

Контрольная работа №2

Вариант 14

- Язык всех слов, порождаемых грамматикой $S \rightarrow bSaS, S \rightarrow aSbS, S \rightarrow a$ из начального нетерминала S , таких что в них максимальный отрезок только из букв b длиннее, чем максимальный отрезок из букв a .

- Язык $\left\{ a^n b^m c^k \mid n \neq m \vee k > n + m \right\}$.

- Язык, определяемый следующей атрибутной грамматикой:

$$\begin{aligned} S &\rightarrow ASA & A_1.b &> A_2.b, S_1.b > A_2.b, S_0.b := S_1.b + 2 \cdot A_1.b \\ S &\rightarrow b & S.b &:= 1 \\ A &\rightarrow aA & A_0.b &:= A_1.b \\ A &\rightarrow bbA & A_0.b &:= A_1.b + 2 \\ A &\rightarrow \varepsilon & A.b &:= 0 \end{aligned}$$

Задание 1

Слова образованы с помощью грамматики $S \rightarrow bSaS \mid aSbS \mid a$, и в них максимальный отрезок из букв b длиннее максимального отрезка из букв a .

Рассмотрим семейство слов языка, получающиеся алгоритмом $b\underline{S}aS \rightarrow bb\underline{S}aSaS \rightarrow bbb\underline{S}aSaSaS \rightarrow \dots$

Изначально слово имеет одну букву b , одну буквы a и два нетерминала S . После каждого раскрытия S общее количество a , b и S увеличивается на 1.

Рассмотрим слова у которых по итогу количество букв b в начале будет нечетно.

$$b^{2p+1}(Sa)^{2p+1}S.$$

Преобразуем выражение:

$$b^{2p+1}(Sa)^{2p+1}S = b^{2p+1}(Sa)^p Sa(Sa)^p S.$$

Далее применяем правила грамматики:

$$b^{2p+1}(Sa)^p Sa(Sa)^p S \Rightarrow b^{2p+1}a^{2p}b(Sa)^p SaSaS$$

$$\Rightarrow b^{2p+1}a^{2p}ba^{2p}bSaSaSaS$$

$$\Rightarrow b^{2p+1}a^{2p}ba^{2p}ba^7.$$

Таким образом, слово

$$b^{2p+1}a^{2p}ba^{2p}ba^7$$

принадлежит рассматриваемому языку при $p \geq 4$.

По лемме о накачке для контекстно-свободных языков существует p_1 , удовлетворяющее условиям леммы. Рассмотрим $n > p_1$ и слово

$$w = b^{2n+1}a^{2n}ba^{2n}ba^7.$$

Разобъём слово на шесть частей:

$$w = \underbrace{b^{2n+1}}_{q_1} \underbrace{a^{2n}}_{q_2} \underbrace{b}_{q_3} \underbrace{a^{2n}}_{q_4} \underbrace{b}_{q_5} \underbrace{a^7}_{q_6}.$$

Рассмотрим нетривиальные положения накачиваемой подстроки vxy :

- q_1 : при $i = 0$ длина блока из b уменьшается, и слово перестаёт принадлежать языку.
- q_2 : при $i \rightarrow \infty$ длина подстроки из a больше чем из b .
- q_3 : при $i = 0$ получаем блок a^{4n} , что невозможно по определению языка.
- q_4 : аналогично случаю q_2 .
- q_5 : при $i = 0$ остаётся суффикс a^{2n+7} , что не допускается.
- q_6 : при $i \rightarrow \infty$ длина подстроки из a больше чем из b .

Рассмотрим случаи, когда vxy пересекает границы частей:

- q_1, q_2 : при $i = 0$ длина подстроки из b становится не больше длины подстроки из a , что противоречит свойствам языка.
- q_2, q_3 : при $i = 0$ и $i \rightarrow \infty$ нарушается структура слова или длина подстроки из a больше чем из b .
- q_3, q_4 : аналогично предыдущему случаю.
- q_4, q_5 : аналогично предыдущему случаю.
- q_5, q_6 : при $i \rightarrow \infty$ слово не принадлежит языку.
- q_2, q_3, q_4 : при $i \rightarrow \infty$ количество подряд идущих букв a становится больше количества подряд идущих букв b .
- q_3, q_4, q_5 : при $i \rightarrow \infty$ количество букв b становится больше количества букв a или длина подстроки из a больше чем из b .
- q_4, q_5, q_6 : при $i \rightarrow \infty$ количество подряд идущих букв a становится больше количества подряд идущих букв b .

Прочие тривиальные случаи также не удовлетворяют условиям леммы о накачке. Следовательно, данный язык не является контекстно-свободным.

Задание 2

$$L = \{ a^n b^n c^k \mid n \neq m \vee k > n + m \}.$$

Возьмём дополнение языка:

$$L' = \{ a^n b^n c^k \mid \underbrace{n = m}_{(1)} \wedge \underbrace{k \leq n + m}_{(2)} \}.$$

Пусть по лемме о накачке для контекстно-свободных языков существует число p , удовлетворяющее условиям леммы. Рассмотрим $n > p$ и слово

$$w = \underbrace{a^n}_{q_1} \underbrace{b^n}_{q_2} \underbrace{c^{2n}}_{q_3}.$$

Рассмотрим возможные случаи накачки подстроки vxy :

- q_1 : при любом $i \neq 1$ слово перестаёт удовлетворять правилу (1).
- q_2 : аналогично случаю q_1 .
- q_3 : при $i \rightarrow \infty$ слово нарушает правило (2).
- q_1, q_2 : либо слово перестанет принадлежать языку по своей структуре, либо при $i = 0$ нарушается правило (2).
- q_2, q_3 : при $i \rightarrow \infty$ нарушаются правила (2).

Следовательно, дополнение языка L' не является контекстно-свободным, а значит исходный язык L не DCFL.

Задание 3

$$\begin{aligned}
 S &\rightarrow ASA ; A_1.b > A_2.b, S_1.b > A_2.b, S_0.b := S_1.b + 2 \cdot A_1.b \\
 S &\rightarrow b ; S.b := 1 \\
 A &\rightarrow aA ; A_0.b := A_1.b \\
 A &\rightarrow bbA ; A_0.b := A_1.b + 2 \\
 A &\rightarrow \varepsilon ; A.b := 0
 \end{aligned}$$

Проанализируем слова, принадлежащие языку. Фактически слова разбиваются на две части, где разделителем служит буква b . В левой части количество b больше, чем в правой, а общее количество букв b нечётное. Атрибуты грамматики отвечают за подсчёт количества букв b .

Пересечём язык с регулярным выражением $(bb)^*ba(bb)^*$. Тогда получаем язык

$$L = \{ b^n ab^m \mid n > m, n \text{ нечёт}, m \text{ чёт} + \text{еще условия} \}.$$

Так как есть условие $A_1.b > A_2.b$, выясним, при каких значениях m возможно минимальное значение n :

- $m = 0$: $n = 3$, слово $bbba$ принадлежит L
- $m = 2$: $n \neq 5$, $n = 7$, слово $bbbbbbabb$ принадлежит L , так как при $n = 5$ невозможно разбиение слова без нарушения условия грамматики
- $m = 4$: $n = 9$
- $m = 6$: $n \neq 11$, минимальное $n = 13$, аналогично при $m = 2$
- $m = 8$: $n = 15$
- $m = 10$: $n = 17$
- $m = 12$: $n = 19$
- $m = 14$: $n = 21$

Рассмотрим последовательности слов:

$$\begin{aligned}
 b^3a \in L, & \quad (n, m) = (3, 0); \quad (n, m) \neq (5, 2) \implies (n, m) = (7, 2) \\
 b^6(b^3a)b^4 \in L, & \quad (n, m) = (9, 4); \quad (n, m) \neq (11, 6) \implies (n, m) = (13, 6) \\
 b^{18}(b^6(b^3a)b^4)b^{16} \in L, & \quad (n, m) = (27, 20); \quad (n, m) \neq (29, 22) \implies (n, m) = (31, 22) \\
 b^{54}(b^{18}(b^6(b^3a)b^4)b^{16})b^{52} \in L, & \quad (n, m) = (81, 72); \quad (n, m) \neq (83, 74) \implies (n, m) = (85, 74) \\
 \dots
 \end{aligned}$$

В парах (n, m) , указанных в неравенствах, не существует сценария, при котором не нарушается правило

$$S_1.b > A_2.b \quad \text{при} \quad S_0.b := S_1.b + 2 \cdot A_1.b.$$

Можно заметить, что количество букв b слева удовлетворяет $n \neq 3^l + 2$, где $l \geq 1$. Как только $n = 3^l$, следующее возможное значение $n = 3^l + 4$, и разность $n - m$ увеличивается на 2.

Таким образом, для $n \in [3^l + 4, 3^{l+1}]$ имеем:

$$n - m = 2l + 3; \quad m = n - 2l - 3.$$

Введём функцию $f(x)$, которая для натуральных нечётных чисел $x \geq 3$ находит подходящее l через промежуток $[3^l + 4, 3^{l+1}]$ и вычисляет $m = f(x)$. Тогда слова языка можно строить как

$$b^n a b^{f(n)}.$$

Применим лемму о накачке. Пусть существует число p , удовлетворяющее условию леммы. Рассмотрим слова

$$w = b^n a b^{f(n)}, \quad \text{где } f(n) > p.$$

Разобьём слово на три части для применения леммы о накачке:

$$w = \underbrace{b^n}_{q_1} a \underbrace{b^{f(n)}}_{q_2}.$$

- q_1 : при $i = 0$ уменьшается количество букв b слева, и меньше быть не может, если зафиксировать $f(n)$.
- q_2 : при $i \rightarrow \infty$ получаем $f(n) > n$, что невозможно.
- q_1, a : нарушается структура слова или количество $a > 1$.
- a, q_2 : нарушается структура слова или количество $a > 1$.
- q_1, a, q_3 : если v или y в vxy содержат букву a , то портится структура слова; если же $v \subseteq q_1$ и $y \subseteq q_3$, тогда для определенного i получаем

$$f(n') \neq m',$$

где n' и m' — новое количество букв b слева и справа. Это связано с тем, что количество букв b растёт линейно (если накачивается одна буква) или экспоненциально, а функция $f(x)$ кусочно-линейная.

Следовательно, такие слова невозможны накачать, и по лемме о накачке язык L не является контекстно-свободным. А значит, исходный язык также не контекстно-свободный.