Zadanie 3.2

Jarosław Socha

26 listopada 2023

1 Treść zadania

Zadanie 2
Niech G będzie gramatyką

$$S \to aS|aSbS|\varepsilon.$$

Udowodnić, że

$$L(G) = \{x : \text{każdy przedrostek } x \text{ ma co najmniej tyle symboli } a, \text{ co symboli } b\}.$$

Niech $p_k(x)$ to k-elementowy przedrostek słowa x. Język możemy równoważnie zapisać jako

$$L(G) = \{x : (\forall k \in [0, |x|]) \quad |p_k(x)|_a \geqslant |p_k(x)|_b\}$$

2 Teoria

Aby udowodnić, że gramatyka G generuje język L musimy pokazać dwie rzeczy:

- każde słowo generowane przez gramatykę G należy do języka L
- każde słowo z języka L da się wyprowadzić z gramatyki G

3 Dowód

3.1 Każde słowo generowane przez gramatykę G należy do języka L - dowód indukcyjny

Dowód indukcyjny po długości wyprowadzenia. Dla długości wyprowadzenia 1 możemy jedynie otrzymać ϵ, a $|p_k(\epsilon)|_a = 0 \geqslant 0 = |p_k(\epsilon)|_b\}$, więc słowo należy do języka. Załóżmy indukcyjnie, że słowo generowane przez wyprowadzenie długości co najwyżej n należy do języka. Teraz udowodnimy, że dla każdego możliwego kroku wyprowadzenia, niezależnie od tego, w którym momencie wyprowadzania go zastosujemy, nowe słowo o długości wyprowadzenia $n + 1$ należy do języka (rozpatrujemy tylko kroki, które powiększą długość wyprowadzenia).

- $S \to aS$
  Jeśli słowo, które powiększaliśmy to $w = w_1w_2$, to nowe słowo $w' = w_1aw_2$. Na mocy założenia indukcyjnego, w_1 i w_2 spełniają warunek należenia do języka. Weźmy dowolne $k \in [0, n+1]$.
  - Jeśli $k \leqslant |w_1|$ to $p_k(w)$ się nie zmienił.
  - Dla $|w_1| < k$ oba przedrostki zawierają w_1, więc weźmy $j = k - |w_1|$. Otrzymujemy:
    $$p_j(aw_2) = ap_{j-1}(w_2)$$
    Czyli liczba liter a w przedrostku zwiększyła się o 1, lub się nie zmieniła.
  - Dla $k = n+1$, liczba liter a w przedrostku zwiększyła się o 1.

  W każdym przypadku liczba liter a nie zmalała, a liczba liter b nie wzrosła, zatem słowo należy do języka.
- $S \to aSbS$
  Dla słowa $w = w_1w_2w_3$, nowe słowo to $w' = w_1aw_2bw_3$. Na mocy założenia indukcyjnego, w_1, w_2 i w_3 spełniają warunek należenia do języka. Weźmy dowolne $k \in [0, n+1]$.
  - dla $k \leqslant |w_1| + |w_2|$, dowód przebiega analogicznie do powyższego
  - dla $k > |w_1| + |w_2|$ podobnie do powyższego dowodu, liczba liter a się nie zmieni, za to liczba liter b będzie stała lub zwiększy się o 1. Jako że jednak dla tych k liczba liter a już zwiększyła się o 1, to liczba liter a wciąż jest nie mniejsza niż liczba liter b

  Podobnie jak powyżej, dla każdego każdego przedrostka słowo spełnia warunek $|p_k(x)|_a \geqslant |p_k(x)|_b$, więc należy ono do języka.

3.2 Każde słowo z języka L da się wyprowadzić z gramatyki G - algorytm wyprowadzenia

Zamieńmy słowo na równoważny ciąg liczb według zasady:

1. $c = 0$
2. Wypisz c
3. Przeczytaj literę
4. Jeśli to a, to $c = c + 1$
5. Jeśli to b, to $c = c - 1$
6. wróć do 2.

Przykład:

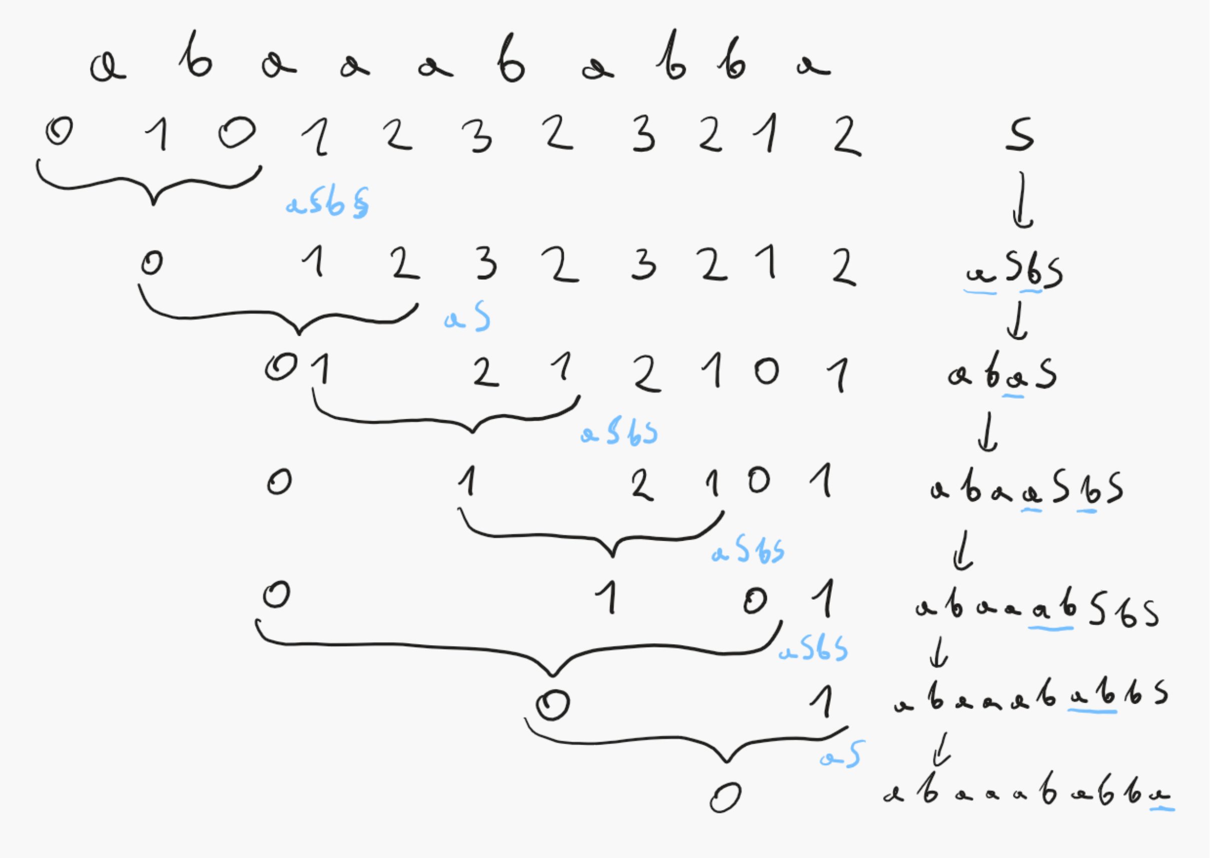
$$a \quad b \quad a \quad a \quad b \quad a \quad b \quad b \quad a$$

$$0 \quad 1 \quad 0 \quad 1 \quad 2 \quad 1 \quad 2 \quad 1 \quad 0 \quad 1$$

Zauważmy, że każdy krok takiego wyprowadzenia należy do języka (dzięki $S \to \epsilon$) i na każdym kroku ciąg z nim powiązany będzie zaczynał się na 0*12 lub będzie miał fragment (n)(n+1)(n) (dla słów większych niż jedna litera). Algorytm:

- Jeśli początek to 0*12 i każda następna liczba jest większa niż 0, to zmieniamy dopasowane 12 na 1 i każdą następną liczbę zmniejszamy o 1 (odpowiada to produkcji $S \to aS$)
- W przeciwnym wypadku znajdujemy najbliższe dopasowanie podciągu do (n)(n+1)(n) i zastępujemy ten fragment liczbą (n) (odpowiada to produkcji $S \to aSbS$)

Powtarzamy, aż nie zostaniemy z jednym zerem (koniec), lub ciągiem 01 (mamy jeszcze jedną produkcję $S \to aS$). Otrzymujemy w ten sposób ciąg produkcji prowadzących do słowa. Przykład:



Zatem $L = L(G)$, co kończy dowód.