

Nyelvek felismerése

Formális nyelvek, 5. gyakorlat

Célja: A nyelvtanokkal való nyelvmegadás gyakorlása, a nyelvtani típusok, illetve a nyelvtípusok fogalmának elmélyítése.

Fogalmak: A Chomsky-féle nyelvtanosztályozás és nyelvosztályozás, Chomsky-hierarchia (gyenge, erős).

Feladatok jellege: Egyszerűbbtől a bonyolultabbig nyelvekhez nyelvtan, nyelvtanhoz a generált nyelv, a nyelvtan típusának detektálásával. 1 példán a kétirányú tartalmazás bizonyítása.

2008/09 I. félév

Házi feladatok megoldása

1. feladat

Adjunk a következő nyelvet generáló 3. típusú nyelvtant!

Azon M -áris számrendszerbeli számok, melyek d -vel osztva k maradékot adnak. (Nem állhat az elején 0.)

Megoldás:

$$\begin{aligned} S &\rightarrow iS_i \pmod{d} & \forall 1 \leq i \leq M-1 \\ S_i &\rightarrow jS_{Mi+j \pmod{d}} & \forall 0 \leq i, j \leq M-1 \\ S_k &\rightarrow \varepsilon \end{aligned}$$

Házi feladatok megoldása

2. feladat

Adjunk a következő nyelvet generáló 3. típusú nyelvtant!

$L = \{u \in \{a, b, c\}^*; ab, bc, ca \not\subseteq u\}$.

Megoldás:

Tiltott = $\{ab, bc, ca\}$

$$\begin{aligned} S &\rightarrow tS_t & \forall t \in T \\ S_t &\rightarrow t'S_{t'} & \forall t, t' \in T, tt' \notin \text{Tiltott} \\ S &\rightarrow \varepsilon \\ S_t &\rightarrow \varepsilon & \forall t \in T \end{aligned}$$

Házi feladatok megoldása

3. feladat

Adjunk nyelvtant! $T = \{a\}$

$L = \{a^{2^n}; n \geq 0\}$.

Megoldás:

$$\begin{aligned} S &\rightarrow LDaR \\ LD &\rightarrow \varepsilon \\ R &\rightarrow \varepsilon \\ Da &\rightarrow aaD \\ DR &\rightarrow ER \\ ER &\rightarrow \varepsilon \\ L &\rightarrow \varepsilon \\ aE &\rightarrow Eaa \\ LE &\rightarrow LD \end{aligned}$$

Nyelvtanok típusai

Típus	Megszorított típus szabályai	Alaptípus szabályai
0.	nincs további megkötés	
1.	$\gamma_1 A \gamma_2 \rightarrow \gamma_1 q \gamma_2$, ahol $\gamma_1, \gamma_2 \in (T \cup N)^*$, $A \in N, q \in (T \cup N)^+$; $S \rightarrow \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$p \rightarrow q$ ahol $\ell(p) \leq \ell(q)$; $S \rightarrow \varepsilon$, ez esetben S nem szerepel szabály jobboldalán
2.	$A \rightarrow q$, ahol $A \in N, q \in (T \cup N)^+$; $S \rightarrow \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$A \rightarrow q$, ahol $A \in N, q \in (T \cup N)^*$
3.	$A \rightarrow aB$ vagy $A \rightarrow a$ ahol $A, B \in N, a \in T$; $S \rightarrow \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$A \rightarrow uB$ vagy $A \rightarrow u$, ahol $A, B \in N$, és $u \in T^*$

Nyelvek típusai

Jelölje \mathcal{G}_i illetve $\mathcal{G}_{\text{megsz}_i}$, ($i = 0, 1, 2, 3$) az i . típusú illetve a megszorított i . típusú nyelvtanok osztályát.

$$\mathcal{L}_{(\text{megsz})_i} = \{L \mid L \text{ nyelv, és van olyan } G \in \mathcal{G}_{(\text{megsz})_i}, \text{ amelyre } L(G) = L\}. \quad (i = 0, 1, 2, 3)$$

Kiterjesztési tétel

$$\mathcal{L}_i = \mathcal{L}_{\text{megsz}_i}, \quad (i = 1, 2, 3).$$

Az 1. típusú nyelveket **környezetfüggő**, a 2. típusúakat **környezetfüggetlen**, a 3. típusúakat **reguláris** nyelveknek is nevezzük. Az utóbbit a Kleene tétel miatt:

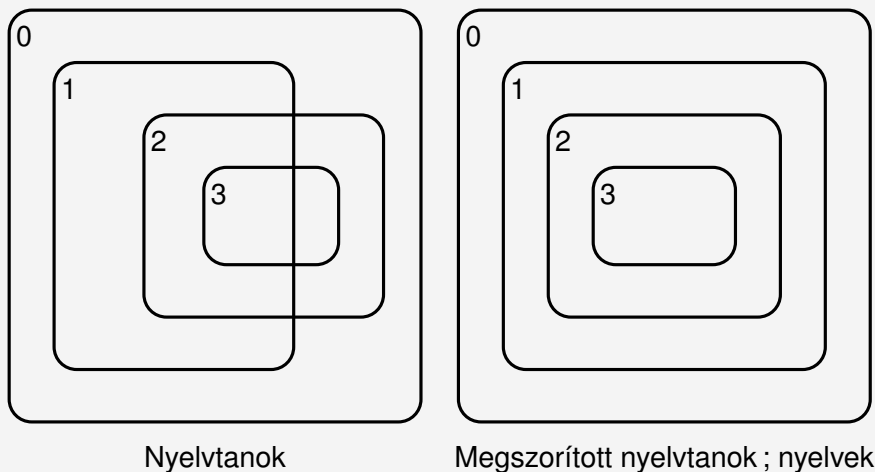
\mathcal{L}_{REG} a reguláris kifejezéssel leírható nyelvek (**reguláris nyelvek**) nyelvcsaládja.

Kleene tétele

$$\mathcal{L}_{\text{REG}} = \mathcal{L}_3.$$

Chomsky nyelvhierarchia

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



Milyen típusúak a következő nyelvtanok?

- $S \rightarrow \varepsilon \mid aSb \mid bSa \mid SS$
2. típus (nem megszorított)
- $S \rightarrow \varepsilon \mid aSB \mid BSa \mid SS$
 $B \rightarrow \varepsilon \mid b$
2. típus (nem megszorított)
- $S \rightarrow \varepsilon \mid aS \mid aSb \mid bSa \mid Sa \mid SS$
2. típus (nem megszorított)
- $S \rightarrow tSt \mid \varepsilon \quad \forall t \in T$
2. típus (nem megszorított)
- $S \rightarrow X_t t \mid X_t t S'$
 $S' \rightarrow t Y_t \mid t Y_t S'$
 $t Y_{t'} \rightarrow Y_{t'} t$
 $X_t Y_{t'} \rightarrow t X_{t'}$
 $X_t \rightarrow t \quad \forall t, t' \in T$
1. típus (nem megszorított)

Melyik nyelvet generálják a következő nyelvtanok?

$$1. S \rightarrow \varepsilon \mid aSb \mid bSa \mid SS$$

$$L_1 = \{u \in \{a, b\}^* \mid \ell_a(u) = \ell_b(u)\}$$

$$2. S \rightarrow \varepsilon \mid aSB \mid BSa \mid SS$$

$$B \rightarrow \varepsilon \mid b$$

$$L_2 = \{u \in \{a, b\}^* \mid \ell_a(u) \geq \ell_b(u)\}$$

$$3. S \rightarrow \varepsilon \mid aS \mid aSb \mid bSa \mid Sa \mid SS$$

$$L_3 = \{u \in \{a, b\}^* \mid \ell_a(u) \geq \ell_b(u)\}$$

$$4. S \rightarrow tSt \mid \varepsilon \quad \forall t \in T$$

$$L_4 = \{u \mid u = vv^{-1}, v \in T^*\}$$

$$5. S \rightarrow X_t t \mid X_t t S'$$

$$S' \rightarrow t Y_t \mid t Y_t S'$$

$$t Y_{t'} \rightarrow Y_{t'} t$$

$$X_t Y_{t'} \rightarrow t X_{t'}$$

$$X_t \rightarrow t \quad \forall t, t' \in T$$

$$L_5 = \{v \mid v = uu, u \in T^+\} \quad (\text{dadogós nyelv})$$

Ugyanannyi a -t és b -t tartalmazó szavak nyelve

Megoldás (vázlat):

Az első G nyelvtanra belátjuk, hogy $L(G) = L_1$.

“ \subseteq ”

A levezetés hosszára vonatkozó teljes indukcióval belátjuk, hogy minden levezetett $\beta \in \{a, b, S\}^*$ mondatformára $\ell_a(\beta) = \ell_b(\beta)$. Az $n = 0$ eset nyilvánvaló. Tegyük fel, hogy minden n -nél rövidebb levezetésre igaz az állítás. Legyen $S \xrightarrow{n} \beta$. Ekkor létezik egy olyan $\alpha \in \{a, b, S\}^*$ mondatforma, hogy $S \xrightarrow{n-1} \alpha \rightarrow \beta$. Az indukciós feltevés alapján $\ell_a(\alpha) = \ell_b(\alpha)$. Mivel $\alpha \rightarrow \beta$, ezért léteznek $\gamma_1, \gamma_2 \in \{a, b, S\}^*$, hogy $\alpha = \gamma_1 S \gamma_2$ és $\beta = \gamma_1 q \gamma_2$, ahol $q \in \{\varepsilon, aSb, bSa, SS\}$. Az első és a negyedik esetben ugyanannyi az α -ban és a β -ban szereplő a -k és b -k száma, míg a második és harmadik esetben β -ban eggyel több a és b van, mint α -ban, tehát $\ell_a(\beta) = \ell_b(\beta)$ mind a négy esetben.

(folyt.) és a dadogós szavak nyelve

“ \supseteq ”

Legyen $u \in L_1$. $\ell(u)$ -ra vonatkozó teljes indukcióval bizonyítjuk, hogy u levezethető. Ha $\ell(u) = 0$, akkor nyilvánvaló. Tegyük fel, hogy minden $\ell(u)$ -nál rövidebb L_1 -beli szó levezethető. Négy eset lehetséges: vagy $u = \varepsilon$, vagy $u = au'b$, vagy $u = bu'a$ (ezen esetekben nyilván $u' \in L_1$), vagy pedig létezik $u', u'' \in L_1$, $u' \neq \varepsilon$, $u'' \neq \varepsilon$, hogy $u = u'u''$. Például a negyedik esetben ekkor az indukciós feltevés alapján $S \xrightarrow{*} u'$ és

$S \xrightarrow{*} u''$, tehát $S \rightarrow SS \xrightarrow{*} u'S \xrightarrow{*} u'u'' = u$. (A többi eset hasonló.)

Megoldás (vázlat):

Az ötödik nyelvtan szabályainak segítségével ($S \rightarrow X_t t \rightarrow tt$ kivételével) csak az alábbi módon vezethetünk le S -ből terminális szót:

$$S \rightarrow X_t t S' \xrightarrow{*} X_t t t_1 Y_{t_1} \cdots t_k Y_{t_k} \xrightarrow{*} X_t Y_{t_1} t t_1 t_2 Y_{t_2} \cdots t_k Y_{t_k} \rightarrow \\ \rightarrow t X_{t_1} t t_1 t_2 Y_{t_2} \cdots t_k Y_{t_k} \xrightarrow{*} t t_1 \cdots t_{k-1} X_{t_k} t t_1 \cdots t_k \rightarrow t t_1 \cdots t_{k-1} t_k t t_1 \cdots t_k.$$

Nyelvtani transzformációk I.

Nyelvtanosztályok zárttsági tulajdonságai (1. típus)

1. Feladat

Készítsünk olyan 1. típusú G nyelvtant, melyre $L(G) = L(G_1) \cup L(G_2)$; $L(G) = L(G_1)L(G_2)$; $L(G) = L(G_1)^*$.

$$G_1 = \langle \{a, b\}, \{S, A, B\}, \{S \rightarrow \varepsilon \mid ASB, AAB \rightarrow aaa, Aa \rightarrow aA, \\ B \rightarrow b\}, S \rangle$$

$$G_2 = \langle \{a\}, \{S, A, B\}, \{S \rightarrow ABS \mid A, ABA \rightarrow aaaa, BA \rightarrow aaa\}, S \rangle$$

Megoldás

Először is átnevezzük a nyelvtani jeleket:

$$G'_1 = \langle \{a, b\}, \{S_1, A_1, B_1\}, \{S_1 \rightarrow \varepsilon \mid A_1 S_1 B_1, A_1 A_1 B_1 \rightarrow aaa, \\ A_1 a \rightarrow a A_1, B_1 \rightarrow b\}, S_1 \rangle$$

$$G'_2 = \langle \{a\}, \{S_2, A_2, B_2\}, \{S_2 \rightarrow A_2 B_2 S_2 \mid A_2, A_2 B_2 A_2 \rightarrow aaaa, \\ B_2 A_2 \rightarrow aaa\}, S_2 \rangle$$

Nyelvtani transzformációk I.

Nyelvtanosztályok zárttsági tulajdonságai (1. típus)

Unió:

$$G_{\cup} = \langle \{a, b\}, \{S, S_1, A_1, B_1, S_2, A_2, B_2\}, \{S \rightarrow S_1 \mid S_2 \mid \varepsilon, S_1 \rightarrow A_1 S_1 B_1, A_1 A_1 B_1 \rightarrow aaa, A_1 a \rightarrow aA_1, B_1 \rightarrow b, S_2 \rightarrow A_2 B_2 S_2 \mid A_2, A_2 B_2 A_2 \rightarrow aaaa, B_2 A_2 \rightarrow aaa\}, S \rangle$$

Konkatenáció:

$$G_{\text{konk}} = \langle \{a, b\}, \{S, S_1, A_1, B_1, S_2, A_2, B_2\}, \{S \rightarrow S_1 S_2 \mid S_2, S_1 \rightarrow A_1 S_1 B_1, A_1 A_1 B_1 \rightarrow aaa, A_1 a \rightarrow aA_1, B_1 \rightarrow b, S_2 \rightarrow A_2 B_2 S_2 \mid A_2, A_2 B_2 A_2 \rightarrow aaaa, B_2 A_2 \rightarrow aaa\}, S \rangle$$

Lezárás:

$$G_1^* = \langle \{a, b\}, \{S_1, A_1, B_1\}, \{S \rightarrow \varepsilon \mid S_1 \mid S_1 S', S_1 \rightarrow \varepsilon \mid A_1 S_1 B_1, A_1 A_1 B_1 \rightarrow aaa, A_1 a \rightarrow aA_1, B_1 \rightarrow b, aS' \rightarrow aS_1, bS' \rightarrow bS_1, aS' \rightarrow aS_1 S', bS' \rightarrow bS_1 S'\}, S \rangle$$

Nyelvtani transzformációk I.

Nyelvtanosztályok zárttsági tulajdonságai (2. típus)

2. Feladat

Készítsünk olyan 2. típusú G nyelvtant, melyre $L(G) = L(G_1) \cup L(G_2)$;
 $L(G) = L(G_1)L(G_2)$; $L(G) = L(G_2)^*$.

$$G_1 = \langle \{a\}, \{S\}, \{S \rightarrow Saa \mid a\}, S \rangle$$
$$G_2 = \langle \{a, b\}, \{S, A\}, \{S \rightarrow aSA \mid \varepsilon, A \rightarrow b \mid \varepsilon\}, S \rangle$$

Megoldás

Először is átnevezzük a nyelvtani jeleket:

$$G'_1 = \langle \{a\}, \{S_1\}, \{S_1 \rightarrow S_1 aa \mid a\}, S_1 \rangle$$
$$G'_2 = \langle \{a, b\}, \{S_2, A_2\}, \{S_2 \rightarrow aS_2 A_2 \mid \varepsilon, A_2 \rightarrow b \mid \varepsilon\}, S_2 \rangle$$

Nyelvtani transzformációk I.

Nyelvtanosztályok zárttsági tulajdonságai (2. típus)

Unió:

$$G_{\cup} = \langle \{a, b\}, \{S, S_1, S_2, A_2\}, \{S \rightarrow S_1 \mid S_2, S_1 \rightarrow S_1 aa \mid a, S_2 \rightarrow aS_2 A_2 \mid \varepsilon, A_2 \rightarrow b \mid \varepsilon\}, S \rangle$$

Konkatenáció:

$$G_{\text{konk}} = \langle \{a, b\}, \{S, S_1, S_2, A_2\}, \{S \rightarrow S_1 S_2, S_1 \rightarrow S_1 aa \mid a, S_2 \rightarrow aS_2 A_2 \mid \varepsilon, A_2 \rightarrow b \mid \varepsilon\}, S \rangle$$

Lezárás:

$$G_2^* = \langle \{a, b\}, \{S, S_2, A_2\}, \{S \rightarrow S_2 S \mid \varepsilon, S_2 \rightarrow aS_2 A_2 \mid \varepsilon, A_2 \rightarrow b \mid \varepsilon\}, S \rangle$$

Nyelvtani transzformációk I.

Nyelvtanosztályok zárttsági tulajdonságai (3. típus)

3. Feladat

Készítsünk olyan 3. típusú G nyelvtant, melyre $L(G) = L(G_1) \cup L(G_2)$;
 $L(G) = L(G_1)L(G_2)$; $L(G) = L(G_2)^*$.

$$G_1 = \langle \{a\}, \{S\}, \{S \rightarrow aaS \mid a\}, S \rangle$$
$$G_2 = \langle \{a, b\}, \{S, A\}, \{S \rightarrow aS \mid A, A \rightarrow abA \mid S \mid a\}, S \rangle$$

Megoldás

Először is átnevezzük a nyelvtani jeleket:

$$G'_1 = \langle \{a\}, \{S_1\}, \{S_1 \rightarrow aaS_1 \mid a\}, S_1 \rangle$$
$$G'_2 = \langle \{a, b\}, \{S_2, A_2\}, \{S_2 \rightarrow aS_2 \mid A_2, A_2 \rightarrow abA_2 \mid S_2 \mid a\}, S_2 \rangle$$

Nyelvtani transzformációk I.

Nyelvtanosztályok zárttsági tulajdonságai (3. típus)

Unió:

$$G_{\cup} = \langle \{a, b\}, \{S, S_1, S_2, A_2\}, \{S \rightarrow S_1 \mid S_2, S_1 \rightarrow aaS_1 \mid a, S_2 \rightarrow aS_2 \mid A_2, A_2 \rightarrow abA_2 \mid S_2 \mid a\}, S \rangle$$

Konkatenáció:

$$G_{\text{konk}} = \langle \{a, b\}, \{S_1, S_2, A_2\}, \{S_1 \rightarrow aaS_1 \mid aS_2, S_2 \rightarrow aS_2 \mid A_2, A_2 \rightarrow abA_2 \mid S_2 \mid a\}, S_1 \rangle$$

Lezárás:

$$G_2^* = \langle \{a, b\}, \{S, S_2, A_2\}, \{S \rightarrow \varepsilon \mid S_2, S_2 \rightarrow aS_2 \mid A_2, A_2 \rightarrow abA_2 \mid S_2 \mid a \mid aS_2\}, S \rangle$$

Nulladik típus esetén unió, konkatenáció úgy, mint a 2. típusnál, lezárás úgy, mint 1. típusnál.

3. típusú nyelvtan készítése reguláris kifejezéshez

4. Feladat

Készítsünk olyan 3. típusú G nyelvtant, melyre $L(G) = ((ab \cup a)^* b \cup \varepsilon)^* \cup a!$

Megoldás

Aulról felfelé építkezve kövessük az előző feladat módszerét!

$$S \rightarrow \varepsilon \mid A \mid a$$

$$A \rightarrow B \mid C \mid \varepsilon$$

$$B \rightarrow abC \mid aC \mid abB \mid aB$$

$$C \rightarrow b \mid bA$$

Házi feladat

1. Melyik nyelvet generálja a következő nyelvtan? $T = \{ (,) \}$

a. $S \rightarrow (S) \mid SS \mid \varepsilon$

b. $S \rightarrow XS \mid \varepsilon$ és $X \rightarrow (S)$

c. $S \rightarrow (SS \mid)$

2. Adjunk az $L = \{v \mid v = uu, u \in T^*\}$ ($T = \{a, b\}$) nyelvet generáló nyelvtant a "vv⁻¹" alakú szavak nyelvénél látott módszerre való visszavezetéssel!

3. Készítsünk olyan 3. típusú G nyelvtant, melyre $L(G) = (b(ab \cup a)^* \cup ab)^* \cup aa!$