A számításelmélet alapjai I.
(Második gyakorlat)

Dr. Lázár Katalin Anna

Eötvös Loránd Tudományegyetem, Informatikai Kar
1117 Budapest, Pázmány Péter sétány 1/C.
e-mail: lazarkati@elte.hu

2024. február 20.

Tematika

- A generatív grammatika fogalma, generatív grammatikák típusai.
- A levezetés fogalma (közvetlen (egylépéses) levezetés, k lépéses levezetés, levezetés, mondatforma), generált nyelv.
- Chomsky-féle hierarchia.
- Nyelvosztályok zártsági tulajdonságai.

Generatív grammatika

Példa 1

Generatív grammatikák-e a következők? Indokoljuk!

1. $G_1 = (\{S, A, B\}, \{a, b\}, \{A \to aAb, A \to ab, B \to ba\}, S)$.
2. $G_2 = (\{S, A, B\}, \{a, b\}, \{S \to \varepsilon, S \to AB, A \to aA, B \to ab, abb \to aSb\}, S)$.
3. $G_3 = (\{S, A, B\}, \{a, b\}, S)$.
4. $G_4 = (\{S\}, \{a, b\}, \{S \to \varepsilon, S \to SS, S \to aSb, S \to bSa\}, S)$.

Generatív grammatika

Példa 1

1. IGEN (G_1 az $\emptyset$ nyelvet generálja).
2. NEM, G_2 $abb \to aSb$ szabály miatt nem generatív grammatika (minden szabály bal oldalán legalább egy nemterminálisnak szerepelnie kell).
3. NEM, G_3 nem generatív grammatika, mert nem adtuk meg a szabályhalmazát.
4. IGEN, $N = \{S\}$, $T = \{a, b\}$, $P = \{S \to \varepsilon, S \to SS, S \to aSb, S \to bSa\}$ szabályai megfelelő alakúak (minden szabály bal oldalán legalább egy nemterminális szerepel), S a kezdőszimbólum.

Generált nyelv

Definíció

Legyen $G = \langle N, \Sigma, P, S \rangle$ egy tetszőleges grammatika. A G által **generált nyelv** alatt az $L(G) := \{w \mid S \Rightarrow_G^* w, w \in \Sigma^*\}$ szavakból álló halmazt értjük.

Példa 2

Milyen nyelvet generálnak a következő grammatikák? Adjunk példát egy-egy lehetséges levezetésre!

1. $G_1 = (N, T, P, S)$, ahol $N = \{S\}$, $T = \{a, b\}$ és $P = \{S \to \varepsilon, S \to SS, S \to aS, S \to Sa, S \to aSb, S \to bSa\}$.
2. $G_2 = (N, T, P, S)$, ahol $N = \{S\}$, $T = \{a, b\}$ és $P = \{S \to \varepsilon, S \to aSa, S \to bSb\}$.
3. $G_3 = (N, T, P, S)$, ahol $N = \{S, A, B\}$, $T = \{a, b, c\}$ és $P = \{S \to \varepsilon, S \to AB, A \to \varepsilon, A \to aAb, B \to \varepsilon, B \to bBc\}$.

Generált nyelv

Példa 2

1. $L(G_1) = \{u \in \{a,b\}^* \mid |u|_a \geq |u|_b\}$, ahol $|u|_a$, $|u|_b$ a és b u-beli előfordulásainak számát jelöli. Példa egy lehetséges G_1-beli levezetésre: $S \Longrightarrow_{G_1} SS \Longrightarrow_{G_1} aSS \Longrightarrow_{G_1} aaSbS \Longrightarrow_{G_1} aabS \Longrightarrow_{G_1} aab$.
2. $L(G_2) = \{uu^R \mid u \in \{a,b\}^*\}$. Példa egy lehetséges G_2-beli levezetésre: $S \Longrightarrow_{G_2} aSa \Longrightarrow_{G_2} aaSaa \Longrightarrow_{G_2} aabSbaa \Longrightarrow_{G_2} aabbaa$.
3. $L(G_3) = \{a^k b^{k+l} c^l \mid k,l \geq 0\}$. Példa egy lehetséges G_3-beli levezetésre: $S \Longrightarrow_{G_3} AB \Longrightarrow^k_{G_3} a^k Ab^k B \Longrightarrow^l_{G_3} a^k Ab^k b^l Bc^l \Longrightarrow_{G_3} a^k b^k b^l Bc^l \Longrightarrow_{G_3} a^k b^k b^l c^l = a^k b^{k+l} c^l$ (k- és l-lépéses levezetéseket alkalmaztunk, $k,l \geq 0$).

Generált nyelv

Példa 3

Legyen $G_4 = (N, T, P, S)$, ahol $N = \{S\}$, $T = \{a, b\}$ és $P = \{S \to \varepsilon, S \to SS, S \to aSb, S \to bSa\}$. Milyen nyelvet generál a G_4 grammatika? Bizonyítsuk!

Generált nyelv

Példa 3

A generált nyelv: $L(G_4) = \{u \in \{a,b\}^* \mid |u|_a = |u|_b\}$, ahol $|u|_a$, $|u|_b$ a és b u-beli előfordulásainak számát jelöli. Bizonyítás (vázlat):

- $\subseteq$: A levezetés hosszára vonatkozó teljes indukcióval belátjuk, hogy minden levezetett $p \in \{a,b,S\}^*$ mondatformára $|p|_a = |p|_b$. Az $n = 0$ eset nyilvánvaló. Tegyük fel, hogy minden n-nél rövidebb levezetésre igaz az állítás. Legyen $S \Longrightarrow^n p$. Ekkor létezik egy olyan $q \in \{a,b,S\}^*$ mondatforma, hogy $S \Longrightarrow^{n-1} q \Longrightarrow p$. Az indukciós feltevés alapján $|q|_a = |q|_b$. Mivel $q \Longrightarrow p$, ezért léteznek olyan $y_1, y_2 \in \{a,b,S\}^*$ mondatformák, amelyekre $q = y_1 S y_2$, $p = y_1 w y_2$, ahol $w \in \{\varepsilon, SS, aSb, bSa\}$. Az első és második esetben $|q|_a = |p|_a$ és $|q|_b = |p|_b$, míg a harmadik és negyedik esetben $|p|_a = |q|_a + 1$ és $|p|_b = |q|_b + 1$, vagyis $|p|_a = |p|_b$ mind a négy esetben.

Generált nyelv

Példa 3

- $\supseteq$: Legyen $u \in L_4 = \{\{a, b\}^* \mid |u|_a = |u|_b\}$. u hosszára vonatkozó teljes indukcióval bizonyítjuk, hogy u levezethető. Ha $|u| = 0$, akkor az állítás triviális. Tegyük fel, hogy minden u-nál rövidebb L_4-beli szó levezethető. Négy eset lehetséges: $u = \varepsilon, u = avb, u = bva$ (ekkor v nyilván L_4-beli), vagy pedig létezik $u_1, u_2 \in L_4$, $u_1 \neq \varepsilon, u_2 \neq \varepsilon$: $u = u_1 u_2$. Ekkor az indukciós feltevés alapján $S \Longrightarrow^* u_1$ és $S \Longrightarrow^* u_2$, tehát $S \Longrightarrow SS \Longrightarrow^* u_1 S \Longrightarrow^* u_1 u_2 = u$.