JFiZO 2020

Wojciech Pawlik

26 marca 2020

Zadanie 74. Pokaż, że istnieje konfluentny język bezkontekstowy, który nie jest jednostajnie konfluentny.

Definicja 1: Język $A \subseteq \Sigma^*$ jest *konfluenty* jeśli:

$$\forall w, v \in \Sigma^* \ \exists x \in \Sigma^* \ \forall y \in \Sigma^* \ (wxy \in A \Leftrightarrow vxy \in A)$$

Definicja 2: Język $A \subseteq \Sigma^*$ jest *jednostajnie konfluenty* jeśli istnieje taka stała $c \in \mathbb{N}$, że:

$$\forall w, v \in \Sigma^* \ \exists x \in \Sigma^* \ (|x| \leq c \land \forall y \in \Sigma^* \ (wxy \in A \Leftrightarrow vxy \in A))$$

Rozwiązanie Niech $A = \{0^n(0+1)^m : n \geq m\} \subseteq \{0,1\}^*$.
Równoważnie: A jest zbiorem słów złożonych z zer i jedynek, w których liczba zer wiodących jest nie mniejsza niż połowa długości słowa. Pokażemy, że A jest konfluentny, bezkontekstowy, ale nie jest jednostajnie konfluentny.

A **jest konfluentny** Weźmy dowolne słowa $w_1, w_2 \in \{0,1\}^*$. Niech $n_i = |w_i|$, natomiast k_i oznaczać będzie liczbę zer wiodących w słowie w_i. Z definicji języka zachodzi $k_i \geq \frac{n_i}{2} \Leftrightarrow w_i \in A$. Rozważmy dwa przypadki:

- $k_1 < \frac{n_1}{2} \ \land \ k_2 \leq \frac{n_2}{2}$: Oznacza to, że słowa w_1 i w_2 nie należą do języka A. Weźmy $x = 1$. Dla dowolnego słowa $y \in \{0,1\}^*$ zachodzi $w_1xy \notin A$ oraz $w_2xy \notin A$, ponieważ takie konkatenacje nie zmieniają długości prefiksów zer wiodących, natomiast zwiększają długości słów.
- $k_1 \geq \frac{n_1}{2} \ \lor \ k_2 \geq \frac{n_2}{2}$: Oznacza to, że co najmniej jedno ze słów w_1 lub w_2 należy do języka A. Niech $m = max(k_1 - \lfloor\frac{n_1}{2}\rfloor, k_2 - \lfloor\frac{n_2}{2}\rfloor)$. Niech $x = 1^{m+1}$. Wtedy słowa w_1x i w_2x nie należą do języka, co oznacza że dla dowolnego $y \in \{0,1\}^*$ słowa w_1xy i w_2xy również nie należą do języka.

A jest bezkontekstowy Pokażemy gramatykę $G = \langle T, N, S, \pi \rangle$ taką, że $L_G = A$. Niech:

$$T = \{0, 1\}$$
$$N = \{S, P\}$$
$$\pi : S \to \epsilon \mid 0SP,\ P \to \epsilon \mid 0 \mid 1$$

- $A \subseteq L_G$: Weźmy dowolne słowo $w \in A$. Pokażemy jak wyprowadzić to słowo w gramatyce G. Za pomocą produkcji $S \to \epsilon \mid 0SP$ możemy wyprowadzić ciąg $0^n P^n$ dla dowolnego n naturalnego. Niech m będzie liczbą zer wiodących w słowie w. Najpierw wyprowadzimy słowo $0^m P^m$. Teraz pozostał nam do wyprowadzenia sufiks słowa w o długości co najwyżej m. Wykonujemy jeszcze m produkcji typu $P \to \epsilon \mid 0 \mid 1$, które pozwalają nam na odpowiednie przekształcenie ciągu symboli nieterminalnych P^m na symbole tworzące sufiks słowa w.
- $L_G \subseteq A$: Niech $A' = \{0^n(S + \epsilon)(0 + 1 + P)^m : n \geq m\}$. Niech $\underline{L}_G$ będzie zbiorem ciągów symboli terminalnych i nieterminalnych produkowanych przez gramatykę G. Ponieważ $L_G = \underline{L}_G \cap \{0, 1\}^*$ oraz $A = A' \cap \{0, 1\}^*$ wystarczy pokazać, że zachodzi $\underline{L}_G \subseteq A'$. Niech $p(w)$ oznacza liczbę zer wiodących w słowie w. Niech $s(w)$ oznacza długość sufiksu słowa w rozpoczynającego się od 1 lub P. Niech $c(w)$ oznacza, że w słowie w znajduje się co najwyżej jeden symbol S, jest poprzedzony samymi zerami, a prefiks zakończony symbolem S jest dłuższy niż połowa długości w[1]. Wtedy $A' = \{w \in \{0, 1, S, P\}^* : p(w) \geq s(w) \land c(w)\}$.
  Chcemy pokazać, że dla każdego $i \in \mathbb{N}$ zachodzi:
  $$\Phi(i) :\ S \to_\pi^i w \Rightarrow p(w) \geq s(w) \land c(w)$$
  1. $i = 0$: Wtedy $S \to_\pi^0 = S$ i warunek $p(S) \geq s(S) \land c(S)$ jest spełniony.
  2. Załóżmy, że $\Phi(i)$ zachodzi dla pewnego i naturalnego. Pokażemy, że $\Phi(i + 1)$ również zachodzi. Weźmy dowolne wyprowadzenie o długości $i+1$. Do długości i tworzyliśmy tylko słowa spełniające hipotezę indukcyjną. Popatrzmy na ostatnią produkcję:
     - $S \to \epsilon$: Taka produkcja nie zmienia wartości długości sufiksu. Może jedynie wydłużyć prefiks zer.
     - $S \to 0ST$: Skoro w słowie mieliśmy symbol S to musiał się on znajdować nie wcześniej niż w połowie tego słowa. Po tej operacji nadal tak jest, zatem sufiks jest co najwyżej tak samo długi jak prefiks złożony z zer. Warunki pozostają spełnione.
     - $T \to \epsilon \mid 0 \mid 1$: Tymi produkcjami możemy w zależności od umiejscowienia symbolu T wydłużyć prefiks, skrócić lub nie zmienić rozmiaru sufiksu. Nierówność pozostaje spełniona.

     Zatem $\Phi(i + 1)$ zachodzi.

  Własności Φ zachodzi dla dowolnego $i \in \mathbb{N}$ co implikuje $L_G \subseteq A$.

[1] Inaczej: jedyny symbol S może znajdować się nie wcześniej niż w połowie słowa w

A **nie jest jednostajnie konfluentny** Załóżmy nie wprost, że A jest jednostajnie konfluentny. Niech c będzie stałą z definicji jednostajnej konfluentności. Weźmy słowa $w_1 = 1$ i $w_2 = 0^c$. Zachodzi $w_1 \notin A$ oraz $w_2 \in A$. Na pewno każde słowo, którego prefiksem jest w_1 również nie należy do tego języka. Zatem $\forall x, y \in \{0,1\}^*\ w_1xy \notin A$. Skoro język ma być jednostajnie konfluentny, to aby zachodziła równoważność z definicji musimy znaleźć słowo x takie, że $|x| \leq c\ \wedge\ \forall y \in \{0,1\}^*\ w_2xy \notin A$. Najkrótsze słowo mające na początku prefiks zer o długości c, ale nie należące do języka A musi mieć długość większą niż $2c$. Ale $|w_2| = c$ zatem $|x| > c$, co prowadzi do sprzeczności.