Формальные грамматики. HW#1

Тураев Тимур, SPbAU, SE, 604 group

1. *Построить обыкновенную грамматику для языка всех палиндромов* $L_1 = \{w \in \{a, b\}^* \mid w = w^R\}$. *Показать, как строка* $ababa$ *выводится с помощью перезаписи строк. Показать, что эта же строка принадлежит наименьшему решению системы языковых уравнений, построив несколько шагов последовательности* $\varphi^k(\perp)$

Обыкновенная грамматика:

$$S \to aSa \mid bSb \mid a \mid b \mid \varepsilon$$

Вывод строки **ababa** с помощью перезаписи строк:

$$S \to aSa \to abSba \to \mathbf{ababa}$$

Несколько шагов последовательности $\varphi^k(\perp)$:
$S = (\{a\} \cdot S \cdot \{a\}) \cup (\{b\} \cdot S \cdot \{b\}) \cup \{a\} \cup \{b\} \cup \{\varepsilon\}$

$$\begin{aligned}
\varphi^0(\perp) =& \varnothing \\
\varphi^1(\perp) =& \{a, b, \varepsilon\} \\
\varphi^2(\perp) =& \{aaa, bab, aba, bbb, aa, bb, a, b, \varepsilon\} \\
\varphi^3(\perp) =& \{aaaaa, \mathbf{ababa}, aabaa, abbba, aaaa, abba, aaa, aba, aa, \\
& baaab, bbabb, babab, bbbbb, baab, bbbb, bab, bbb, bb, a, b, \varepsilon\}
\end{aligned}$$

2. *Доказать, что не существует обыкновенной грамматики для языка* $L_2 = \{a^{k_1}b \dots a^{k_n}b \mid n \geqslant 1, 0 \leqslant k_1 \leqslant \dots \leqslant k_n\}$

Предположим, грамматика существует и пусть для константы $p \geqslant 1$ выполняется лемма накачки. Рассмотрим строку $w = a^p b a^p b a^p b$. Для этой строки должно существовать разбиение, допускающее накачку. Рассмотрим ряд случаев:

- Можно легко показать (практически по аналогии с лекционными заметками), что, если либо в u или v есть символ b, то строка не допускает накачки: можно накачать так, что пропадет условие неубывания длины блоков.
- Если оба u или v лежат в одном блоке a^p, то можно этот блок размножить так, что условие неубывания не выполнится, или наоборот, удалить его (в случае, если они лежат в третьем блоке), что тоже приведет к нарушению условия.
- Пусть u лежит в первом блоке a^p, а v во втором, тогда накачав их, получим, что в первом и втором блоке число символов a больше, чем в третьем.
- Пусть u лежит во втором блоке a^p, а v в третьем, тогда удалив их, получим, что в первом блоке число символов a больше, чем в во втором или третьем.
- u и v не могут лежать в первом и третьем блоках в силу условия $|uyv| \leqslant p$.

3. *Построить конъюнктивную грамматику для языка* L_2.

$$\begin{aligned}
S &\to Ab \mid SAb \;\&\; EBb \\
A &\to aA \mid \varepsilon \\
B &\to aBa \mid bA \\
E &\to aE \mid bE \mid \varepsilon
\end{aligned}$$

Пояснение: правило SAb просто определяет набор блоков. Нужно еще убедиться, что последний дописанный блок по мощности не меньше чем предпоследний. Это проверяет правило B.

4. *Построить однозначную обыкновенную грамматику для языка* $L_3 = \{c^m a^{\ell_0} b \dots a^{\ell_{m-1}} b a^{\ell_m} b \dots a^{\ell_z} b d^n \mid m, n, \ell_i \geqslant 0, z \geqslant 1, \ell_m = n\}$

$$
\begin{aligned}
&S \to B_1 B_2 \\
&B_1 \to c B_1 A \mid \varepsilon \\
&B_2 \to a B_2 d \mid b E \\
&E \to A E \mid A \\
&A \to a A \mid b
\end{aligned}
$$

Пояснение: строку S можно построить путем конкатенации двух строк: сначала мы генерируем m символов c и такое же количество блоков A. Затем $l_m = n$ символов a и такое же количество символов d, это заканчивается символом b, который завершает блок мощности n, и ненулевым количеством блоков A (из-за условия $z \geqslant 1$)

5. *Является ли язык* L_3 *линейным конъюнктивным?*

6. *Пусть* $D = \{\varepsilon, ab, aabb, abab, aaabbb, \dots\}$ *— язык Дика над алфавитом* $\{a, b\}$*. Существует ли грамматика обёртывания пар для языка* $L_4 = \{wc^{|w|} \mid w \in D\} = \{\varepsilon, abcc, aabbcccc, ababcccc, aaabbbcccccc, \dots\}$*?*

$$S \to (a : c) S (b : c) \mid SS \mid (\varepsilon : \varepsilon)$$

Пояснение: логика построения языка практически идентична грамматике обертывания пар для языка $\{ww \mid w \in \{a, b\}^*\}$.

7. *Построить грамматику 1-го порядка для языка* $\{ww \mid w \in \{a, b\}^*\}$.

$$
\begin{aligned}
S(x, y) &= (\exists z)(A(x, z, z, y)) \\
A(x, y, u, v) &= (x = y \wedge u = v) \vee (((a(x+1) \wedge a(y+1)) \vee (b(x+1) \wedge b(y+1))) \wedge A(x+1, y, u+1, v))
\end{aligned}
$$

Пояснение: существует такая позиция z в строке, которая делит эту строку на две равных. Далее простой логикой первого порядка проверяем, что эти строки либо пусты, либо равны (равны их первые символы и равны суффиксы длиной на единицу меньше, чем длина исходной строки)