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:

$$\begin{array}{ll} S \to iS_{i \pmod d} & \forall\, 1 \le i \le M-1 \\ S_i \to jS_{Mi+j \pmod d} & \forall\, 0 \le i,j \le M-1 \\ S_k \to \varepsilon & \end{array}$$

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:

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

$$\begin{array}{ll} S \to tS_t & \forall t \in T \\ S_t \to t'S_{t'} & \forall t,t' \in T,\ tt' \not\in \text{Tiltott} \\ S \to \varepsilon & \\ S_t \to \varepsilon & \forall t \in T \end{array}$$

Házi feladatok megoldása

3. feladat

Adjunk nyelvtant! $T = \{a\}$
$L = \{a^{2^n};\, n \ge 0\}$.

Megoldás:

$$\begin{array}{l} S \to LDaR \\ LD \to \varepsilon \\ R \to \varepsilon \\ Da \to aaD \\ DR \to ER \\ ER \to \varepsilon \\ L \to \varepsilon \\ aE \to Eaa \\ LE \to LD \end{array}$$

Nyelvtanok típusai

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

Nyelvek típusai

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

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

Kiterjesztési tétel

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

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

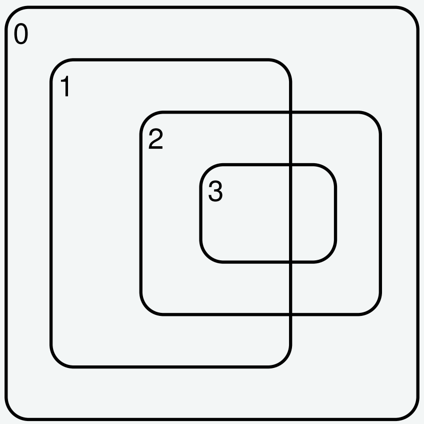
$\mathcal{L}_{REG}$ a reguláris kifejezéssel leírható nyelvek (reguláris nyelvek) nyelvcsaládja.

Kleene tétele

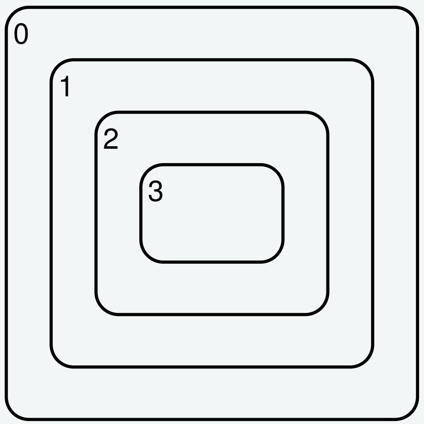
$\mathcal{L}_{REG} = \mathcal{L}_3$.

Chomsky nyelvhierarchia

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



Nyelvtanok



Megszorított nyelvtanok ; nyelvek

Milyen típusúak a következő nyelvtanok?

1. $S \to \varepsilon \mid aSb \mid bSa \mid SS$
   2. típus (nem megszorított)
2. $S \to \varepsilon \mid aSB \mid BSa \mid SS$
   $B \to \varepsilon \mid b$
   2. típus (nem megszorított)
3. $S \to \varepsilon \mid aS \mid aSb \mid bSa \mid Sa \mid SS$
   2. típus (nem megszorított)
4. $S \to tSt \mid \varepsilon \quad \forall t \in T$
   2. típus (nem megszorított)
5. $S \to X_t t \mid X_t t S'$
   $S' \to tY_t \mid tY_t S'$
   $tY_{t'} \to Y_{t'} t$
   $X_t Y_{t'} \to tX_{t'}$
   $X_t \to t \qquad \forall t, t' \in T$
   1. típus (nem megszorított)

Melyik nyelvet generálják a következő nyelvtanok?

1. $S \to \varepsilon \mid aSb \mid bSa \mid SS$
   $L_1 = \{u \in \{a,b\}^* \mid \ell_a(u) = \ell_b(u)\}$
2. $S \to \varepsilon \mid aSB \mid BSa \mid SS$
   $B \to \varepsilon \mid b$
   $L_2 = \{u \in \{a,b\}^* \mid \ell_a(u) \geq \ell_b(u)\}$
3. $S \to \varepsilon \mid aS \mid aSb \mid bSa \mid Sa \mid SS$
   $L_3 = \{u \in \{a,b\}^* \mid \ell_a(u) \geq \ell_b(u)\}$
4. $S \to tSt \mid \varepsilon \quad \forall t \in T$
   $L_4 = \{u \mid u{=}vv^{-1}, v{\in}T^*\}$
5. $S \to X_t t \mid X_t t S'$
   $S' \to tY_t \mid tY_t S'$
   $tY_{t'} \to Y_{t'} t$
   $X_t Y_{t'} \to tX_{t'}$
   $X_t \to t \qquad\qquad \forall t, t' \in T$
   $L_5 = \{v \mid v = uu,\ u \in T^+\}$ (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 $S \xrightarrow{n-1} \alpha \to \beta$. Az indukciós feltevés alapján $\ell_a(\alpha) = \ell_b(\alpha)$. Mivel $\alpha \to \beta$, 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 $S \to SS \xrightarrow{*} u'S \xrightarrow{*} u'u'' = u$. (A többi eset hasonló.)

Megoldás (vázlat):

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

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

Nyelvtani transzformációk I.

Nyelvtanosztályok zártsági tulajdonságai (1. típus)

1. Feladat

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

$$G_1 = \langle \{a,b\}, \{S,A,B\}, \{S \to \varepsilon \mid ASB, AAB \to aaa, Aa \to aA, B \to b\}, S \rangle$$
$$G_2 = \langle \{a\}, \{S,A,B\}, \{S \to ABS \mid A, ABA \to aaaa, BA \to aaa\}, S \rangle$$

Megoldás

Először is átnevezzük a nyelvtani jeleket:

$$G_1' = \langle \{a,b\}, \{S_1,A_1,B_1\}, \{S_1 \to \varepsilon \mid A_1S_1B_1, A_1A_1B_1 \to aaa, A_1a \to aA_1, B_1 \to b\}, S_1 \rangle$$
$$G_2' = \langle \{a\}, \{S_2,A_2,B_2\}, \{S_2 \to A_2B_2S_2 \mid A_2, A_2B_2A_2 \to aaaa, B_2A_2 \to aaa\}, S_2 \rangle$$

Unió:

$$G_{\cup} = \langle \{a, b\}, \{S, S_1, A_1, B_1, S_2, A_2, B_2\}, \{S \to S_1 \mid S_2 \mid \varepsilon, S_1 \to A_1S_1B_1,$$
$$A_1A_1B_1 \to aaa, A_1a \to aA_1, B_1 \to b, S_2 \to A_2B_2S_2 \mid A_2,$$
$$A_2B_2A_2 \to aaaa, B_2A_2 \to aaa\}, S\rangle$$

Konkatenáció:

$$G_{\text{konk}} = \langle \{a, b\}, \{S, S_1, A_1, B_1, S_2, A_2, B_2\}, \{S \to S_1S_2 \mid S_2,$$
$$S_1 \to A_1S_1B_1, A_1A_1B_1 \to aaa, A_1a \to aA_1, B_1 \to b,$$
$$S_2 \to A_2B_2S_2 \mid A_2, A_2B_2A_2 \to aaaa, B_2A_2 \to aaa\}, S\rangle$$

Lezárás:

$$G_1^* = \langle \{a, b\}, \{S_1, A_1, B_1\}, \{S \to \varepsilon \mid S_1 \mid S_1S', S_1 \to \varepsilon \mid A_1S_1B_1,$$
$$A_1A_1B_1 \to aaa, A_1a \to aA_1, B_1 \to b, aS' \to aS_1, bS' \to bS_1,$$
$$aS' \to aS_1S', bS' \to bS_1S'\}, S\rangle$$

2. Feladat

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

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

Megoldás

Először is átnevezzük a nyelvtani jeleket:

$$G_1' = \langle \{a\}, \{S_1\}, \{S_1 \to S_1aa \mid a\}, S_1\rangle$$
$$G_2' = \langle \{a, b\}, \{S_2, A_2\}, \{S_2 \to aS_2A_2 \mid \varepsilon, A_2 \to b \mid \varepsilon\}, S_2\rangle$$

Unió:

$$G_{\cup} = \langle \{a, b\}, \{S, S_1, S_2, A_2\}, \{S \to S_1 \mid S_2, S_1 \to S_1aa \mid a,$$
$$S_2 \to aS_2A_2 \mid \varepsilon, A_2 \to b \mid \varepsilon\}, S\rangle$$

Konkatenáció:

$$G_{\text{konk}} = \langle \{a, b\}, \{S, S_1, S_2, A_2\}, \{S \to S_1S_2, S_1 \to S_1aa \mid a,$$
$$S_2 \to aS_2A_2 \mid \varepsilon, A_2 \to b \mid \varepsilon\}, S\rangle$$

Lezárás:

$$G_2^* = \langle \{a, b\}, \{S, S_2, A_2\}, \{S \to S_2S \mid \varepsilon, S_2 \to aS_2A_2 \mid \varepsilon, A_2 \to b \mid \varepsilon\}, S\rangle$$

3. Feladat

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

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

Megoldás

Először is átnevezzük a nyelvtani jeleket:

$$G_1' = \langle \{a\}, \{S_1\}, \{S_1 \to aaS_1 \mid a\}, S_1\rangle$$
$$G_2' = \langle \{a, b\}, \{S_2, A_2\}, \{S_2 \to aS_2 \mid A_2, A_2 \to abA_2 \mid S_2 \mid a\}, S_2\rangle$$

Nyelvtani transzformációk I.

Nyelvtanosztályok zártsági tulajdonságai (3. típus)

Unió:

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

Konkatenáció:

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

Lezárás:

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

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!
$S \to \varepsilon \mid A \mid a$
$A \to B \mid C \mid \varepsilon$
$B \to abC \mid aC \mid abB \mid aB$
$C \to b \mid bA$

Házi feladat

1. Melyik nyelvet generálja a következő nyelvtan? $T = \{(,)\}$
   - a. $S \to (S) \mid SS \mid \varepsilon$
   - b. $S \to XS \mid \varepsilon$ és $X \to (S)$
   - c. $S \to (SS \mid)$
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 $L(G) = (b(ab \cup a)^* \cup ab)^* \cup aa$!