Mateusz Basiak, 27.03.2020r.

Definicja Dla języka L nad alfabetem Σ^* definiujemy język $L/2$ jako:

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

Zadanie 56. Pokaż że istnieje taki CFL L, dla którego $L/2$ nie jest CFL.

Rozwiązanie
Weźmy język:

$$L = \{a^i b^j c^j aad^{3i} \ : \ i, j \in \mathbb{N}\}$$

Pokażemy, że ten język jest bezkontekstowy oraz że $L/2$ nie jest bezkontekstowy.

Lemat 1. *L jest bezkontekstowy*

Dowód. L jest generowany przez gramatykę bezkontekstową G o terminalach a, b, c, d, nieterminalach S, T, symbolu startowym S i przejściach:

$$S \to aSddd|Taa$$

$$T \to bTc|\varepsilon$$

Trzeba teraz pokazać $L = L_G$.
1. $L \subseteq L_G$
Weźmy dowolne słowo $x \in L$. Słowo to jest postaci $x = a^i b^j c^j aad^{3i}$ dla pewnych $i, j \in \mathbb{N}$. Można je uzyskać w gramatyce G wykonując najpierw i razy przejście $S \to aSddd$, następnie raz $S \to Taa$, j razy $T \to bTc$ i na koniec z T uzyskać słowo puste. Oznacza to, że $x \in L$.

2. $L_G \subseteq L$
Weźmy teraz dowolne słowo $x \in L_G$. Na początku zauważmy, że w każdym przejściu w G z prawej strony jest tylko jeden nieterminal, więc na każdym etapie tworzenia słowa przez tę gramatykę może mieć co najwyżej jeden nieterminal. Dodatkowo zauważmy, że jest tylko jeden sposób pozbycia się nieterminala ($T \to \varepsilon$), że z S do T da się przejść na dokładnie jeden sposób, oraz że dla każdego z nieterminali da się przejść z niego do słowa z nim na jeden sposób. Wynika stąd, że każde słowo w tej gramatyce można wyprowadzić na dokładnie jeden sposób.
Niech więc $S, x_1, x_2, \cdots x_{n-1}, x_n = x$ będą kolejnymi słowami przez które przechodzimy konstruując słowo x. Jako, że przejście z x_{n-1} do x musi zajść używając $T \to \varepsilon$, w słowie x_{n-1} musi być nieterminal T, czyli musi istnieć takie $i < n-1$, że między x_i a x_{i+1} jest przejście $S \to Taa$. Wówczas dla $j < i$ wszystkie przejścia były postaci $S \to aSddd$, czyli $x_i = a^i Sd^{3i}$ oraz $x_{i+1} = a^i Taad^{3i}$. Niech teraz $k = n - 1 - (i + 1)$. Pomiędzy x_{i+1} a x_{n-1} wszystkie przejścia muszą być postaci $T \to bTc$, czyli $x_{n-1} = a^i b^k Tc^k aad^{3i}$ i ostatecznie $x = x_n = a^i b^k c^k aad^{3i} \in L$. $\square$

Lemat 2. *$L/2$ nie jest bezkontekstowy*

Dowód. Załóżmy nie wprost, że $L/2$ jest bezkontekstowy. Wówczas istnieje dla niego stała $n \in \mathbb{N}$ z lematu o pompowaniu dla języków bezkontekstowych. Weźmy słowo $w = a^n b^n c^n a$. Należy ono do $L/2$ ponieważ dla słowa $v = ad^{3n}$ słowo $wv = a^n b^n c^n aad^{3n} \in L$ oraz $|w| = |v|$. Ponadto $|w| > n$, więc istnieje podział $w = uvxyz$ zgodny z zasadami z lematu o pompowaniu.
Zauważmy na początek, że jeśli v lub y obejmuje a będące ostatnim elementem w, to słowo uv^3xy^3z ma po literkach c trzy literki a, co sprawia, że niezależnie od dołączonego do niego sufiksu s słowo uv^3xy^3zs nie będzie w L, czyli uv^3xy^3z nie będzie w $L/2$. Stąd v ani y na pewno nie obejmują tego końcowego a.
Dalej zauważmy, że ze względu na długość segmentów literek a, b i c oraz warunek $|vxy| \leq n$ z lematu o pompowaniu, słowa v i y muszą się w całości zawierać w jednym ($uxy \in a^n$ lub $uxy \in b^n$ lub $uxy \in c^n$) lub dwóch sąsiadujących segmentach ($uxy \in a^n b^n$ lub $uxy \in b^n c^n$) tych samych literek. Daje to 5 możliwości:

- v, y należą tylko do b^n lub tylko do c^n lub do $a^n b^n$
  Weźmy dowolne $k \geq 2$. Wówczas w słowie $uv^k xy^k z$ będzie różna liczba liter b od liczby liter c, bo zwiększamy liczbę tylko jednej z tych liter. Załóżmy teraz, że jest takie $s \in \{a, b, c, d\}^*$, że $uv^k xy^k zs \in L$. Zauważmy, że przez a na końcu słowa $uv^k xy^k z$ nie możemy w s dodać literek b ani c, bo wtedy nasze słowo na pewno nie będzie w L - po parze literek a mogą być już tylko literki d. Stąd ciągi liter b i c będą miały różną długość, czyli $uv^k xy^k zs$ nie należy do L, czyli $uv^k xy^k z$ nie należy do $L/2$.

- v, y należą do $b^n c^n$
  Weźmy dowolne $k \geq 2$. Wówczas dla pewnych $g, h > 0$ słowo $uv^k xy^k z$ można przedstawić jako $a^n b^{n+g} c^{n+h} a$. Załóżmy teraz, że jest takie $s \in \{a, b, c, d\}^*$, że $uv^k xy^k zs \in L$. Wówczas $|s| = |a^n b^{n+g} c^{n+h} a| = 3n+g+h+1$. Jako że w słowie $uv^k xy^k z$ są już segmenty literek a, b, c i pojedyncze a, to aby całe słowo było postaci słów z L, sufiks s musi być postaci ad^p dla pewnego p. Z długości słowa s wiemy, że $p = 3n + g + h$. Jednak z definicji języka L mamy, że długość segmentu literek d powinna być trzykrotnie większa niż długość segmentu literek a, a w tym wypadku $3n \neq 3n + g + h$. Stąd nie istnieje takie s i $uv^k xy^k z$ nie należy do $L/2$.

- v, y należą tylko do a^n
  Weźmy dowolne $k \geq 2$. Wówczas dla pewnego $g > 0$ słowo $uv^k xy^k z$ można przedstawić jako $a^{n+g} b^n c^n a$. Załóżmy teraz, że jest takie $s \in \{a, b, c, d\}^*$, że $uv^k xy^k zs \in L$. Wówczas $|s| = |a^{n+g} b^n c^n a| = 3n+g+1$. Jako że w słowie $uv^k xy^k z$ są już segmenty literek a, b, c i pojedyncze a, to aby całe słowo było postaci słów z L, sufiks s musi być postaci ad^p dla pewnego p. Z długości słowa s wiemy, że $p = 3n + g$. Jednak z definicji języka L mamy, że długość segmentu literek d powinna być trzykrotnie większa niż długość segmentu literek a, a w tym wypadku $3(n + g) \neq 3n + g$. Stąd nie istnieje takie s i $uv^k xy^k z$ nie należy do $L/2$.

W każdym z powyższych przypadków da się tak napompować słowo w, że po napompowaniu nie będzie w $L/2$. Daje to sprzeczność z lematem o pompowaniu. Czyli $L/2$ nie jest bezkontekstowy. $\square$