

1. Podaj gramatykę bezkontekstową dla języka

$$L = \{ a^p b^q c^r d^s : p + q = r + s, \quad p, q, r, s \geq 0 \}.$$

WSKAZÓWKA: zadanie rozbij na 2 przypadki: $p \leq s$ czyli $s = p + k$ i $p > s$ czyli $p = s + k$.

2. Przekształć gramatykę do postaci normalnej Chomsky'ego

$$\begin{array}{ll} S \rightarrow P \mid Q & X \rightarrow aXb \mid \lambda \\ P \rightarrow aPc \mid X & Y \rightarrow bYc \mid \lambda \\ Q \rightarrow XY & \end{array}$$

Jest to nb. gramatyka opisująca język z jednego z zadań domowych, $L = \{ a^k b^m c^{|k-m|} : k, m \geq 0 \}$.

3. Pokaż, że $L = \{ a^p b^{pq} c^q : p, q > 0 \}$ nie jest językiem bezkontekstowym.

4. Czy język $L = \{ \omega \in \{a, b, c\}^* : |\omega| = 3|\omega|_a \}$ jest regularny, liniowy, bezkontekstowy nieliniowy, czy też nie jest nawet bezkontekstowy? Odpowiedź uzasadnij przez zastosowanie odpowiedniego lematu o pompowaniu lub przez podanie gramatyki właściwej klasy.

5. Podaj gramatykę kontekstową lub graf maszyny Turinga dla języka $L = \{ \omega \in \{a, b, c\}^* : |\omega|_a = |\omega|_b = |\omega|_c \}$.

Rozwiązania

- Jeśli $p \leq s$, to $s = p + k$ dla $k \geq 0$. Wtedy zgodnie z warunkiem $p + q = r + s = r + p + k$ mamy $q = k + r$. Słowa L mają więc postać $\omega = a^p b^{k+r} c^r d^{k+p}$. Wyprowadzimy je stosując kolejno produkcje
 - $S \rightarrow aSd$ do wygenerowania p liter a i d , a następnie $S \rightarrow X$ by bezpowrotnie przejść do dalszej części wyprowadzenia
 - $X \rightarrow bXd$ do utworzenia k liter b i d , po czym $X \rightarrow Z$, by przejść do generowania r liter b i c
 - $Z \rightarrow bZc \mid \lambda$ użyta r razy kończy wyprowadzenie słowa ω

Z drugiej strony, jeśli $p > s$, napiszemy $p = s + k$ i podobną drogą otrzymamy $p + q = s + k + q = r + s$, a więc $r = q + k$. Budowa produkcji wyprowadzających słowa $\omega = a^{s+k} b^q c^{q+k} d^s$ jest analogiczna i wykorzystuje nowy nieterminal Y by nie mieszać dwóch części gramatyki. Ostatecznie

$$S \rightarrow aSd \mid X \mid Y, \quad X \rightarrow bXd \mid Z, \quad Y \rightarrow aYc \mid Z, \quad Z \rightarrow bZc \mid \lambda.$$

- Wszystkie symbole nieterminalne gramatyki mogą wyprowadzić napis pusty. Stąd λ -wolna postać gramatyki to

$$\begin{aligned} S &\rightarrow P \mid Q \mid \lambda & X &\rightarrow aXb \mid ab \\ P &\rightarrow aPc \mid ac \mid X & Y &\rightarrow bYc \mid bc \\ Q &\rightarrow XY \mid X \mid Y \end{aligned}$$

Eliminacja produkcji jednostkowych prowadzi z kolei do gramatyki

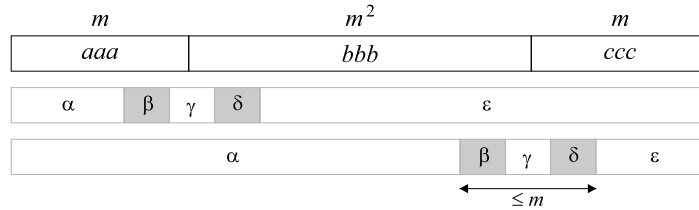
$$\begin{aligned} S &\rightarrow aPc \mid ac \mid aXb \mid ab \mid XY \mid bYc \mid bc \mid \lambda & X &\rightarrow aXb \mid ab \\ P &\rightarrow aPc \mid ac \mid aXb \mid ab & Y &\rightarrow bYc \mid bc \\ Q &\rightarrow XY \mid aXb \mid ab \mid bYc \mid bc \end{aligned}$$

Na koniec dodajemy produkcje $A \rightarrow a$, $B \rightarrow b$ i $C \rightarrow c$ by zastąpić a, b, c przez A, B, C i przekształcamy 3-literowe prawe strony produkcji do dwuliterowych w standardowy sposób.

- Założmy, że L jest bezkontekstowy i niech m będzie stałą z lematu o pompowaniu. Weźmy słowo $\omega = a^m b^{m^2} c^m$. Jeśli w podziale $\omega = \alpha\beta\gamma\delta\varepsilon$ obydwa fragmenty podlegające pompowaniu β i δ zbudowane są wyłącznie z a, b lub c , pompowanie w oczywisty sposób naruszy relację między liczbą liter b oraz a i c . Jedynie przypadki $\beta = a^r$ i $\delta = b^s$ oraz symetrycznie $\beta = b^s$ i $\delta = c^r$, $r, s > 0$, mogłyby prowadzić do pompowania zachowującego strukturę słów języka. Musi przy tym zachodzić $r + s \leq m$. W pierwszym przypadku pompowanie produkuje słowa

$$\omega_k = a^{m+kr} b^{m^2+ks} c^m, \quad k = 1, 2, \dots$$

dla których powinna być spełniona relacja $m^2 + ks = (m + kr)m$ jeśli ω_k miałyby należeć do L . To jest jednak możliwe jedynie gdy $s = rm$, co w połączeniu z $s \leq m - r$ prowadzi natychmiast do sprzeczności $rm \leq m - r$, czyli $r(m + 1) \leq m$. Drugi przypadek jest analogiczny.



- Język L nie jest liniowy (więc tym bardziej nie jest regularny). Weźmy np. słowo $\omega = a^m b^{2m} c^{2m} a^m \in L$, gdzie m jest stałą, o której mowa w lemacie o pompowaniu dla języków liniowych. Dziąc ω na części mamy $\omega = \alpha\beta\gamma\delta\eta$, przy czym $|\alpha\beta\delta\eta| \leq m$, a więc części podlegające pompowaniu β i δ muszą składać się wyłącznie z liter a . Stąd pompowanie produkuje słowa ω' spoza języka L , dla których $|\omega'| < 3|\omega'|_a$. L jest jednak nieliniowym językiem bezkontekstowym, generowanym przykładowo przez gramatykę:

$$S \rightarrow SaSXXS \mid SXSaXS \mid SXSXSaS \mid \lambda, \quad X \rightarrow b \mid c.$$

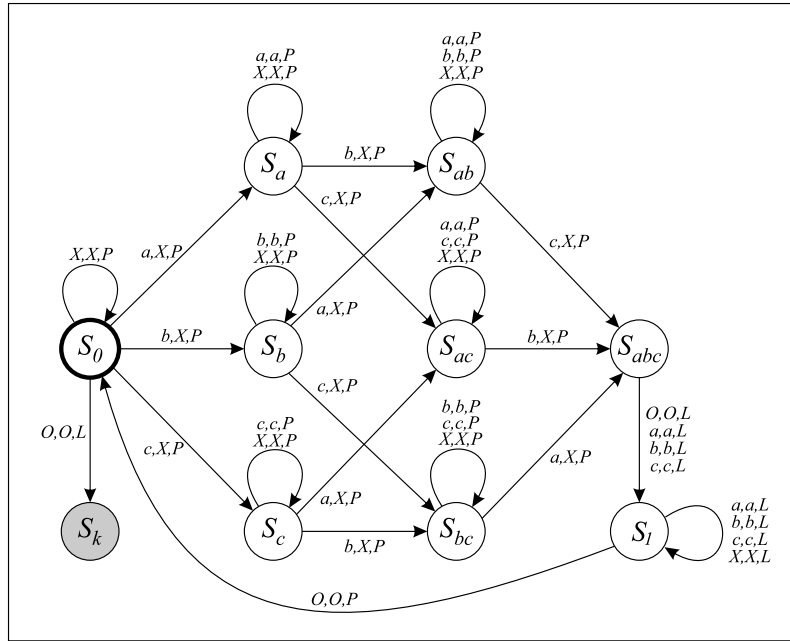
Warunek $|\omega| = 3|\omega|_a$ jest równoważny $2|\omega|_a = |\omega|_b + |\omega|_c$, stąd powyższa postać gramatyki.

5. Gramatyka $G(L)$ zawiera następujące produkcje (nieskracające, jest więc kontekstowa):

$$\begin{array}{lll}
 S & \rightarrow & ABCS \mid \lambda \\
 AB & \rightarrow & BA \\
 BA & \rightarrow & AB \\
 A & \rightarrow & a \\
 AC & \rightarrow & CA \\
 CA & \rightarrow & AC \\
 B & \rightarrow & b \\
 BC & \rightarrow & CB \\
 CB & \rightarrow & BC \\
 C & \rightarrow & c
 \end{array}$$

S generuje napisy $ABCABC \dots ABC$, pozostałe produkcje pozwalają dowolnie zmienić kolejność A, B i C , by w końcu zastąpić je terminalami a, b, c .

Maszyna Turinga akceptująca L przegląda taśmę od lewej do prawej eliminując na niej litery terminalne trójkami, po jednym a, b i c , zastępując je symbolami X . Po eliminacji trójki liter, przesuwa głowicę w lewo do początku taśmy i powtarza cykl. Jeśli na taśmie znajdują się już tylko symbole X , przechodzi do stanu akceptującego, natomiast jeśli wciąż pozostają symbole terminalne "nie do pary", maszyna zatrzyma się w jednym ze swoich stanów nieakceptujących.



UWAGA: L nie jest językiem bezkontekstowym. Gdyby był, jego przekrój z językiem regularnym $L(a^*b^*c^*)$ musiałby także być bezkontekstowy, a oczywiście nie jest.