A számításelmélet alapjai I. (Kilencedik gyakorlat)

Dr. Lázár Katalin Anna

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

2024. április 16.

Tematika

- A CYK algoritmus.
- Bar-Hillel vagy pumpáló lemma.

A CYK algoritmus

Példa 1

Tekintsük a $G = (\{S, A, B, X, Y, Z\}, \{a, b\}, P, S)$ grammatikát, ahol $P = \{S \to AY, Y \to XB, X \to BA, X \to ZA, Z \to BX, A \to a, B \to b\}$! Döntsük el, benne van-e a grammatika által generált nyelvben az *abbaab* szó?

A CYK algoritmus

Példa 1

Megjegyzés

Bármely környezetfüggetlen grammatika és $w \in T^*$ esetében el tudjuk dönteni, hogy $w \in L(G)$ teljesül-e vagy sem.
A CYK (Cocke-Younger-Kasami) algoritmus:

1. Legyen $G = (N, T, P, S)$ Chomsky normálformájú grammatika.
2. Kitöltünk egy háromszög alakú táblázatot, amelyben a sorok az $a_1 \ldots a_n$ szót reprezentálják.
3. A táblázat $x_{i,j}$ eleme azon A nemterminálisokat tartalmazza, amelyekre $A \Longrightarrow^* a_i \ldots a_j$ teljesül.

A CYK algoritmus

Példa 1

Megjegyzés

4. Az első sorban (alulról-felfelé) x_{ii} minden olyan A nemterminálist tartalmaz, amelyre $A \to a_i \in P$ teljesül.
5. A $(j-i+1)$-edik sorban levő x_{ij}-t a következőképpen számítjuk ki. Az x_{ij} minden olyan A nemterminálist tartalmaz, amelyre $A \Longrightarrow^* a_i \dots a_j$ teljesül. Ennek megfelelően minden olyan A nemterminálist tartalmazni fog, amelyre $A \to BC \in P$ fennáll, ahol $B \in x_{ik}$ és $C \in x_{k+1j}$, ahol $i \leq k < j$.
6. $w = a_1 \dots a_n \in L(G)$, akkor és csak akkor, ha $S \in x_{1n}$.

A CYK algoritmus

Példa 1

Megjegyzés

x_{14}			
x_{13}	x_{24}		
x_{12}	x_{23}	x_{34}	
x_{11}	x_{22}	x_{33}	x_{44}
a_1	a_2	a_3	a_4

A CYK algoritmus

Példa 1

A feladat megoldása: $S \in x_{16}$, tehát $abbaab \in L(G)$.

S					
$\emptyset$	Y				
$\emptyset$	X	$\emptyset$			
$\emptyset$	Z	$\emptyset$	$\emptyset$		
$\emptyset$	$\emptyset$	X	$\emptyset$	$\emptyset$	
A	B	B	A	A	B
a	b	b	a	a	b

A CYK algoritmus

Példa 2

Tekintsük a $G = (\{S, A, B, C, D\}, \{a, b, c\}, P, S)$ grammatikát, ahol $P = \{S \to AB, A \to CA, A \to SS, B \to CD, A \to b, D \to a, C \to c, C \to b\}$. Döntsük el, benne van-e a grammatika által generált nyelvben az *abcacb* és a *bbcbba* szó?

A CYK algoritmus

Példa 2

$S \notin x_{16}$, tehát $abcacb \notin L(G)$.

$\emptyset$					
$\emptyset$	$\emptyset$				
$\emptyset$	$\emptyset$	$\emptyset$			
$\emptyset$	S	$\emptyset$	$\emptyset$		
$\emptyset$	$\emptyset$	B	$\emptyset$	A	
D	A, C	C	D	C	A, C
a	b	c	a	c	b

A CYK algoritmus

Példa 2

$S \in x_{16}$, tehát $bbcbba \in L(G)$.

S					
A	S				
A	A	S			
$\emptyset$	A	A	S		
A	$\emptyset$	A	A	B	
A, C	A, C	C	A, C	A, C	D
b	b	c	b	b	a

Bar-Hillel vagy pumpáló lemma

Lemma 1

Minden L környezetfüggetlen nyelvhez meg tudunk adni két természetes számot, p-t és q-t úgy, hogy minden olyan szó L-ben, amely hosszabb, mint p, $uvxyz$ alakú, ahol $|vxy| \leq q$, $vy \neq \varepsilon$, és minden uv^ixy^iz szó is benne van az L nyelvben minden $i \geq 0$ egész számra ($u, v, x, y, z \in T^*$).

Bar-Hillel vagy pumpáló lemma

Példa 3

Bizonyítsuk be, hogy az alábbi nyelvek nem környezetfüggetlenek!

- $L_1 = \{a^n b^m a^n \mid n \geq m\}$.
- $L_2 = \{ww \mid w \in \{a, b\}^*\}$.
- $L_3 = \{a^{n^2} \mid n \geq 1\}$.

Bar-Hillel vagy pumpáló lemma

Példa 3

Legyen p tetszőleges természetes szám, $w = a^p b^p a^p \in L_1$, amelyre $|w| \geq p$. Ekkor $\forall u, v, x, y, z \in \{a, b\}^*$ esetén, amelyre $w = uvxyz$, $|vxy| \leq p$, $|vy| > 0$, ha $|vy|_b \neq 0$, akkor $uv^2xy^2z \notin L_1$, egyébként $(|vy|_b = 0)$ $uxz \notin L_1$.

Bar-Hillel vagy pumpáló lemma

Példa 3

A nyelvre nem teljesül a pumpáló lemma. Legyen p tetszőleges természetes szám, $w = a^p b^p a^p b^p \in L_2$, amelyre $|w| \geq p$. Ekkor $\forall u, v, x, y, z \in \{a, b\}^*$ esetén, amelyre $w = uvxyz$, $|vxy| \leq p$, $|vy| > 0$, $uxz \notin L_2$.

Bar-Hillel vagy pumpáló lemma

Példa 3

A nyelvre nem teljesül a pumpáló lemma. Legyen p tetszőleges természetes szám, $w = a^{p^2} \in L_2$, amelyre $|w| \geq p$. Ekkor $\forall u, v, x, y, z \in \{a\}^*$ esetén, amelyre $w = uvxyz$, $|vxy| \leq p$, $|vy| > 0$, $uv^2xy^2z \notin L_3$, mivel $|w| = p^2 < |uv^2xy^2z| \leq p^2 + p < (p+1)^2 = p^2 + 2p + 1$.

Bar-Hillel vagy pumpáló lemma

Példa 4

Bizonyítsuk be, hogy az $L = \{w \mid w \in \{a, b, c\}^*, |w|_a = |w|_b = |w|_c\}$ nyelv nem környezetfüggetlen!

Bar-Hillel vagy pumpáló lemma

Példa 4

- 1. megoldás (Bar-Hillel vagy pumpáló lemma segítségével:) Legyen p tetszőleges természetes szám, $w = a^p b^p c^p \in L$, amelyre $|w| \geq p$. Ekkor $\forall u, v, x, y, z \in \{a, b, c\}^*$ esetén, amelyre $w = uvxyz$, $|vxy| \leq p$, $|vy| > 0$, $uxz \notin L$, mivel vy $\{a, b, c\}$-ből legalább egy betűt nem tartalmaz, vagyis a nyelv nem teljesíti a Bar-Hillel vagy pumpáló lemma feltételeit, tehát nem környezetfüggetlen.

Bar-Hillel vagy pumpáló lemma

Példa 4

- 2. megoldás (a környezetfüggetlen nyelvek zártak a reguláris nyelvvel való metszetre): Legyen $L_1 = \{a^n b^n c^n \mid n \geq 0\}$. Tudjuk, hogy L_1 nem környezetfüggetlen nyelv. Mivel $L \cap \{a\}^*\{b\}^*\{c\}^* = L_1$ és $\{a\}^*\{b\}^*\{c\}^*$ reguláris, ezért L nem környezetfüggetlen nyelv.

Bar-Hillel vagy pumpáló lemma

Példa 5

Bizonyítsuk be, hogy az

- $L_1 = \{a^i b^{2i} c^j \mid i, j \geq 0\}$ nyelv környezetfüggetlen!
- $L_2 = \{a^i b^j c^{2j} \mid i, j \geq 0\}$ nyelv környezetfüggetlen!
- $L = L_1 \cap L_2$ nyelv nem környezetfüggetlen!

Bar-Hillel vagy pumpáló lemma

Példa 5

- Legyen $G_1 = (\{S, X, C\}, \{a, b, c\}, P_1, S)$, ahol $P_1 = \{S \to XC, X \to aXbb, X \to \varepsilon, C \to cC, C \to \varepsilon\}$. Ekkor $L(G_1) = L_1$.
- Legyen $G_2 = (\{S, A, X\}, \{a, b, c\}, P_2, S)$,ahol $P_2 = \{S \to AX, X \to bXcc, X \to \varepsilon, A \to aA, A \to \varepsilon\}$. Ekkor $L(G_2) = L_2$.
- Az $L = L_1 \cap L_2 = \{a^i b^{2i} c^{4i} \mid i \geq 0\}$ nyelvre nem teljesül a környezetfüggetlen nyelvek pumpáló lemmája. Legyen p tetszőleges természetes szám, $w = a^p b^{2p} c^{4p} \in L$. Ekkor $|w| \geq p$ és $\forall u, v, x, y, z \in \{a, b, c\}^*$ esetén, amelyre $w = uvxyz$, $|vxy| \leq p$, $|vy| > 0$, $uxz \notin L$, mivel a vy szóban legfeljebb kétféle betű szerepelhet $\{a, b, c\}$-ből.