Zadanie 56 z “Około dwustu łatwych zadań z języków formalnych i złożoności obliczeniowej (i jedno czy dwa trudne) 2020 edition”

Łukasz Klasiński

28 marca 2020

Zadanie 56.

Wiadomo z jednego z poprzednich zadań, że język $L \subseteq \Sigma^*$ jest regularny, to również język

$$L/2 = \{w \in \Sigma^* : \exists v \in \Sigma^* \; |w| = |v| \wedge wv \in L\}$$

jest regularny.

Pokaż, że podobna implikacja nie zachodzi dla języków bezkontekstowych. To znaczy istnieje taki CFL L, dla którego $L/2$ nie jest CFL.

Rozwiązanie.

Najpierw musimy znaleźć odpowiedni L. Jego budowa musi być taka, aby łatwo było “wypompować” brak bezkontekstowości odpowiadającemu mu $L/2$.

Niech takim językiem będzie:

$$L = \{a^n b^m c^m d^{3n} \mid m, n \in \mathbb{N}\}$$

Możemy do niego bardzo łatwo stworzyć akceptujący go język CFG, co wykazuje bezkontekstowość:

$$T = \{a, b, c\}$$

$$N = \{S, B\}$$

$$\Pi = S \to aSddd \mid B, \;\; B \to bBc \mid \varepsilon$$

Ponieważ do udowodnienia tego, że w takim przypadku $L/2$ nie będzie CFL skorzystamy z twierdzenia o pompowaniu dla języków bezkontekstowych (którego nie zdążyliśmy pokazać na wykładzie), przypomnijmy jego treść:

Lemat o pompowaniu dla języków bezkontekstowych

Dla każdego języka bezkontekstowego L istnieje taka stała n, że dla każdego słowa w należącego do tego języka o długości co najmniej n, możemy podzielić to słowo na $uvtxy$ w taki sposób, że:

- $|vx| > 0$
- $|vtx| \leq n$
- $\forall k \in \mathbb{N} \; uv^k tx^k y \in L$, w szczególności $uty \in L$

D-d, że dla naszego L, $L/2$ nie jest CFG

Załóżmy nie wprost, że $L/2$ jest CFG. Wtedy istnieje dla niego stała z lematu o pompowaniu - n.

Rozważmy:

$$w = a^n b^n c^n$$

Oczywiście $w \in L/2$, ponieważ z definicji $L/2$ mamy: $|w| = |v| \; : v = d^{3n} \, , w \wedge wv \in L$

Podzielmy nasze słowo w na $uvwxy$ i sprawdźmy przypadki:

1. $vtx = a^k : k \leq n$, wtedy otrzymujemy słowo: $w' = a^{n+k}b^nc^n$.

   Załóżmy, że $w' \in L/2$. Wtedy z definicji $L/2$, musi istnieć słowo $w'v \in L$ takie, że $|w'| = |v|$. Jedyne słowo $w'' \in L$ które to spełnia, jest równe:

   $$w'' = a^{n+k}b^nc^nd^{3(n+k)} = w'v$$

   Ale wtedy $|v| = 3n + 3k$, a $|w'| = 3n + k$, zatem dla dowolnego $k > 0$ otrzymujemy sprzeczność

2. $vtx = b^k$ - otrzymujemy słowo: $w' = a^nb^{n+k}c^n$, ale $|b^{n+k}| = |c^n|$, bo ilość c w słowie musi być równa ilości b, czyli dla $k \neq 0$ - w' nie należy do $L/2$

3. $vtx = c^k$ - wtedy idąc tym samym rozumowaniem co w (2), nie może należeć do $L/2$

4. $vtx = a^kb^j : k + j \leq n$ - wtedy $j = 0$, bo inaczej będzie więcej b niż c, ale wtedy otrzymamy $vtx = a^k$, czyli przypadek (1)

5. $vtx = b^kc^j$ - mamy podprzypadki:

   - $k > j$ oraz $k < j$ - wtedy ilość b i c jest różna, zatem nie należy do $L/2$
   - $k = j$ - otrzymamy słowo: $w' = a^nb^{n+k}c^{n+k}$.

     Załóżmy, że $w' \in L/2$. Wtedy podobnie jak w (1) musi istnieć słowo

     $$w'' = a^nb^{n+k}c^{n+k}d^{3n} = w'v$$

     Ale $|v| = 3n$, a $|w'| = 3n + 2k$, zatem dla dowolnego $k > 0$ otrzymujemy sprzeczność