

A számításelmélet alapjai I.

2. gyakorlat

Boda Bálint

2023. tavaszi félév

1. Generatív grammatika

Definíció (grammatika). Egy $G = (N, T, P, S)$ rendezett négyest (generatív) grammatikának vagy nyelvtannak nevezünk, ha N és T diszjunkt (azaz $N \cap T = \emptyset$) véges ábécék. Ekkor

- N a nem terminális szimbólumok halmaza,
- T (vagy Σ) a terminális szimbólumok halmaza,
- $S \in N$ a grammatika kezdőszimbóluma,
- $P = \{(x, y) \mid x, y \in (N \cup T)^* \text{ szavak úgy, hogy } x \text{ legalább egy nem terminális betűt tartalmaz}, \text{ az ún. (átírási) szabályok (vagy produkciók) halmaza.}$

Jelölés. Gyakran (x, y) helyett az $x \rightarrow y$ jelölést használjuk, egy szabály leírására. Természetesen ez csak akkor lehetséges ha az adott ábécének nem eleme \rightarrow .

Definíció (egylépéses levezetés). Legyen $G = (N, T, P, S)$ egy grammatika és legyen $u, v \in (N \cup T)^*$. Azt mondjuk v közvetlen levezethető az u szóból G -ben (jelekkel: $u \Rightarrow_G v$), ha

$$\exists (x, y) \in P : u = u_1 x u_2 \text{ és } v = u_1 y u_2, \quad (u_1, u_2 \in (N \cup T)^*)$$

Definíció (többlépéses levezetés). Legyen $G = (N, T, P, S)$ egy grammatika és legyen $u, v \in (N \cup T)^*$. Azt mondjuk v több lépésben levezethető az u szóból G -ben (jelekkel: $u \Rightarrow_G^* v$), ha

$$u = v \vee \exists (n \geq 1 \wedge w_0, \dots, w_n \in (N \cup T)^*), \text{ hogy } w_{i-1} \Rightarrow_G w_i \ (1 \leq i \leq n), \ w_0 = u \text{ és } w_n = v$$

Definíció (generált nyelv). Legyen $G = (N, T, P, S)$ egy grammatika, ekkor a G által generált nyelvnek nevezzük az S kezdőszimbólumból több lépésben levezethető terminális szavak halmazát, azaz a

$$L(G) = \{u \in T^* \mid S \Rightarrow_G^* u\}$$

nyelvet.

Példa. Legyen

$$G = (\{S, A, B\}, \{a, b\}, P, S) \\ P = \{S \rightarrow B|bb, B \rightarrow aaA, A \rightarrow a|\varepsilon\}$$

Adjuk meg $L(G)$ -t!

Megjegyzés. Egy $S \rightarrow B|bb$ az $S \rightarrow B$ és $S \rightarrow bb$ szabályokat jelöli.

$$S \rightarrow bb$$

$$S \rightarrow B \rightarrow aaA \rightarrow a$$

$$S \rightarrow B \rightarrow aaA \rightarrow \varepsilon$$

Így: $L(G) = \{bb, aa, a\}$.

1.1. Feladatok

1. Legyen $G_i = (\{S, A, B\}, \{a, b\}, P_i, S)$. Határozzuk meg az $L(G_i)$ nyelvet, ha

- $P_1 = \{S \rightarrow aaS|a\}$
- $P_2 = \{S \rightarrow aSb|\varepsilon\}$
- $P_3 = \{S \rightarrow ASB|\varepsilon, AB \rightarrow BA, BA \rightarrow AB, A \rightarrow a, B \rightarrow b\}$

Megoldás.

$$L(G_1) = \{a, aaa, aaaaa \dots\} = \{a^{(2n+1)} \mid n \geq 0\}$$

$$L(G_2) = \{\varepsilon, ab, aabb \dots\} = \{a^n b^n \mid n \geq 0\}$$

A harmadik nyelv meghatározása már nehezebb feladat. Tekintsünk pár példa levezetést:

$$\begin{aligned} S &\rightarrow \underline{ASB} \rightarrow \underline{AB} \rightarrow a\underline{B} \rightarrow ab \\ S &\rightarrow ASB \rightarrow AB \rightarrow BA \rightarrow bA \rightarrow ba \\ S &\rightarrow \underline{ASB} \rightarrow A\underline{ASBB} \rightarrow A\underline{ABBB} \rightarrow \underline{ABAB} \rightarrow BA\underline{AAB} \rightarrow \underline{BABA} \rightarrow \dots \rightarrow baba \\ S &\rightarrow ASB \rightarrow AASBB \rightarrow AABBB \rightarrow ABAB \rightarrow BAAB \rightarrow baab \end{aligned}$$

Ezek alapján $L(G_3) = \{u \in \{a, b\}^* \mid l_a(u) = l_b(u)\}$.

2. A grammatikák Chomsky féle osztályozása

Legyen $G = (N, T, P, S)$ egy grammatika. A G grammatika i -típusú ($i = 0, 1, 2, 3$), ha a P szabályhalmazra teljesülnek a következők:

- $i = 0$ (mondatszerkezetű grammatika): nincs korlátozás
- $i = 1$ (környezetfüggő grammatika):
 - P minden szabálya $u_1 A u_2 \rightarrow u_1 v u_2$ alakú, ahol $u_1, u_2, v \in (N \cup T)^*$, $A \in N$ és $v \neq \varepsilon$
 - Kivétel: P tartalmazhatja az $S \rightarrow \varepsilon$ szabályt, de csak akkor, ha S nem fordul elő egyetlen szabály jobb oldalán sem.
- $i = 2$ (környezetfüggetlen): P minden szabálya $A \rightarrow v$ alakú ($A \in N$, $v \in (N \cup T)^*$)
- $i = 3$ (reguláris): P minden szabálya $A \rightarrow uB$ vagy $A \rightarrow u$ alakú ($A, B \in N$, $u \in T^*$)

Az adott osztályokat \mathcal{G}_i -vel jelöljük.

Definíció (nyelvosztály). Az i típusú nyelvek osztályának nevezzük a

$$\mathcal{L}_i = \{L \mid \exists G \in \mathcal{G}_i, \text{ hogy } L = L(G)\}$$

1. Tétel (Chomsky nyelvhierarchia tétel).

$$\mathcal{L}_3 \subset \mathcal{L}_2 \subset \mathcal{L}_1 \subset \mathcal{L}_0$$

2.1. Feladatok

1. Írjuk fel azt a grammatikát, mely a 4-el osztható bináris számok nyelvét generálja! Milyen osztályba sorolható a generált nyelv?

Megoldás.

Egy kettes számrendszerbeli szám akkor osztható négygyel, ha utolsó két számjegye 0. Gondoskodnunk kell továbbá arról, hogy ne legyenek felesleges nullák az elején. Így

$$G = \left(\{S, B\}, \{0, 1, \varepsilon\}, \left\{ S \rightarrow \underbrace{0}_{3.} | \underbrace{1B00}_{2.}, B \rightarrow \underbrace{\varepsilon}_{3.} | \underbrace{0B}_{3.} | \underbrace{1B}_{3.} \right\}, S \right)$$

(Az adott szabály jobb oldala alatt tüntettem fel annak szintjét.) Mivel kettes a legkisebb szint ezért a generált nyelv is 2-es szintű.

A feladat megoldható más módon is:

$$G = (\{S, A\}, \{0, 1, \varepsilon\}, \{S \rightarrow 0|1A, A \rightarrow \varepsilon|0A|1A\}, S)$$

amiből már reguláris nyelv adódik.

2. Írjuk fel azt a grammatikát, ami az $L(G) = \{a^n, b^m, c^n \mid n \geq 0, m \geq 3\}$ nyelvet generálja!

3. Reguláris kifejezések