

LIMBAJE LEONTIEF PROBABILISTE

AURORA BACIU

Universitatea Spiru Haret

Rezumat: În articol se introduce noțiunea de probabilitate la gramaticile Leontief studiate în [1] definind gramaticile Leontief probabiliste, care generează limbaje Leontief probabiliste. Pentru aceste limbaje se studiază două proprietăți de închidere.

Cuvinte cheie: gramatică Leontief probabilistă, limbaj Leontief probabilist.

1. INTRODUCERE

În teoria gramaticilor și limbajelor formale, prin definirea noțiunilor de mulțime fuzzy și de probabilitate, s-au definit conceptele de gramatică fuzzy și de gramatică probabilistică, în scopul creerii unor noi mecanisme cu o putere generativă mai mare decât cele imaginate anterior.

Modelul input-output, conceput de Leontief [3], a sugerat definirea unei noi clase de limbaje, care reprezintă un model lingvistic al sistemelor input-output, și anume clasa limbajelor Leontief.

2. GRAMATICI ȘI LIMBAJE LEONTIEF

Sistemul economic al legăturilor dintre ramuri este un sistem de tip input-output de forma:

$$X = AX + Y,$$

unde $A = (a_{ij})$, $i, j = \overline{1, n}$ este matricea coeficienților cheltuielilor directe, numită matricea Leontief cu $0 \leq a_{ij} \leq 1$, X și Y sunt vectorii $X = (x_1, \dots, x_n)^T$ și $Y = (y_1, \dots, y_n)^T$.

Fie A matricea Leontief corespunzătoare ramurilor R_1, \dots, R_n și două mulțimi M și K .

Definiție. O gramatică Leontief de ordin n este o gramatică

$$G = (V_n, \Sigma, S, P, M = \{M_1, \dots, M_n\}, K = \{K_1, \dots, K_n\}),$$

unde $V_n = \{S, R_1, \dots, R_n\}$ reprezintă mulțimea neterminalelor, $\Sigma = \{a_{ij} / i, j = 1, \dots, n\}$ este alfabetul terminal, S este simbolul start definit de reguli de producție P de forma:

$$\begin{aligned}
S &\rightarrow R_i & \forall i = 1, \dots, n \\
R_i &\rightarrow a_{ij} R_j & \forall j \in M_i \subset \{1, \dots, n\} \\
R_i &\rightarrow a_{ij} & \forall j \in K_j \subset \{1, \dots, n\}
\end{aligned}$$

Este probabil ca $M_i = \emptyset$ sau $K_j = \emptyset$ sau amândouă.

Un limbaj Leontief generat de o gramatică Leontief este notat $L = L(G)$. Clasa limbajelor Leontief L , se notează prin \mathcal{L}_L .

În [1] se arată că limbajele Leontief care reprezintă modele lingvistico-matematice ale sistemelor input-output formează o subclasă a clasei limbajelor regulate.

3. GRAMATICI ȘI LIMBAJE LEONTIEF PROBABILISTE

Definiție. O gramatică Leontief probabilistică este

$$G_p = (V_n, \Sigma, S, P, M, K, x_S, x_{R_1}, \dots, x_{R_n})$$

unde V_n, Σ, S, P, M, K au semnificația prezentată la gramatici Leontief iar $x_S, x_{R_1}, \dots, x_{R_n}$ sunt variabile aleatoare atașate neterminalelor gramaticii G .

În gramatica Leontief probabilistică o regulă de producție poate fi aplicată cu o anumită probabilitate. Toate regulile care au în membrul stâng același terminal R_i sunt considerate ca evenimente ale variabilei aleatoare x_{R_i} .

Un cuvânt generat prin derivatele:

$$D_1 : S \xrightarrow{p_1} \alpha_1 \xrightarrow{p_2} \alpha_2, \dots, \xrightarrow{p_n} \alpha_n = v$$

$$D_2 : S \xrightarrow{q_1} \beta_1 \xrightarrow{q_2} \beta_2, \dots, \xrightarrow{q_n} \beta_n = v$$

cu $p = p_1, p_2, \dots, p_n$ și $q = q_1, q_2, \dots, q_n$, atunci în gramatica G cuvântul v va fi generat cu probabilitatea $p \oplus q$, unde:

$$p \oplus q = p + q - pq$$

Se poate demonstra teorema:

Teorema 1. Clasa limbajelor Leontief \mathcal{L}_L este inclusă în clasa limbajelor Leontief probabiliste, notată \mathcal{L}_{LP} , generată de gramatica Leontief probabilistă.

Demonstrație. Fie $L \subset \mathcal{L}_L$ și $G = (V_n, \Sigma, S, P, M, K)$ gramatica Leontief de ordin n care generează limbajul L .

Fie o gramatică Leontief probabilistă G_p unde :

$$x_s : \begin{pmatrix} S \rightarrow R_1 \dots S \rightarrow R_n \\ 1/n \quad 1/n \end{pmatrix}$$

Dacă $\text{card}(M_i \cup K_i) = m_i + k_i$

$$x_{R_i} : \left(\begin{array}{cccc} R_1 \rightarrow a_{ij_1} R_{j_1}, \dots, R_i \rightarrow a_{ij_s} R_{j_s} & R_i \rightarrow a_{ij_{k_1}}, \dots, R_i \rightarrow a_{ij_{k_1}} \\ \frac{1}{m_i + k_i}, \dots, \frac{1}{m_i + k_i} & \frac{1}{m_i + k_i}, \dots, \frac{1}{m_i + k_i} \end{array} \right)$$

Se arată că orice cuvânt din $L = L(G)$ este generat de G_p cu probabilitatea

$$\frac{1}{n} \cdot \frac{1}{m_{i_1} + k_{i_1}} \cdot \frac{1}{m_{i_2} + k_{i_2}} \cdot \dots \cdot \frac{1}{m_{i_{r-1}} + k_{i_{r-1}}} \cdot \frac{1}{m_{i_r} + k_{i_r}} = \frac{1}{n(m_{i_1} + k_{i_1})(m_{i_2} + k_{i_2}) \dots (m_{i_{r-1}} + k_{i_{r-1}})(m_{i_r} + k_{i_r})}$$

Gramatica G_p este consistentă în sensul definiției din [4].

L nu este un limbaj Leontief [1], dar este un limbaj Leontief probabilist fiind generat de gramatica

$$G_p = (\{S, R_1, R_2\}, \{a_{11}, a_{12}, a_{21}, a_{22}\}, P, S, M = (M_1, M_2), K = (K_1, K_2), x_S, x_{R_1}, x_{R_2})$$

unde P :

$$\begin{array}{ll} S \rightarrow R_1 & \text{cu probabilitatea } 1 \\ S \rightarrow R_2 & \text{cu probabilitatea } 0 \\ R_1 \rightarrow a_{11}R_1 & \text{cu probabilitatea } 1/4 \\ R_1 \rightarrow a_{12}R_2 & \text{cu probabilitatea } 1/4 \\ R_2 \rightarrow a_{21}R_1 & \text{cu probabilitatea } 1 \\ R_1 \rightarrow a_{12}R_2 & \text{cu probabilitatea } 1/4 \\ R_1 \rightarrow a_{12} & \text{cu probabilitatea } 1/4 \end{array}$$

$$M_1 = \{1, 2\} \quad K_2 = \{1\}$$

$$x_S : \left(\begin{array}{cc} S \rightarrow R_1 & S \rightarrow R_2 \\ 1 & 0 \end{array} \right)$$

$$x_{R_1} : \left(\begin{array}{cccc} R_1 \rightarrow a_{11}R_1 & R_1 \rightarrow a_{12}R_2 & R_1 \rightarrow a_{12}R_2 & R_1 \rightarrow a_{12} \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right)$$

$$x_{R_2} : \left(\begin{array}{c} R_2 \rightarrow a_{21}R_1 \\ 1 \end{array} \right)$$

Se observă că cuvântul $a_{11}^n a_{12} (a_{21} a_{12})^n$ va fi generat cu probabilitatea

$$\left(\frac{1}{4}\right)^n \cdot \frac{1}{4} \cdot \left(1 \cdot \frac{1}{4}\right)^n = \frac{1}{4^{2n+1}}. \quad \text{Deci } L = \left\{ a_{11}^n a_{12} (a_{21} a_{12})^n, \frac{1}{4^{2n+1}} \mid n \geq 1 \right\} \quad \text{ceea ce}$$

demonstrează teorema.

Teorema 2. Fie $L_1, L_2 \in \mathcal{L}_{\mathcal{G}_p}$ peste același alfabet. Atunci și $L_1 \cup L_2$ și $L_1 \cap L_2 \in \mathcal{L}_{\mathcal{G}_p}$.

Demonstrație. Fie $L_1 = L(G_p^1)$ și $L_2 = L(G_p^2)$ unde $G_p^1 = (V_N, \Sigma, P^1, S, x_{R_1}^1, \dots, x_{R_n}^1)$ și $G_p^2 = (V_N, \Sigma, P^2, S, x_{R_1}^2, \dots, x_{R_n}^2)$.

Atunci $L_1 \cup L_2 = L(G_p)$, unde $G_p = (V_n, \Sigma, P, S, x_{R_1}, \dots, x_{R_n})$ cu $P = P^1 \cup P^2$ și $x_{R_i} = x_{R_i}^1 + x_{R_i}^2$, $i = \overline{1, n}$.

Limbajul $L_1 \cap L_2$ va fi generat de gramatica G_p^1 care este definită: $G_p^1 = (V_n, \Sigma, S, P^1, x_{R_1}^1, \dots, x_{R_n}^1)$ unde $P^1 = P_1 \cap P_2$.

Fie $q_{ir} = \max(P_{ir}^1, P_{ir}^2)$ unde P_{ir}^1 este probabilitatea cu care se aplică regula $R_i \rightarrow a_{ij_r} R_{j_r}$ cu gramatica G_p^1 și P_{ir}^2 în G_p^2

Atunci

$$x_{R_i}^1 : \left(\begin{array}{c} R_i \rightarrow a_{ij_1} R_{j_1} \dots R_i \rightarrow a_{ij_k} R_{j_k} \\ \hline q_{ik} / \sum_{r=1}^k q_{ir} \quad \dots \quad q_{ik} / \sum_{r=1}^k q_{ir} \end{array} \right)$$

Limbajele Leontief probabiliste modelează sisteme input-output probabiliste care țin seama de aspectul natural al interdependențelor dintre ramurile unui sistem economic, interdependențe care la un moment dat apar cu anumite probabilități.

BIBLIOGRAFIE

1. Baciu A., Pascu A., *Leontief Type Grammars and Languages*, Foundations of Control Engineering, Nr. 3, Vol. 1, (1979), 91-106.
2. Bucurescu I., Pascu A., *Gramatici contextuale fuzzy*, Rev. Roum. Math. Pures et. Appl, (1989).
3. Leontief W., *Input-output Analysis*, Scientific Publishing House, (1970).
4. Rafael C., Gonzales G., *Syntactic Pattern Recognition in Introduction Addison Wesley*, Publishing Company, Massachusettes, (1978).
5. Tilanus C.B., *Intra Firm Input-Output Analysis*, Proceedings of Euro II, (1976).

Abstract: This paper introduces the probability for Leontief grammars studied in [1] defining probabilistic Leontief grammars generating probabilistic Leontief languages. Two closing properties are studied.