Formális nyelvek és automaták

Nagy Sára gyakorlatai alapján
Készítette: Nagy Krisztián

3. gyakorlat

Ismétlés:

1. beadható feladat megoldása (saját):

Feladat:
$(L' \cap L'')L''' =^? L'L''' \cap L''L'''$
Ha teljesül, akkor bizonyítás, ha nem, akkor ellenpélda!

Megoldás:
Válasz: Nem teljesül.
Ellenpélda: $L' = \{a, \varepsilon\}; L'' = \{a^2, a^3\}; L''' = \{a, \varepsilon\}$
Bizonyítás:
$(\boldsymbol{a})\ (\boldsymbol{L'} \cap \boldsymbol{L''})\boldsymbol{L'''}\,\boldsymbol{eset}$:
$L' \cap L'' = \{a, \varepsilon\} \cap \{a^2, a^3\} = \emptyset$
$\emptyset L''' = \emptyset$
$(\boldsymbol{b})\ \boldsymbol{L'L'''} \cap \boldsymbol{L''L'''}\,\boldsymbol{eset}$:
$L'L''' = \{a, \varepsilon\}\{a, \varepsilon\} = \{a^2, a, \varepsilon\}$
$L''L''' = \{a^2, a^3\}\{a, \varepsilon\} = \{a^4, a^3, a^2\}$
$L'L''' \cap L''L''' = \{\varepsilon, a, a^2\} \cap \{a^2, a^3, a^4\} = \{a^2\}$
Mivel $\emptyset \neq \{a^2\} \rightarrow (L' \cap L'')L''' \neq L'L''' \cap L''L'''$ ∎

Ismétléshez egy feladat:

Határozzuk meg az alábbi nyelv három hosszú szavait!
$L = (a|ab)^*(b|\varepsilon)$

Megoldás:

$\{a^3, a^2b, ab^2, aba\}$

Megjegyzés: Amennyiben $L = \{a, b\}^*$ nyelvre lenne ez a feladatunk, halmazokra vonatkozóan meg tudjuk határozni, hogy hány darab adott hosszúságú szó szerepel a nyelvben.

Lásd:
$0\ hossz\acute{u}: 2^0 = 1; 1\ hossz\acute{u}: 2^1 = 2; 2\ hossz\acute{u}: 2^2 = 4; 3\ hossz\acute{u}: 2^3 = 8$
Amennyiben a nyelvünkben több megadott betű található, úgy a hatványunk alapja a betűk elem számára „változik” .
Például: $L = \{a, b, c\}^*$

$$0\ hossz\acute{u}: 3^0 = 1; 1\ hossz\acute{u}: 3^1 = 3; 2\ hossz\acute{u}: 3^2 = 9$$

Reguláris nyelv: Lásd 2. gyakorlat ☺

Grammatika → Lehetséges nyelvek körét kibővíti

$G = (T, N, S, P)$
T → szavakat gyártjuk – Terminálisok
N → szavak gyártása közben alkalmazott szimbólumok – Nem terminálisok
S → kezdő szimbólum $S \in N$
P → szabály rendszer, amit a gyártás során alkalmazunk

$L(G) \coloneqq \{u \in T^* \mid S \to_G^* u\}$ - szabályrendszer alapján legyártható a nyelv

$p \to q \in P\ szabály$
Mondat forma:
$p \in (T \cup N)^* N (T \cup N)^*$ Megjegyzés: $(T \cup N)^*$ halmazok lezártja.
Azt írja le, hogy az adott szabály jobb oldala legalább 1 jelet tartalmazzon, ami Nem terminális

$q \in (T \cup N)^*$

Két grammatika ekvivalens, ha ugyan az a nyelv levezethető belőle:
$G' \sim G'': L(G') = L(G'')$

Nyelvek típusai:

0. típusú: Nincs plusz megkötés

1. típusú: $l(p) \leq l(q)$ - hosszúság nem csökkentő szabály/grammatika
KES szabály:

Korlátozott epszilonszabály (KeS): Egy $G = \langle T, N, \mathcal{P}, S \rangle$ nyelvtanra teljesül a korlátozott epszilonszabály, ha nincsenek epszilonszabályai az $S \to \varepsilon$ szabály esetleges kivételével, de ebben az esetben minden $p \to q \in \mathcal{P}$ szabályra $q \in (T \cup N \setminus \{S\})^*$ (azaz semelyik szabály jobboldala sem tartalmazza a kezdőszimbólumot).

Ha S előfordulna jobb oldalt → előfordulna hosszúság csökkentés

2. típusú: $A \to q \quad A \in N, q \in (T \cup N)^*$ - itt megszorítjuk a szabályok bal oldalát, viszont a jobb oldalára nincs egyéb megkötésünk.

ε szabály: $A \to \varepsilon$
Megjegyzés: A 2. típusú nyelvek környezet független nyelvek

3. típusú: $A \to uB \quad A, B \in N, u \in T^*$
$A \to u$
Most a jobb oldalt is megszorítottuk. Láthatjuk, hogy a grammatikánk, ilyenkor balról jobbra gyárt.

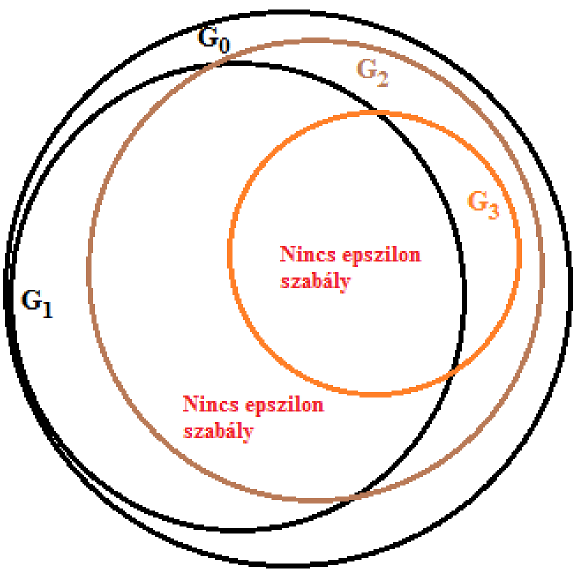
$A \to B\ típusú\ szabályok - lánc\ szabályok$
$L_{reg} = L_3$ - azaz megegyezik a reguláris nyelvekkel

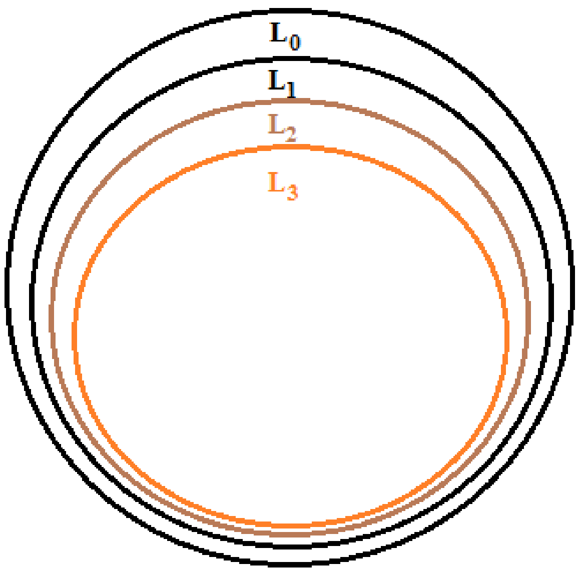
Összefoglalva egy táblázatban:

Típus	Alaptípus szabályai	Megszorított típus szabályai	Normálforma szabályai
0.	nincs további megkötés	$p \to q$, ahol $q \in (T \cup N)^+$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$AB \to A$ $(A, B \in N)$; $BA \to A$ $(A, B \in N)$; +Kuroda NF szabálysémái
1.	$p \to q$, ahol $\ell(p) \le \ell(q)$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$\gamma_1 A \gamma_2 \to \gamma_1 q \gamma_2$, 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	(Kuroda) $A \to a$ $(A \in N, a \in T)$; $A \to BC$ $(A, B, C \in N)$; $AB \to AC$ $(A, B, C \in N)$; $BA \to CA$ $(A, B, C \in N)$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán
2.	$A \to q$, ahol $A \in N, q \in (T \cup N)^*$	$A \to q$, ahol $A \in N, q \in (T \cup N)^+$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	(Chomsky) $A \to a$ $(A \in N, a \in T)$; $A \to BC$ $(A, B, C \in N)$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán
3.	$A \to uB$ vagy $A \to u$, ahol $A, B \in N$, és $u \in T^*$	$A \to aB$ vagy $A \to a$ ahol $A, B \in N, a \in T$; $S \to \varepsilon$, ez esetben S nem szerepel szabály jobboldalán	$A \to \varepsilon$ vagy $A \to aB$, ahol $A, B \in N$ és $a \in T$

(A táblázatban konvencionálisan S a kezdőszimbólum)

Grammatikák közötti kapcsolatok és az általuk gyártott nyelvek közötti kapcsolatok:





Feladat:

Milyen típusú az alábbi grammatika?

G_1:

$S \to XAKBY$
$K \to AKB$
$K \to AB$
$AB \to BaA$
$XB \to aX$
$AY \to Ya$
$aB \to Ba$
$Aa \to aA$
$X \to aa$
$Ya \to aa$

A szabály rendszerek alapján vizsgáljuk meg: (lásd: Nyelvek típusai)
$S \to XAKBY$ 2. típusú grammatikára illik
$K \to AKB$ 2. típusú grammatikára illik
$K \to AB$ 2. típusú grammatikára illik
$AB \to BaA$ 1. típusú grammatikára illik
$XB \to aX$ 1. típusú grammatikára illik
$AY \to Ya$ 1. típusú grammatikára illik
$aB \to Ba$ 1. típusú grammatikára illik
$Aa \to aA$ 1. típusú grammatikára illik
$X \to aa$ 3.típusú grammatikára illik
$Ya \to aa$ 1.típusú grammatikára illik

Ezek alapján ez 1. típusú grammatika.

Feladat:

$L_1 = \{u \in \{a\}^* | \, l(u) \, mod \, 2 = 0\} = (aa)^* \in L_{reg}$
Adjunk rá 3. típusú grammatikát!

G_2:
$S \to aaS$
$S \to \varepsilon$
Ez valóban 3. típusú grammatika, viszont nem 1. típusú [Ha a kezdő szimbólumból levezethető az ε, akkor a kezdőszimbólum nem szerepelhet szabály jobb oldalán]

Alakítsuk át: ε mentesítés [Legyen az új kezdőszimbólum S']
$S' \to \varepsilon | S$
$S \to aaS$
$S \to aa$

$L(G_2) = L(G_3)$, azaz a G_2 és G_3-as grammatika ugyan azt a nyelvet generálja.

Bizonyítási módszerek

$L_2 = \{u \in \{a,b\}^* | l_a(u) = l_b(u)\} = \{\varepsilon, ab, ba, aabb, abab, abba, baab, baba, bbaa, \dots\}$

G_4:

$$\left.\begin{array}{ll} S \to aSb & \text{2.típusú} \\ S \to \varepsilon & \text{3.típusú} \\ aB \to Ba & \text{1.típusú} \\ B \to b & \text{3.típusú} \end{array}\right\} \text{0. típusú lesz}$$

$L(G_4) = L_2$
$L_2 \subseteq L(G_4)$ ✔
$L(G_4) \subseteq L_2$
Vegyünk egy szót, legyártva jó szót kapunk?

Megjegyzés:
$aB \to Ba$ csereszabály előre viszi a B-ket majd a $B \to b$ szabály lecseréli $b - re\ a\ nagy\ B$-t

$S \to aSb$: Amikor születnek a-k, akkor ugyan annyi B is születik.
$aB \to Ba$: csak cserél, de nem módosítja ezeket a születéseket.
$S \to \varepsilon$: nem módosítja a, b-t

Tehát teljesül $L(G_4) \subseteq L_2$ és mivel $L_2 \subseteq L(G_4)$ is teljesült, ezért $L(G_4) = L_2$

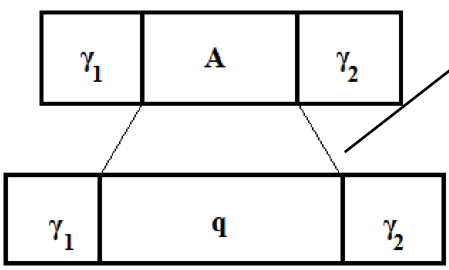
G_5: [Legyen az új kezdőszimbólum S']

$$\left.\begin{array}{ll} S' \to \varepsilon & \text{KES szabály} \\ S' \to S & \text{3. típusú (Megjegyzés: Lánc)} \\ S \to aSB & \text{2.típusú} \\ S \to aB & \text{3.típusú} \\ aB \to Ba & \text{1.típusú} \\ B \to b & \text{3.típusú} \end{array}\right\} \text{1.típusú}$$

G_6:

$$\left.\begin{array}{ll} S \to aSbS & \text{2.típusú} \\ S \to bSaS & \text{2.típusú} \\ S \to \varepsilon & \text{3.típusú} \end{array}\right\} \text{2.típusú}$$

$L_2 = L(G_6)$
$L(G_6) \subseteq L_2 \to$ Legyártunk egy szót a grammatikában, biztos hogy benne van-e a szó?
Ha bejön egy a, akkor egy b is bejön, az ε nem zavarja meg ezt a folyamatot, tehát az összes szót le lehet vezetni.
$L_2 \subseteq L(G_6) \to$ 2.típusú grammatika $\to$ Teljes indukció



$A \to q$
bele konkatenálódik a környezetébe $\to$ ezért lehet teljes indukció

Teljes indukció:

$n = 2k$

$k = 0 \;\rightarrow\; \varepsilon : S \rightarrow \varepsilon$ ✔

$k = 1 \;\rightarrow ab : S \rightarrow aSbS \rightarrow abS \rightarrow ab$ ✔

$ba : S \rightarrow bSaS \rightarrow bSa \rightarrow ba$ ✔

Tegyük fel, hogy $2k$-ra teljesül, nézzük meg, hogy le tudjuk-e vezetni $2(k+1) = 2k+2$-re!

$(\boldsymbol{a})$:

$l(u) = 2k + 2, u \in L_2$

$u = au_1bu_2 \qquad \rightarrow \qquad a\,\overbrace{ab}^{u_1}\,b\,\underbrace{abba}_{u_2}$

párja

$au_1b \in L_2$ ugyan annyi a- és b- vel rendelkező prefixet keresünk

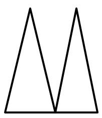
$u_1 \in L_2 \quad u_2 \in L_2$ b-nél vágás van: $(au_1b|u_2)$

$0 \leq l(u_1) \leq 2k$

$0 \leq l(u_2) \leq 2k$ mert a, b-t már mondtuk

Gyártás:

$$S \rightarrow aSbS \xrightarrow[\;]{teljes\ indukció}{}^{*}\ au_1bS \xrightarrow[\;]{teljes\ indukció}{}^{*}\ au_1bu_2$$



$u_1 \quad u_2$

$(\boldsymbol{b})$:

$l(u) = 2k + 2, u \in L_2$

$u = bu_1au_2 \qquad \rightarrow \qquad b\,\overbrace{ba}^{u_1}\,a\,\underbrace{baab}_{u_2}$

párja

$bu_1a \in L_2$ ugyan annyi b- és a- val rendelkező prefixet keresünk

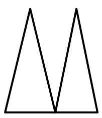
$u_1 \in L_2 \quad u_2 \in L_2$ a-nál vágás van: $(bu_1a|u_2)$

$0 \leq l(u_1) \leq 2k$

$0 \leq l(u_2) \leq 2k$ mert a, b-t már mondtuk

Gyártás:

$$S \rightarrow bSaS \xrightarrow[\;]{teljes\ indukció}{}^{*}\ bu_1aS \xrightarrow[\;]{teljes\ indukció}{}^{*}\ bu_1au_2$$



$u_1 \quad u_2$

Megjegyzés: Addig lehet mutatni, amíg $\boldsymbol{u_1, u_2}$ -ből $\boldsymbol{\varepsilon}$-ok lesznek. → $\boldsymbol{u_1, u_2}$ eltűnik ∎