Задачи с доказателство на контекстно-свободни граматики

Ангел Димитриев

Задача 1

Докажете, че езикът $L = \{a^n b^n | n \in \mathbf{N}\}$ е контекстно- свободен.

Решение:

L е контекстно-свободен, ако съществува контекстно-свободна граматика G за езика L. Да разгледаме граматика G:

$$S \to aSb|\epsilon$$

Ще докажем, че граматиката генерира точно думите от L. Т.е че: $L(G) = L$.
1.$L(G) \subseteq L$.

Трябва да покажем, че всички думи от граматиката (съставени само от символи от Σ) принадлежат на езика L. За целта трябва да дефинираме точно какво е поведението на променливата S. Твърдим, че:

$$S \to^* \underset{\triangle}{\{a^m S b^m \ | m \in \mathbf{N}\}} \cup \underset{\circ}{\{a^m b^m | m \in \mathbf{N}\}}$$

(S $\to^*$ w означава, че w е изводимо от S за краен брой стъпки)

Ще докажем твърдението с **индукция по дължината на извода**.
База: За 0 стъки: $S \to^0 S = a^0 S b^0 \triangle$.

Индукционно предположение: Допускаме, че за k стъпки:
$S \to^k a^q S b^q$ или $S \to^k a^q b^q$

Индукционно стъпка: Разглеждаме за k+1 стъпка. Ще приложим всяко едно от правилата:
•Знаем, че $S \to^k a^q S b^q$:
Правило 1: $S \to^k a^q S b^q \ \to a^q a S b b^q = a^{q+1} S b^{q+1} \ \triangle$
Правило 2: $S \to^k a^q S b^q \ \to a^q b^q \ \circ$
Което и правило да приложим за $k + 1$-та стъпка, получаваме дума във вида, който искаме.
•Знаем, че $S \to^k a^q b^q$
Но тук нямаме променливи, така че няма как да приложим $k + 1$ стъпка.

Доказахме, че от S можем да изведем **само** думи от този вид:
$\{a^m S b^m \ | m \in \mathbf{N}\} \cup \{a^m b^m | m \in \mathbf{N}\}$
Но единствените думи, които са от Σ^* са $\{a^m b^m | m \in \mathbf{N}\}$, за които е ясно, че са $\subseteq L$.
Доказахме, че $L(G) \subseteq L$

2.$L \subseteq L(G)$. Трябва да покажем, че всяка дума, която е от езика, може да се изведе от G. Т.е $(\forall i \in \mathbf{N}) S \rightarrow^* a^i b^i$
Ще докажем твърдението с **индукция по дължината на думата**.

База: Най-малката дума от езика е с дължина 0. Това е думата ϵ. Можем да изведем: $S \rightarrow \epsilon$

Индукционно предположение: Допускаме, че можем да изведем **всички думи** с дължина <k (k е произволно естествено число (k>1))
Т.е допускаме, че $(\forall j) j < k : S \rightarrow a^j b^j$

Индукционна стъпка:
Разглеждаме произволна дума с дължина k: $a^t b^t$,където $t + t = k$.
Ще покажем, че можем да я изведем. **Трябва да изразим дадената дума, чрез по-къса дума от същия език + някое правило от граматиката.**
Думата може да разбием така:
$a^t b^t = a a^{t-1} b^{t-1} b$.
Вижда се, че $a^{t-1} b^{t-1}$ е от езика и че $|a^{t-1} b^{t-1}| < k$. Т.е от и.п. можем да я изведем от граматиката. Т.е $S \rightarrow^* a^{t-1} b^{t-1}$
Сега да изведем желаната дума ($a^t b^t$)

$$S \rightarrow aSb \rightarrow^{*\text{и.п.}} a a^{t-1} b^{t-1} b = a^t b^t$$

Показахме, че можем да изведем думата $a^t b^t$.
От тук следва, че $(\forall i \in \mathbf{N}) S \rightarrow^* a^i b^i$.
Т.е $L \subseteq L(G)$.
От тук вече е ясно, че $L = L(G)$. Следователно L е контекстно свободен.

Задача 2

Докажете, че езикът $L = \{a^n b^k | n, k \in \mathbf{N} \wedge n > k\}$ е контекстно-свободен.

Решение:

L е контекстно-свободен, ако съществува контекстно-свободна граматика G за езика L. Да разгледаме граматика G:

$$S \rightarrow aSb | aS | a$$

Ще докажем, че граматиката генерира точно думите от L. Т.е че: $L(G) = L$.

1.$L(G) \subseteq L$. Трябва да покажем, че всички думи от граматиката (съставени само от символи от Σ) принадлежат на езика L. За целта трябва да дефинираме точно какво е поведението на променливата S. Твърдим, че:

$$S \to^* \underset{\triangle}{\{a^{t_1} S b^{t_2} \ | t_1 \geq t_2\}} \cup \underset{\circ}{\{a^{s_1} b^{s_2} | s_1 > s_2\}}$$

(S $\to^*$ w означава, че w е изводимо от S за краен брой стъпки)

Ще докажем твърдението с **индукция по дължината на извода**.
База: За 0 стъки: $S \to^0 S = a^0 S b^0 \triangle$.

Индукционно предположение: Допускаме, че за k стъпки:
$S \to^k a^{q_1} S b^{q_2}$ $(q_1 \geq q_2)$ или $S \to^k a^{q_1} b^{q_2}$ $(q_1 > q_2)$

Индукционно стъпка: Разглеждаме за k+1 стъпка.
•Знаем, че $S \to^k a^{q_1} S b^{q_2}$. Ще приложим всяко едно от правилата:
1. $S \to^k a^{q_1} S b^{q_2} \ \to a^{q_1} a S b b^{q_2} = a^{q_1+1} S b^{q_2+1} \triangle \ (q_1 \geq q_2 \implies q_1 + 1 \geq q_2 + 1))$
2. $S \to^k a^{q_1} S b^{q_2} \ \to a^{q_1} a S b^{q_2} = a^{q_1+1} S b^{q_2} \triangle \ (q_1 \geq q_2 \implies q_1 + 1 \geq q_2))$
3. $S \to^k a^{q_1} S b^{q_2} \ \to a^{q_1} a b^{q_2} = a^{q_1+1} b^{q_2} \circ (q_1 \geq q_2 \implies q_1 + 1 > q_2))$

Което и правило да приложим за $k+1$-та стъпка, получаваме дума във вида, който искаме.

•Знаем, че $S \to^k a^{q_1} b^{q_2}$.
Но тук нямаме променливи, така че няма как да приложим $k+1$ стъпка.

Доказахме, че от S можем да изведем **само** думи от този вид:
$\{a^{t_1} S b^{t_2} \ | t_1 \geq t_2\} \cup \{a^{s_1} b^{s_2} | s_1 > s_2\}$
Но единствените думи, които са от Σ^* са $\{a^{s_1} b^{s_2} | s_1 > s_2\}$, за които е ясно, че са $\subseteq L$.
Доказахме, че $L(G) \subseteq L$

2.$L \subseteq L(G)$. Трябва да покажем, че всяка дума, която е от езика, може да се изведе от G. Т.е $(\forall i, j \in \mathbf{N}) i > j \implies S \to^* a^i b^j$
Ще докажем твърдението с **индукция по дължината на думата**.

База: Най-малката дума от езика е с дължина 1. Това е думата **a**. Можем да изведем: $S \to a$

Индукционно предположение: Допускаме, че можем да изведем **всички думи** с дължина <k (k е произволно естествено число (k>1))
Т.е допускаме, че $(\forall i, j) i > j \wedge i + j < k \implies S \to a^i b^j$

Индукционна стъпка:
Разглеждаме произволна дума с дължина k: $a^t b^r$,където $t > r \wedge t + r = k$.

Ще покажем, че можем да я изведем. **Трябва да изразим дадената дума, чрез по-къса дума от същия език + някое правило от граматиката.**

1 сл. Ако в думата има поне едно b. Тогава тя изглежда така: $a^t b^r$, където $t > r \land t + r = k \land t > 1 \land r \geq 1$. Думата може да разбием така:
$a^t b^r = aa^{t-1}b^{r-1}b$. Думата $a^{t-1}b^{r-1}$ е част от езика, защото:
$t > r \implies t - 1 > r - 1$. Освен това $|a^{t-1}b^{r-1}| < k$, защото $t + r = k$. Т.е от и.п можем да я изведем от граматиката ($S \rightarrow^* a^{t-1}b^{r-1}$). Сега остана да изведем желаната дума:

$$S \rightarrow aSb \rightarrow^{*\text{и.п.}} aa^{t-1}b^{r-1}b = a^t b^r$$

2 сл. Ако в думата няма b-та. Тогава тя изглежда така: $a^k b^0 = a^k$. Можем да я разбием по следния начин: $a^k = aa^{k-1}$. Но думата a^{k-1} е в езика, защото $a^{k-1} = a^{k-1}b^0$, но $k - 1 > 0$, защото $k > 1$. Освен това $|a^{k-1}| < k$, т.е от и.п $S \rightarrow^* a^{k-1}$. Сега остана да изведем думата:

$$S \rightarrow aS \rightarrow^{*\text{и.п.}} aa^{k-1} = a^k$$

Показахме, че можем да изведем думата $a^t b^r (t > r) \land (t + r = k)$. От тук следва, че $(\forall i, j \in \mathbf{N}) i > j \implies S \rightarrow^* a^i b^j$.
Т.е $L \subseteq L(G)$.
От тук вече е ясно, че $L = L(G)$. Следователно L е контекстно-свободен.