vor fi notate cu c_i .

Operațiile induse de perechile grupei Q_0 se definesc astfel: Dacă $(A_{0i}, w_{0i}) \in Q_0$, atunci θ_i^0 se definește prin schema:

$$\theta_i^0(x) = \begin{cases} x_1 A_{0i} x_2 & \text{dacă } x \in W \text{ și } x = x w_{0i} x_2, x_1 \neq w_{0i} x_2. \\ \varepsilon & \text{în caz contrar.} \end{cases}$$

Presupunem acum că $(A_i, w_i) \in \mathcal{Q}_j$, $j = 1, 2, \dots, N$. Deoarece G este o CF-gramatică normală, w_i are una din formele:

- (b) $w_i = B$, $B \in V_N$
- (c) $w_i = AB$, $A, B \in V_N$
- (d) $w_i = t_1 B t_2$, $B \in V_N$, $t_1, t_2 \in W(V_T)$, $t_1, t_2 \neq \varepsilon$
- (e) $w_i = AtB$, $A, B \in V_N$, $t \in W(V_T)$, $t \neq \varepsilon$.

Operația θ_i^j indusă de perechea (A_i, w_i) în cazul (b) se definește astfel

$$\theta_i^j(x) = \begin{cases} x_1 A_i x_2 & \text{dacă } x \in W \text{ și } x = x_1 B x_2, x_1 \neq v B z. \\ \varepsilon & \text{în caz contrar.} \end{cases}$$

În cazul (c) vom avea

$$\theta_i^j(x) = \begin{cases} x, A_i x_2 & \text{dacă } x \in W \text{ și } x = x_1 A B x_2, x_1 \neq v A B z \\ \varepsilon & \text{în caz contrar,} \end{cases}$$

În cazul (d) vom avea

$$\theta_i^j(x) = \begin{cases} x_1 A_i x_2 & \text{dacă } x \in W \text{ și } x = x_1 t_1 B t_2 x_2, x_1 \neq v t_1 B t_2 z \\ \varepsilon & \text{în caz contrar.} \end{cases}$$

În cazul (e) vom avea

$$\theta_i^j(x) = \begin{cases} x_1 A_i x_2 & \text{dacă } x \in W \text{ și } x = x_1 A t B x_2, x_1 \neq v A t B z, \\ \varepsilon & \text{în caz contrar.} \end{cases}$$

În toate aceste cazuri x_1, x_2, V, z pot fi egale cu cuvîntul vid ε .

Dacă notăm cu $\Theta_0, \Theta_1, \dots, \Theta_N$ operațiile algebrice astfel definite de către toate regulile de producere din $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_N$ respective, atunci

$$\Omega = \mathcal{C} \cup \Theta_0 \cup \Theta_1 \cup \dots \cup \Theta_N.$$

Fie acum limbajul formalizat $L(G)$ generat de către CF-gramatica normală G , adică

$$L = \{x \in W(V_T) \mid (\exists A_i \in V_N) A_i \models_R x\},$$

unde R este un produs finit de relații binare induse de mulțimea $\mathcal{Q}_G$ în W . Notăm cu $[A_i]$ mulțimea tuturor cuvintelor limbajului $L(G)$ astfel ca

$$[A_i] = \{x \mid A_i \models x, x \in L(G)\}.$$

Atunci este evidentă egalitatea

$$L(G) = \bigcup_{i \in I} [A_i],$$

unde $I = (0, 1, 2, \dots, m)$, $m = \text{Card}(V_N)$.

LEMA 2. Dacă $A_i \in V_N$ și $x \in [A_i]$, atunci există un șir finit de operații $\theta_1, \theta_2, \dots, \theta_p$ în Ω așa ca

$$\theta_1 \theta_2, \dots, \theta_p(x) = A_i,$$

unde prin $\theta_1 \theta_2 \dots, \theta_p(x)$ am notat aplicarea succesivă a operațiilor $\theta_1, \theta_2, \dots, \theta_p$ asupra cuvîntului x .

Demonstrație. Deoarece $x \in [A_i]$, urmează că există $R = \prod_{i=0}^{p-1} R_{j_i}$,

unde R_{j_i} este relația binară indusă de perechea (A_i, w_i) în W , pentru $i = 0, 1, \dots, p-1$, astfel ca $A_i \models_R x$. Prin urmare există șirul de cuvinte $x_0, x_1, \dots, x_p$ astfel ca să avem

$$\begin{aligned} x_0 &= A_i \\ x_i &\vdash_{R_{j_i}} x_{i+1}, \quad i = 0, 1, \dots, p-1 \\ x_p &= x. \end{aligned}$$

Aceasta înseamnă că

$$\begin{aligned} x_{p-1} &= x^1 A_{j_{p-1}} x^2 \\ x_n &= x^1 w_{j_{p-1}} x^2 \end{aligned}$$

unde $x_n = x \in L(G)$, $(A_{j_{p-1}}, w_{j_{p-1}}) \in \mathcal{P}_G$. Prin urmare vom avea $(A_{j_{p-1}}, w_{j_{p-1}}) \in \mathcal{Q}_0$. Fie atunci θ_p operația indusă de perechea $(A_{j_{p-1}}, w_{j_{p-1}})$ în W . Atunci

$$\theta_p(x) = \theta_p(x_p) = x^1 A_{j_{p-1}} x^2 = x_{p-1}.$$

Procedînd analog cu x_{p-1} ca și cu x_p , din aproape în aproape ajungem la egalitatea

$$\theta_1(\theta_2(\dots, (\theta_p(x)))) = A_i$$

sau

$$\theta_1 \theta_2 \dots \theta_p(x) = A_i$$

ceea ce demonstrează lema.

LEMA 3. Dacă $x \in L(G)$ și dacă există șirul de operații $\theta_1, \theta_2, \dots, \theta_p$ din Ω astfel ca $\theta_1 \theta_2 \dots, \theta_p(x) = A_i$, atunci există relațiile $R_{j_1}, R_{j_2}, \dots, R_{j_p}$ induse de către perechile $(A_{j_1}, w_{j_1}), (A_{j_2}, w_{j_2}), \dots, (A_{j_p}, w_{j_p})$ astfel ca să avem

$$A_i \models_R x, \quad R = \prod_{i=1}^p R_{j_i}.$$

Demonstrație. Fie $x \in L(G)$ și $\theta_1, \theta_2, \dots, \theta_p \in \Omega$ astfel ca $\theta_1 \theta_2 \dots \theta_p(x) = A_i$. Conform definiției lui Ω avem

$$x = x_1 w_{j_p} x_2, \quad x_1 \neq v w_{j_p} z, \quad (A_{j_p}, w_{j_p}) \in \mathcal{P}_G$$

iar operația θ_p indusă de această pereche aplicată la x ne dă

$$y = \theta_p(x) = x_1 A_{j_p} x_2.$$

Procedăm analog cu y ca și cu x și apoi din aproape în aproape punem în evidență perechile $(A_{j_1}, w_{j_1}), (A_{j_2}, w_{j_2}), \dots, (A_{j_p}, w_{j_p})$ astfel

ca $A_i \models_R x$, $R = \prod_{i=1}^p R_{j_i}$, ceea ce demonstrează lema 3.

DEFINIȚIA 8. Dacă $x \in L(G)$ și $\theta_1, \theta_2, \dots, \theta_p$ este un șir de operații algebrice din Ω astfel ca $\theta_1 \theta_2, \dots, \theta_p(x) = A_i$, unde $A_i \in V_N$, atunci vom spune că $\theta_1, \theta_2, \dots, \theta_p$ este o analiză a cuvîntului x în mulțimea $[A_i]$.

LEMA 4. Fie $x \in L(G)$ și $x \in [B]$, $B \in V_N$. Dacă în mulțimea $\mathcal{P}_G$ există regula (A_i, B) , atunci $x \in [A_i]$.

Demonstrație. Deoarece $x \in L(G)$ și $x \in [B]$, conform lemei 2 există o analiză $\theta_1, \theta_2, \dots, \theta_p$ a lui x în $[B]$.

$$\theta_1, \theta_2, \dots, \theta_p(x) = B.$$

Dar în $\mathcal{P}$ există perechea (A_i, B) . Atunci în Ω există operația θ_k indusă de această pereche, astfel ca $\theta_k(B) = A_i$. Aceasta înseamnă că șirul $\theta_k, \theta_1, \dots, \theta_p$ este o analiză a cuvîntului x în $[A_i]$. După lema 3 rezultă că $A_i \mid_R x$.

LEMA 5. Dacă $x_1, x_2 \in L(G)$ și $x_1 \in [A]$, $x_2 \in [B]$, iar $A, B \in V_N$ și dacă în $\mathcal{P}$ există perechea (A_i, AB) , atunci avem $x = x_1 x_2 \in L(G)$ și $x = x_1 x_2 \in [A_i]$.

Demonstrație. Fie $\theta_1^1, \theta_2^1, \dots, \theta_{p_1}^1$ și $\theta_1^2, \theta_2^2, \dots, \theta_{p_2}^2$ analizele lui x_1 și respectiv x_2 în $[A]$ respectiv în $[B]$. După lema 3 avem

$$\theta_1^1 \theta_2^1 \dots \theta_{p_1}^1(x_1) = A$$

$$\theta_1^2 \theta_2^2 \dots \theta_{p_2}^2(x_2) = B.$$

Presupunem acum că θ_k este operația indusă de perechea (A_i, AB) . Atunci, evident, vom avea $\theta_k(\theta_1^1 \dots \theta_{p_1}^1(x_1) \circ \theta_1^2 \theta_2^2 \dots \theta_{p_2}^2(x_2)) = A_i$, unde $\circ$ este operația de semigrup. De aici urmează că șirul operațiilor $\theta_k, \theta_1^1, \theta_2^1, \dots, \theta_{p_1}^1, \theta_1^2, \theta_2^2, \dots, \theta_{p_2}^2$, este o analiză a cuvîntului $x = x_1 x_2$ în $[A_i]$ și prin urmare lema este demonstrată.

LEMA 6. Dacă $x_1 \in L(G)$ și $x_1 \in [B]$, $B \in V_N$, și dacă în $\mathcal{P}$ există perechea $(A_i, t_1 B t_2)$, atunci cuvîntul $x = t_1 x_1 t_2$ aparține limbajului $L(G)$ și $t_1 x_1 t_2 \in [A_i]$.

Demonstrație. Fie $\theta_1, \theta_2, \dots, \theta_p$ o analiză a cuvîntului x_1 în B . Considerăm atunci cuvîntul $x = t_1 \circ \theta_1 \theta_2, \dots, \theta_p(x_1) \circ t_2$ și operația θ_k indusă de $(A_i, t_1 B t_2)$ în W . Este evident că

$$\theta_k(t_1 \circ \theta_1 \theta_2 \dots \theta_p(x_1) \circ t_2) = A_i.$$

Prin urmare $\theta_k, \theta_1, \dots, \theta_p$ este o analiză a lui $t_1 x_1 t_2$ în $[A_i]$ și deci $t_1 x_1 t_2 \in L(G)$ și $t_1 x_1 t_2 \in [A_i]$ ceea ce demonstrează lema.

LEMA 7. Dacă $x_1, x_2 \in L(G)$ și $x_1 \in [A]$, $x_2 \in [B]$, $A, B \in V_N$ și dacă există perechea (A_i, AtB) , atunci $x_1 t x_2 \in [A_i]$ și $x_1 t x_2 \in L(G)$.

Demonstrație. Fie $\theta_1^1, \theta_2^1, \dots, \theta_{p_1}^1$ și $\theta_1^2, \theta_2^2, \dots, \theta_{p_2}^2$ analizele lui x_1 și x_2 în $[A]$ respectiv $[B]$. Considerăm cuvîntul $x = \theta_1^1, \theta_2^1, \dots, \theta_{p_1}^1(x_1) \circ t \circ \theta_1^2 \theta_2^2, \dots, \theta_{p_2}^2(x_2) = AtB$. Să presupunem că θ_k este operația indusă de (A_i, AtB) în W . Atunci avem $\theta_k(x) = A_i$ și prin urmare șirul operațiilor $\theta_k, \theta_1^1, \dots, \theta_{p_1}^1, \theta_1^2, \theta_2^2, \dots, \theta_{p_2}^2$ este o analiză a lui $x_1 t x_2$ în $[A_i]$, de unde rezultă lema.

Fie acum un limbaj formalizat $L(G)$ generat de o CF-gramatică G și fie Ω domeniul de operații induse de $\mathcal{P}_G$ în W . Vom defini în continuare

în limbajul $L(G)$ un domeniu de operații algebrice Ω' astfel ca să avem $\text{Card}(\Omega) = \text{Card}(\Omega')$.

Mulțimea Ω' se construiește în felul următor: dacă $c_i \in \mathcal{C}$, atunci evident $c_i \in L(G)$ și prin urmare putem considera $\mathcal{C}$ drept mulțimea constantelor individuale din Ω' . Fie acum operația $\theta_k^i \in \Theta_j$, $j = 0, 1, \dots, N$. Atunci θ_k^i este indusă în W de o pereche de una din formele:

$$(a) (A_k, w_k), A_k \in V_N, w_k = c_k$$

$$(b) (A_k, B), A_k, B \in V_N$$

$$(c) (A_k, AB), A_k, A, B \in V_N$$

$$(d) (A_k, t_1 B t_2), A_k, B \in V_N, t_1 t_2 \in W(V_T), t_1, t_2 \neq \varepsilon$$

$$(e) (A_k, AtB), A_k, A, B \in V_N, t \in W(V_T), t \neq \varepsilon.$$

Dacă θ_k^i este indusă în W de o pereche de forma (a), atunci acesteia îi va corespunde în $L(G)$ o operație $\theta_k'^i$ definită astfel

$$\theta_k'^i(x) \begin{cases} x & \text{dacă } x \in L(G) \text{ și } x \in [A_k] \\ \varepsilon & \text{dacă } x \in L(G) \text{ și } x \notin [A_k]. \end{cases}$$

În felul acesta corespunzător cazurilor (b)–(e) vom avea definită operația $\theta_k'^i$ după schemele următoare:

$$\theta_k'^i(x) = \begin{cases} x & \text{dacă } x \in L(G) \text{ și } x \in [B] \\ \varepsilon & \text{dacă } x \in L(G) \text{ și } x \notin [B] \end{cases}$$

$$\theta_k'^i(x_1, x_2) = \begin{cases} x_1 x_2 & \text{dacă } x_1, x_2 \in L(G) \text{ și } x_1 \in [A], x_2 \in [B] \\ \varepsilon & \text{dacă } x_1, x_2 \in L(G) \text{ și } x_1 \notin [A] \text{ sau } x_2 \notin [B] \end{cases}$$

$$\theta_k'^i(x) = \begin{cases} t_1 x t_2 & \text{dacă } x \in L(G) \text{ și } x \in [B] \\ \varepsilon & \text{dacă } x \in L(G) \text{ și } x \notin [B] \end{cases}$$

$$\theta_k'^i(x_1, x_2) = \begin{cases} x_1 t x_2 & \text{dacă } x_1, x_2 \in L(G) \text{ și } x_1 \in [A], x_2 \in [B] \\ \varepsilon & \text{dacă } x_1, x_2 \in L(G) \text{ și } x_1 \notin [A] \text{ sau } x_2 \notin [B]. \end{cases}$$

Procedînd analog pentru $j = 0, 1, \dots, N$, obținem domeniul de operații Ω' , cu proprietatea indicată mai sus. Construcția de mai sus este posibilă avînd în vedere lemele 2–7.

TEOREMA 4. Mulțimea $L(G) \cup \{\varepsilon\}$ este o algebră universală relativ la domeniul de operații algebrice Ω' .

Demonstrație. Teorema rezultă imediat din construcția mulțimii Ω' . Vom nota în continuare cu $\mathcal{A}_G$ algebra universală pusă în evidență prin teorema 4, și prin urmare avem

$$\mathcal{A}_G = \{L(G) \cup \{\varepsilon\}, \Omega'\}.$$

CONSECINȚA 1. Dacă $\mathcal{A}_G$ este algebra universală a limbajului formalizat $L(G)$ și dacă $A_k \in V_N$, atunci $[A_k] \cup \{\varepsilon\}$ este o subalgebră a algebrei $\mathcal{A}_G$.

Această consecință rezultă imediat din definiția mulțimii $[A_k]$ precum și din definiția algebrei $\mathcal{A}_G$.

CONSECINȚA 2. Dacă $A_1, A_2, \dots, A_p$ sînt toate categoriile sintactice din V_N care apar în partea stîngă a perechilor din mulțimea $\mathcal{Q}_1$ atunci

$$[A_1] \cup [A_2] \cup \dots \cup [A_p] \cup \{\varepsilon\}$$

este o subalgebră a lui $\mathcal{A}_G$.

Această consecință rezultă ușor din consecința 1 și din definiția algebrei $\mathcal{A}_G$. Notăm în continuare cu $\mathcal{A}_G^1$ subalgebra algebrei $\mathcal{A}_G$ pusă în evidență în consecința 2.

TEOREMA 5. Algebra $\mathcal{A}_G$ este o algebră liberă a cărei mulțime de generatori coincide cu mulțimea $\mathcal{C}$.

Demonstrație. Să considerăm toate perechile grupelor

$$\mathcal{Q}_0 = \{(A_{01}, w_{01}), (A_{02}, w_{02}), \dots, (A_{0p}, w_{0p})\}$$

$$\mathcal{Q}_1 = \{(A_{11}, w_{11}), (A_{12}, w_{12}), \dots, (A_{1q}, w_{1q})\}$$

În acest caz $A_1, A_2, \dots, A_p$ din consecința 2 sînt chiar $A_{11}, A_{12}, \dots, A_{1p}$. Cuvintele w_{1i} sînt de una din formele

$$(b) w_{1i} = B, B \in V_N$$

$$(c) w_{1i} = AB, A, B \in V_N$$

$$(d) w_{1i} = t_1 B t_2, B \in V_N, t_1, t_2 \in W(V_T), t_1, t_2 \neq \varepsilon$$

$$(e) w_{1i} = AtB, A, B \in V_N, t \in W(V_T), t \neq \varepsilon.$$

Conform definiției grupelor $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_N$, categoriile sintactice care pot interveni în descompunerea cuvîntului w_{1i} au proprietatea că

$$A, B \in \left(\bigcup_{j=0}^p A_{0j} \right) \cup \{A_{1i}\}.$$

Din proprietatea (e) enunțată la definirea partiționării mulțimii $\mathcal{P}_G$ în grupele $\mathcal{Q}_0, \mathcal{Q}_1, \dots, \mathcal{Q}_N$ urmează că nu există perechi de forma

$$(A_k, w_k) \in \mathcal{P}_G, \text{ cu } w_k \text{ de una din formele (b)–(c) astfel ca } A, B \in \left(\bigcup_{j=0}^p A_{0j} \right) \cup$$

$\{A_k\}$, unde (A_k, w_k) să aparțină uneia din grupele $\mathcal{Q}_2, \mathcal{Q}_3, \dots, \mathcal{Q}_N$. Aceasta înseamnă că $\mathcal{A}_G^1$ este cea mai mică subalgebră a algebrei $\mathcal{A}_G$ diferită de $\mathcal{C}$ care conține mulțimea $\mathcal{C}$. De aici urmează că mulțimea $\mathcal{C}$ este o mulțime de generatori pentru algebra $\mathcal{A}_G^1$.

Vom arăta acum că $\mathcal{C}$ este o mulțime de generatori liberi pentru $\mathcal{A}_G^1$. Pentru aceasta fie $\mathcal{B}$ o altă algebră similară cu $\mathcal{A}_G^1$. Vom demonstra că dacă o transformare

$$f: \mathcal{C} \rightarrow \mathcal{B}$$

poate fi extinsă la un homomorfism

$$h: \mathcal{A}_G^1 \rightarrow \mathcal{B},$$

atunci această extindere este unică. Vom construi homomorfismul h în felul următor:

$$h(x) = f(x) \text{ dacă } x \in \mathcal{C}.$$

Dacă $x \in \mathcal{C}$ deoarece $\mathcal{C}$ este o mulțime de generatori pentru algebra $\mathcal{A}_G^1$, există elementele $x_1, x_2, \dots, x_m \in \mathcal{C}$ și de asemenea operațiile $\theta_1^1, \theta_2^1, \dots, \theta_k^1$ definite în $\mathcal{A}_G^1$ care aplicate la $(x_1, x_2, \dots, x_m)$ într-o anumită ordine să ne dea elementul x . Pentru construirea lui h în continuare vom proceda prin inducție după lungimea lui x , în raport cu $x_1, x_2, \dots, x_m$. Astfel, dacă $x = \theta_1^1(x_1)$, atunci $h(x) = \theta_1^1(h(x_1))$ iar dacă $x = \theta_1^1(x_1, x_2)$, atunci $h(x) = \theta_1^1(h(x_1), h(x_2))$. În felul acesta este evident că transformarea $h: \mathcal{C} \rightarrow \mathcal{B}$ se prelungește la homomorfismul

$$h: \mathcal{A}_G^1 \rightarrow \mathcal{B}.$$

Să arătăm acum că această prelungire homomorfă a lui f este unică. Pentru aceasta să presupunem că există o altă extindere homomorfă h' a lui $f: \mathcal{C} \rightarrow \mathcal{B}$ diferită de h . Dar, conform definiției acestor homomorfisme, avem

$$h(x) = h'(x) = f(x) \text{ pentru } x \in \mathcal{C}.$$

Dacă $x \in \mathcal{C}$ și $x \in \mathcal{A}_G^1$, atunci există operația θ_k^1 și elementele x_1, x_2 astfel ca $x = \theta_k^1(x_1)$ sau $x = \theta_N^1(x_1, x_2)$. Atunci deoarece h' este un homomorfism, avem

$$h'(x) = \theta_k^1(h'(x_1))$$

sau

$$h'(x) = \theta_k^1(h'(x_1), h'(x_2)).$$

Dacă $x_1, x_2 \in \mathcal{C}$, se vede imediat că $h(x) = h'(x)$. Dacă însă $x_1, x_2 \in \mathcal{C}$, asupra acestora se aplică același raționament și deoarece $\mathcal{C}$ este o mulțime de generatori pentru $\mathcal{A}_G^1$ iar elementul x are o lungime finită, din aproape în aproape se ajunge la cazul $x = \theta_1^1, \theta_2^1, \dots, \theta_p^1(x_1, x_2, \dots, x_m)$, unde $x_1, x_2, \dots, x_m \in \mathcal{C}$. Atunci avem

$$h'(x) = \theta_1^1 \theta_2^1 \dots \theta_p^1(h'(x_1), h'(x_2), \dots, h'(x_m)).$$

Deoarece însă h și h' coincid pe mulțimea $\mathcal{C}$ și sînt homomorfisme, deducem ușor că $h(x) = h'(x)$. Această concluzie arată că mulțimea $\mathcal{C}$ este o mulțime de generatori liberi pentru algebra $\mathcal{A}_G^1$, ceea ce demonstrează teorema.

CONSECINȚA 3. Dacă notăm cu $\mathcal{A}_G^i$ subalgebra algebrei $\mathcal{A}_G$ astfel ca

$$\mathcal{A}_G^i = \{[A_{i1}] \cup [A_{i2}] \cup \dots \cup [A_{i\alpha_i}] \cup \{\varepsilon\}, \Omega_j^i\},$$

unde $\Omega_j^i \in \Omega'$ și este mulțimea operațiilor din Ω' definite în $\mathcal{A}_G^i$, atunci $\mathcal{A}_G^i$ este o subalgebră liberă a lui $\mathcal{A}_G$, care are drept mulțime de generatori liberi mulțimea

$$\mathcal{A}_G^{i-1} - \{\varepsilon\}.$$

Demonstrație. Consecința 3 se demonstrează în mod absolut analog ca și teorema 5.

TEOREMA 6. Algebra universală $\mathcal{A}_G$ este o algebră liberă a cărei mulțime de generatori este $\mathcal{A}_G^{N-1} - \{\varepsilon\}$.

Demonstrație. Teorema rezultă imediat din teorema 5 și din consecința 3.

În lucrarea [4] au fost obținute rezultate analoage pentru cazul limbajelor logico-matematice de primul ordin.

Primită la redacție la 5 aprilie 1967

*Institutul de Calcul
al Academiei Republicii Socialiste România — Filiala Cluj*

BIBLIOGRAFIE

1. T. RUS, *Tratarea algebrică a limbajelor formalizate*. St. cerc. mat., **19**, 2, 259—272 (1967).
 2. B. B. БАГНЕР (Editor), *Теория полугрупп и ее приложения*, Сб. статей, Вып. 1, Изд. Саратовского Ун-та, 1965, p. 3—179.
 3. P. M. COHN, *Universal Algebras*. Harper-Row, New-York, 1965.
 4. R. SIKORSKI, *Algebra of formalized languages*. Colloquium Mathematicum, **9**, 1—31 (1962).
-