Упражнение 10

1 Затвореност относно регулярните операции на контекстно-свободните езици

В духа на това, което направихме при регулярните езици, сега ще покажем някои свойства на затвореност на контекстно-свободните езици относно операции върху езици.

Теорема 1. Контекстно-свободните езици са затворени относно обединение, конкатенация и звезда на Клини.

Нека $G_1 = (V_1, \Sigma, R_1, S_1)$ и $G_2 = (V_2, \Sigma, R_2, S_2)$ са контекстно-свободни граматики и без ограничение на общността да допуснем, че $V_1 \cap V_2 = \varnothing$. Конструкциите са следните.

Обединение. Нека S е символ, който не принадлежи на $V_1 \cup V_2$. Езикът $L(G_1) \cup L(G_2)$ се генерира от граматика

$$G = (V_1 \cup V_2 \cup \{S\}, \Sigma, R_1 \cup R_2 \cup \{S \to S_1, S \to S_2\}, S).$$

Конкатенация. Подобно, $L(G_1) \circ L(G_2)$ се генерира от граматиката

$$G = (V_1 \cup V_2 \cup \{S\}, \Sigma, R_1 \cup R_2 \cup \{S \to S_1 S_2\}, S).$$

Звезда на Клини. $L(G_1)^*$ се генерира от

$$G = (V_1 \cup \{S\}, \Sigma, R_1 \cup \{S \to \epsilon, S \to SS_1\}, S).$$

Теорема 1 влече вече доказания факт, че класът на регулярните езици се съдържа в класа на контекстно-свободните такива.

2 Задачи

Задача 1. Използвайте затвореността относно обединение, за да покажете, че следните езици са контекстно-свободни.
(а) $\{a^m b^n \mid m \neq n\}$
(б) $\{a, b\}^* \setminus \{a^n b^n \mid n \in \mathbb{N}\}$
(в) $\{a^m b^n c^p d^q \mid n = q$, или $m \leq p$ или $m + n = p + q\}$

(г) $\{w \in \{a,b\}^* \mid w = w^R\}$

Задача 2. Покажете, че езикът $L = \{a^n b^{n+m} a^n \mid n, m \in \mathbb{N}\}$ е контекстно свободен, използвайки затвореността относно конкатенация.

Задача 3. Покажете, че езикът $L = \{uw^{rev}v \mid w = vu\}$ е контекстно свободен, използвайки затвореността относно конкатенация.

Задача 4. Покажете, че е контекстно-свободен следният език

$$L = \{a^{n_1}\#a^{n_1+n_2}\#a^{n_2+n_3}\#...\#a^{n_{k-1}+n_k}\#a^{n_k} \mid k, n_1, ..., n_k \in \mathbb{N}\}$$

3 Решения

Задача 1.(а) $\{a^m b^n \mid n < m\} \cup \{a^m b^n \mid n > m\}$;
(б) Можем да представим този език като обединението на следните два езика:
(1) $\{a,b\}^* \setminus \mathscr{L}(a^*b^*)$ (думи, в които има поне едно b преди някое a)
(2) $\{a^n b^m \mid n \neq m\}$
За (1) имаме следната граматика:

$S \to AbAaA$
$A \to aA \mid bA \mid \epsilon$.

(в) $\{a^m b^n c^p d^q \mid n = q\} \cup \{a^m b^n c^p d^q \mid m \leq p\} \cup \{a^m b^n c^p d^q \mid m + n = p + q\}$.
Граматиките са съответно следните:

(1) $S \to aS \mid A$
$A \to bAd \mid B$
$B \to cB \mid \epsilon$;

(2) $S \to Sd \mid A$
$A \to aAc \mid Ac \mid B$
$B \to bB \mid \epsilon$;

(3) правена на предишното упражнение.
(г) $\{ww^{rev} \mid w \in \Sigma^*\} \cup \{waw^{rev} \mid w \in \Sigma^*\} \cup \{wbw^{rev} \mid w \in \Sigma^*\}$

Задача 2. $L = \{a^n b^n \mid n \in \mathbb{N}\} \circ \{b^n a^n \mid n \in \mathbb{N}\}$.

Задача 3. $L = \{ww^{rev} \mid w \in \Sigma^*\}^2$.

Задача 4. $L = \{a^n\#a^n \mid n \in \mathbb{N}\}^*$.

4 Коректност на контекстно-свободни граматики

В тази секция ще покажем как се доказва формално коректността на една контекстно-свободна граматика.

Пример 1. Ще докажем коректността на следната граматика G, генерираща езика $L = \{a^n b^n \mid n \in \mathbb{N}\}$.

$$G : \boxed{S \to aSb \mid \epsilon}$$

Твърдението, което се опитваме да докажем е следното:

$$(\forall w \in \Sigma^*)[w \in L(G) \iff w \in L].$$

За целите на доказателстово му, първо ще докажем следната помощна Лема за релацията $\Rightarrow^*$ в контекста на дадената граматика.

Лема 1. За всяка дума $w \in (\Sigma \cup V)^*$, ако $S \Rightarrow^* w$, то w е в един от следните два вида:
(1) $w = a^n S b^n$, за някое $n \in \mathbb{N}$.
(2) $w = a^n b^n$, за някое $n \in \mathbb{N}$.

Доказателство: Еквивалентно, искаме да докажем, че

$$(\forall n \in \mathbb{N})[(\forall w \in (\Sigma \cup V))^*[S \Rightarrow^n w \implies w \text{ е от вид (1) или (2)}]].$$

Това естествено ще сторим с индукция относно n.

База: Ако $n = 0$ и w е такава поредица от терминали и нетерминали, че $S \Rightarrow^0 w$, то съществува извод с дължина 0 на w от S. С други думи $S = w$. Тогава $w = a^0 S b^0$, тоест w е от вид (1).

Стъпка: Ако $n > 0$ и w е такава поредица от терминали и нетерминали, че $S \Rightarrow^n w$, то съществува извод с дължина n на w от S. Да фиксираме един такъв извод

$$S = w_0 \Rightarrow w_1 \Rightarrow ... \Rightarrow w_{n-1} \Rightarrow w_n = w.$$

От този извод можем да извлечем, че $S \Rightarrow^{n-1} w_{n-1}$. Съгласно И.П. това означава, че w_{n-1} е от вид (1) или (2). Но $w_{n-1} \Rightarrow w_n$ и значи няма как w_{n-1} да е от вид (1) (тя трябва да съдържа нетерминали, съгласно дефиницията на $\Rightarrow$). Тогава $w_{n-1} = a^k S b^k$ за някое $k \in \mathbb{N}$. Сега, щом $w_{n-1} \Rightarrow w_n$ и в w_{n-1} има един единствен нетерминал, S, то имаме, че $w_n = a^k a S b b^k$ или $w_n = a^k \epsilon b^k$. Тоест $w_n = a^{k+1} S b^{k+1}$ или $w_n = a^k b^k$. Следователно w_n е от вид (1) или (2). Остана само да си припомним, че $w_n = w$.

Сега преминаваме към доказателството на същинското твърдение. Нека $w \in \Sigma^*$.

($\Rightarrow$) Нека $w \in L(G)$. Тогава $S \Rightarrow^* w$ и $w \in \Sigma^*$. Съгласно **Лема 1** това означава, че $w = a^n b^n$ за някое $n \in \mathbb{N}$. Значи $w \in L$.
($\Leftarrow$) Обратно с индукция относно n ще покажем, че

$$(\forall n \in \mathbb{N})[S \Rightarrow^{n+1} a^n b^n],$$

което е еквивалентно на обратната посока на твърдението.

База: $n = 0$. Имаме извода $S \Rightarrow \epsilon = a^0 b^0$. Значи $S \Rightarrow^1 a^0 b^0$.

Стъпка: Да допуснем, че за някое $n \in \mathbb{N}$ $S \Rightarrow^{n+1} a^n b^n$. Приемаме за очевидно, че това означава, че $aSb \Rightarrow^{n+1} aa^n b^n b = a^{n+1} b^{n+1}$. От друга страна $S \Rightarrow aSb$ и значи общо имаме "извода"

$$S \Rightarrow aSb \Rightarrow^{n+1} a^{n+1} b^{n+1}.$$

Тоест $S \Rightarrow^{n+2} a^{n+1} b^{n+1}$, което искахме да покажем.

5 Задачи

Задача 1. Докажете формално коректностите на граматиките, предложени на последното упражнение за следните езици.
(а) $\{ww^{rev} \mid w \in \{a, b\}^*\}$
(б) $\{a^m b^n \mid m \geq n\}$
(в) $\{w \in \{a, b\}^* \mid$ броят на a-тата в w е равен на два пъти броя на a-тата$\}$