Теория и реализация языков программирования. Задание 6: Грамматики

Сергей Володин, 272 гр.

задано 2013.10.09

Задача 1

Идея взята у Николая Ионанова (272), с Хабрахабра, из книги Серебрякова.
$\Sigma \overset{\text{def}}{=} \{a\}$. $\Sigma^* \supset L \overset{\text{def}}{=} \{a^{n^2} \big| n \in \mathbb{N} \cup \{0\}\}$. $\Gamma \overset{\text{def}}{=} (N, \Sigma, P, S)$. $N \overset{\text{def}}{=} \{S, A, B, C, D\}$. Список правил P:

(1) Правила, порождающие $S \Longrightarrow^* LA^nCB^nR$:

1. $S \longrightarrow LDR$
2. $D \longrightarrow ADB$
3. $D \longrightarrow C$

(2) Правила, реализующие $LA^nCB^nR \Longrightarrow^* a^{n^2}$.

1. $ACB \longrightarrow BACa$ — AC и B меняются с порождением a (AC движется вправо)
2. $ACa \longrightarrow aAC$ — AC и a меняются (AC движется вправо)
3. $ACR \longrightarrow CR$ — символ A удаляется, когда дошел до правого конца строки с R
4. $BC \longrightarrow CB$ — B и C меняются (C движется влево)
5. $aC \longrightarrow Ca$ — a и C меняются (C движется влево)
6. $La \longrightarrow aL$ — L и a меняются (L движется вправо)
7. $LCa \longrightarrow aLC$ — LC и a меняются (LC движется вправо)
8. $LCB \longrightarrow LC$ — «LC удаляет B»
9. $LCR \longrightarrow \varepsilon$ — LCR удаляются (конец)

Объяснение

Первая часть порождает $S \Longrightarrow^* LA^nSB^nR$. Далее последний символ A вместе с C движется направо (правила 21, 22), порождая один символ a при проходе через B (правило 21). Всего символов B n, поэтому каждая из n A породит n символов a. Когда AC дойдет до R, симол A удалится (правило 23). Далее символ C пойдет налево до первого символа A (правила 24 и 25).

Доказательство

1. Рассмотрим вывод слова w в Г. Пусть $\{x_k\}_{k=0}^{I}$ — последовательность слов при выводе, $\{p_k\}_{k=1}^{I}$ — примененные правила.
2. Докажем, что $\exists k_0 \colon \forall k > k_0 \hookrightarrow p_k$ из (2), $\forall k \leqslant k_0 \hookrightarrow p_k$ из (1).
   (a) $x_0 = S$
   (b) S есть слева только в 11, поэтому p_1 =11. Поэтому $w_2 = LDR$.
   (c) К LDR нельзя применить никакие правила, кроме 12 и 13. Получим $LADBR$ или LCR.
   (d) К $LADBR$ аналогично нельзя применить ничего, кроме 12 и 13. Получим $LAADBBR$ или $LACBR$.
   (e) Повторяя рассуждения, получим, что $\exists k_0 \colon x_{k_0-1} \overset{13}{\Longrightarrow} x_{k_0} = LA^nCB^nR$.
   (f) Далее правила из (1) не могут быть применены, так как в слове нет ни S, ни D, и правила из (2) не порождают эти символы. Значит, далее правила из (2).

Демонстрация работы грамматики

После применения правил из (1) дальнейшее применение правил определено однозначно. Следующая программа моделирует вывод слова a^{n^2} из LA^nCB^nR, используя правила из (2):
https://bitbucket.org/etoestja/inf/raw/HEAD/mipt/s3/TIPL/6/KCsim.cpp — исходный код,
https://bitbucket.org/etoestja/inf/raw/HEAD/mipt/s3/TIPL/6/task1KC — правила и входная строка. Запускать:

```
$ g++ KCsim.cpp -o KCsim
$ cat task1KC | ./KCsim
```

Результат работы (пример вывода $LA^2CB^2R \Longrightarrow^* a^{2^2} \equiv a^4$):
Над стрелками указаны правила из (2). Подчеркнута та часть, которая будет заменена.
$LA\underline{ACB}BR \xRightarrow{(1)} LAB\underline{ACa}BR \xRightarrow{(2)} LABa\underline{ACB}R \xRightarrow{(1)} LABaB\underline{ACa}R \xRightarrow{(2)} LABaBa\underline{ACR} \xRightarrow{(3)} LABaB\underline{aC}R \xRightarrow{(5)} LABa\underline{BC}aR \xRightarrow{(4)}$
$LAB\underline{aC}BaR \xRightarrow{(5)} LA\underline{BC}aBaR \xRightarrow{(4)} L\underline{ACB}aBaR \xRightarrow{(1)} LB\underline{ACa}aBaR \xRightarrow{(2)} LBa\underline{ACa}BaR \xRightarrow{(2)} LBaa\underline{ACB}aR \xRightarrow{(1)} LBaaB\underline{ACa}aR \xRightarrow{(2)}$
$LBaaBa\underline{ACa}R \xRightarrow{(2)} LBaaBaa\underline{ACR} \xRightarrow{(3)} LBaaBa\underline{aC}R \xRightarrow{(5)} LBaaB\underline{aC}aR \xRightarrow{(5)} LBaa\underline{BC}aaR \xRightarrow{(4)} LBa\underline{aC}BaaR \xRightarrow{(5)} LB\underline{aC}aBaaR \xRightarrow{(5)}$
$L\underline{BC}aaBaaR \xRightarrow{(4)} \underline{LCB}aaBaaR \xRightarrow{(8)} \underline{LCa}aBaaR \xRightarrow{(7)} a\underline{LCa}BaaR \xRightarrow{(7)} aa\underline{LCB}aaR \xRightarrow{(8)} aa\underline{LCa}aR \xRightarrow{(7)} aaa\underline{LCa}R \xRightarrow{(7)} aaaa\underline{LCR} \xRightarrow{(9)}$
$aaaa$

Задача 2

$\Sigma \overset{\mathrm{def}}{=} \{a, b\}$, $\Sigma^* \supset L \overset{\mathrm{def}}{=} \{w \in \Sigma^* \big| w = w^R\}$ — язык палиндромов из a, b.

i. Определим КС-грамматику $\Gamma \overset{\mathrm{def}}{=} (\{S\}, \Sigma, P, S)$, $P \overset{\mathrm{def}}{=} \Big\{ \underbrace{S \longrightarrow \varepsilon}_{(1)}, \underbrace{S \longrightarrow aSa}_{(2)}, \underbrace{S \longrightarrow bSb}_{(3)}, \underbrace{S \longrightarrow a}_{(4)}, \underbrace{S \longrightarrow b}_{(5)} \Big\}$.

Докажем, что $L(\Gamma) = L$:

(a) $L(\Gamma) \subseteq L$. Пусть $w \in L(\Gamma) \Rightarrow S \Longrightarrow^* w$. $|w| = n$. Рассмотрим последовательность $\{w_k\}_{k=0}^{I} \subset (N \cup \Sigma)^*$ слов в выводе. $w_0 = S$, $w_I = w$. Индукцией по k докажем $P(k) = \big[w_k = w_k^R, \forall i\colon w_k[i] = S \hookrightarrow i = \frac{|w_k|+1}{2}\big]$.

1. $k = 0 \Rightarrow w_k \equiv w_0 = S$. Поэтому $\exists! i = 1\colon w_0[i] = S$. Но $1 \equiv \frac{1+1}{2}$ и $w_0^R = S^R = S = w_0 \Rightarrow P(0)$ ■
2. Пусть $P(n)$, $n < I$. Докажем, $P(n+1)$. $P(n) \Rightarrow w_n = w_n^R$, $\forall i\colon w_n[i] = S \hookrightarrow i = \frac{|w_n|+1}{2}$.
Предположим, что $\nexists i\colon w_n[i] = S \Rightarrow w \in \Sigma^*$. Тогда ни одно правило не может быть применено, так как в левой части каждого правила $S \in N$. Но $n < I$ (это не конец вывода) $\Rightarrow$ противоречие.
Значит, $\exists i\colon w_n[i] = S$. Но $P(n) \Rightarrow \forall i\colon w_n[i] = S \hookrightarrow i = \frac{|w_n|+1}{2}$. Поэтому $\exists! i = \frac{|w_n|+1}{2}\colon w_n[i] = S$. Значит, $w_n = xSy$, $|x| = |y|$, $x, y \in \Sigma^*$. $w_n^R = y^R S x^R$. S в w_n входит один раз $\Rightarrow x = y^R$.
Рассмотрим правила (1)—(4):
(1). $w_n = xSy \xRightarrow{(1)} x\varepsilon y \equiv xy = xx^R = w_{n+1}$ — палиндром: $(xx^R)^R = x^{R^R}x^R = xx^R$. $w_{n+1} = xy \in \Sigma^* \Rightarrow \forall i \hookrightarrow w_{n+1}[i] \neq S \Rightarrow P(n+1)$ ■
(2). $w_n = xSx^R \xRightarrow{(2)} xaSax^R = w_{n+1}$. $w_{n+1}^R = x^{R^R}a^R S^R a^R x^R = xaSax^R \equiv w_{n+1}$. $a \neq S \Rightarrow \exists! i\colon w_{n+1} = S$, $i = \frac{|w_n|+1}{2} + 1 = \frac{|w_n|+3}{2} \equiv \frac{|w_{n+1}|+1}{2} \Rightarrow P(n+1)$ ■.
(3). $w_n = xSx^R \xRightarrow{(3)} xbSbx^R = w_{n+1}$. $w_{n+1}^R = x^{R^R}b^R S^R b^R x^R = xbSbx^R \equiv w_{n+1}$. $b \neq S \Rightarrow \exists! i\colon w_{n+1} = S$, $i = \frac{|w_n|+1}{2} + 1 = \frac{|w_n|+3}{2} \equiv \frac{|w_{n+1}|+1}{2} \Rightarrow P(n+1)$ ■.
(4). $w_n = xSx^R \xRightarrow{(4)} xax^R = w_{n+1}$. $w_{n+1}^R = x^{R^R}a^R x^R = xax^R \equiv w_{n+1}$. $w_{n+1} = xax^R \in \Sigma^* \Rightarrow \forall i \hookrightarrow w_{n+1}[i] \neq S \Rightarrow P(n+1)$ ■
(5). $w_n = xSx^R \xRightarrow{(5)} xbx^R = w_{n+1}$. $w_{n+1}^R = x^{R^R}b^R x^R = xbx^R \equiv w_{n+1}$. $w_{n+1} = xbx^R \in \Sigma^* \Rightarrow \forall i \hookrightarrow w_{n+1}[i] \neq S \Rightarrow P(n+1)$ ■

Итак, доказано $\forall k \in \overline{0, I} \hookrightarrow P(k) \Rightarrow P(I) \Rightarrow w \equiv w_I \overset{P(I)}{=} w_I^R \equiv w^R \Rightarrow w \in L$ ■

(b) $L \subseteq L(\Gamma)$. Пусть $w \in L \Rightarrow w^R = w$. $|w| = n$. Рассмотрим $n \mod 2$:

0. $n \mod 2 = 0 \Rightarrow w = xy$, $|x| = |y|$. $w = w^R \Rightarrow xy = y^R x^R$. Поскольку $|x| = |y|$, $y = x^R \Rightarrow \boxed{w = xx^R}$
0. $n \mod 2 = 1 \Rightarrow w = x\sigma y$, $|x| = |y|$, $\sigma \in \Sigma$. $w = w^R \Rightarrow x\sigma y = y^R \sigma^R x^R = y^R \sigma x^R$. Так как $|x| = |y|$, $y = x^R \Rightarrow \boxed{w = x\sigma x^R}$

Значит, $L = \{xx^R, xax^R, xbx^R \big| x \in \Sigma^*\}$.

Построим вывод $S \Longrightarrow^* xSx^R$:

a. Пусть $x = \varepsilon$. $S \xRightarrow{(1)} \varepsilon = \varepsilon\varepsilon^R = w \Rightarrow w \in L(\Gamma)$ ■.
b. Иначе $x = x_1 ... x_m$, $\forall i \in \overline{1, m} \hookrightarrow x_i \in \Sigma$. Рассмотрим символы $x_m, ..., x_1$. Применим правило (2), если $x_i = a$ и (3) иначе. Примененное правило обозначим за $R(i)$ Получим $S \xRightarrow{(R(m))} x_m S x_m \Longrightarrow ... \xRightarrow{(R(1))} x_1 ... x_m S x_m ... x_1$.

Теперь покажем, как получить w:

1. $w = xx^R$. Было получено $S \Longrightarrow^* xSx^R$. Тогда $S \Longrightarrow^* xSx^R \xRightarrow{(1)} xx^R$ ■
2. $w = xax^R$. Было получено $S \Longrightarrow^* xSx^R$. Тогда $S \Longrightarrow^* xSx^R \xRightarrow{(4)} xax^R$ ■

3. $w = xbx^R$. Было получено $S \Longrightarrow^* xSx^R$. Тогда $S \Longrightarrow^* xSx^R \overset{(5)}{\Longrightarrow} xbx^R$ ■

Получаем $w \in L(\Gamma)$.

$\boxed{\text{Ответ: } \Gamma \overset{\text{def}}{=} (\{S\}, \Sigma, P, S),\ P \overset{\text{def}}{=} \{S \longrightarrow \varepsilon | aSa | bSb | a | b\}.}$

ii. Определим грамматику $\overline{\Gamma} = \{\{S, D\}, \Sigma, \overline{P}, S\}$, $\overline{P} = \{D \longrightarrow \underbrace{aD}_{(1)} | \underbrace{bD}_{(2)} | \underbrace{\varepsilon}_{(3)},\ S \longrightarrow \underbrace{aDb}_{(4)} | \underbrace{bDa}_{(5)} | \underbrace{aSa}_{(6)} | \underbrace{bSb}_{(7)}\}$.

Пояснение: D порождает Σ^*. S порождает непалиндромы. Если первый и последний символ непалиндрома различны, то между ними может быть все, что угодно, а если они одинаковые, то между ними должен быть непалиндром.

1. Докажем $D \Longrightarrow^* w \in \Sigma \Leftrightarrow w \in \Sigma^*$, что равносильно $w \in \Sigma^* \Rightarrow D \Longrightarrow^* w$. Если $w = \varepsilon$, то применим $D \overset{(3)}{\Longrightarrow} \varepsilon \equiv w$. Иначе $w = w_1 ... w_n$, $\forall i \in \overline{1, n} \hookrightarrow w_i \in \Sigma$. Рассмотрим символы $w_1, ..., w_n$, если $w_i = a$, применим (1), иначе применим (2). Примененное правило обозначим за $P(i)$. Тогда $D \overset{P(1)}{\Longrightarrow} w_1 D \overset{P(2)}{\Longrightarrow} ... \overset{P(n)}{\Longrightarrow} w_1 ... w_n D \overset{(3)}{\Longrightarrow} w_1 ... w_n \equiv w$.
2. Рассмотрим вывод в $\overline{\Gamma}$ (из S). Заметим, что правила (6), (7) не изменяют количество вхождений S (оно остается равным 1). Правила (4) и (5) уменьшают количество S на 1. Поэтому в начале вывода несколько применений правил (6) и (7), затем применение (4) или (5).
3. Правила (6) и (7) сохраняют количество символов перед и после $S \Rightarrow$ после применений этих правил перед и после S одинаковое число символов.
4. Докажем $L(\overline{\Gamma}) = \overline{L}$:
   1. $w \in L(\overline{\Gamma})$. Рассмотрим вывод w. Как было показано, перед применением правил (6) или (7) слово имеет вид xSy, $|x| = |y|$, $x, y \in \Sigma^*$. Применим (4), получим $xaDby$. Пусть D порождает $w_D \in \Sigma^*$. Тогда получим $w = xaw_D by \in \Sigma^*$. Но $|x| = |y|$, поэтому $w[|x|+1] = a \neq b = w[|w| - (|x|+1) + 1] \Rightarrow w \notin L \Rightarrow w \in \overline{L}$. Аналогично при применении (5). ■
   2. $w \in \overline{L}$. $n \overset{\text{def}}{=} |w|$. Пусть $i = \min\{i \in \overline{1, n} \,|\, w[i] \neq w[n - i + 1]\}$ *(первое несовпадение)*.
   Пусть $w = xaw_D bx^R$ (случай $w = xbw_D ax^R$ аналогичный). Рассмотрим символы $x = x_1 ... x_k$. Если $x_i = a$, применим правило (6), если $x_i = b$, применим (7). Примененное правило обозначим за $P(i)$. Получаем $S \overset{P(1)}{\Longrightarrow} x_1 S x_1 \overset{P(2)}{\Longrightarrow} ... \overset{P(k)}{\Longrightarrow} x_1 ... x_k S x_k ... x_1 \equiv xSx^R \overset{(4)}{\Longrightarrow} xaDbx^R$. В случае $w = xbw_D ax^R$ нужно последним применить (5) вместо (4). $w_D \in \Sigma^*$. В ii1 было показано, что $D \Longrightarrow^* w_D$. Поэтому $S \overset{\text{см. ранее}}{\Longrightarrow^*} xaDbx^R \Longrightarrow^* xaw_D x^R$.

$\boxed{\text{Ответ: } \overline{\Gamma} = \{\{S, D\}, \Sigma, \overline{P}, S\},\ \overline{P} = \{D \longrightarrow aD | bD | \varepsilon,\ S \longrightarrow aSa | bSb | aDb | bDa\}.}$

Задача 3

$\Sigma \overset{\text{def}}{=} \{a, b\}$. $\Sigma^* \supset L \overset{\text{def}}{=} L^= \overset{\text{def}}{=} \{w \in \Sigma^* \,\big|\, |w|_a = |w|_b\}$. КС-грамматика $\Gamma = \{N, \Sigma, P, S\}$,
$P \overset{\text{def}}{=} \{\underbrace{S \longrightarrow SS}_{(1)}, \underbrace{S \longrightarrow aSb}_{(2)}, \underbrace{S \longrightarrow bSa}_{(3)}, \underbrace{S \longrightarrow \varepsilon}_{(4)}\}$.
Докажем, что $L(\Gamma) = L^=$:

- $L(\Gamma) \subset L$. $w \in L(\Gamma) \Rightarrow S \Longrightarrow^* w$. Пусть $\{w_k\}_{k=0}^I \subset (N \cup \Sigma)^*$ — последовательность слов при выводе.
  $P(k) \overset{\text{def}}{=} \big[|w_k|_a = |w_k|_b\big]$. Докажем, что $\forall k \in \overline{0, I} \hookrightarrow P(k)$:
  1. $k = 0 \Rightarrow w_k \equiv w_0 = S$. $|w_k|_a = |S|_a = 0 = |S|_b = |w_k|_b \Rightarrow P(0)$.
  2. $P(n) \Rightarrow |w_n|_a = |w_n|_b$. $n < I$. Пусть $w_n \overset{(i)}{\Longrightarrow} w_{n+1}$. Каждое из правил (1)—(4) сохраняет равенство между $|w|_a$ и $|w|_b$:
  (1) и (4) не изменяют их, а (2) и (3) увеличивают каждое на 1 $\Rightarrow |w_{n+1}|_a = |w_{n+1}|_b \Rightarrow P(n+1)$

  Получаем $P(I) \Rightarrow |w|_a \equiv |w_I|_a \overset{P(I)}{=} |w_I|_b \equiv |w|_b \Rightarrow w \in L^=$.

- $L \subset L(\Gamma)$.
  Определим $S \colon \Sigma^* \longrightarrow \mathbb{Z}$: $w \in \Sigma^* \Rightarrow S(w) = |w|_a - |w|_b$. $w \in L^= \Leftrightarrow |w|_a = |w|_b \Leftrightarrow S(w) = 0$. $|w| = n$. $w \in \Sigma^* \Rightarrow |w| = |w|_a + |w|_b = 2|w|_a \Rightarrow |w|$ — четно $\Rightarrow n = 2m$.

  Пусть $L \ni w = axa$. Тогда $0 = S(w) = |axa|_a - |axa|_b = 2 + S(x) \Rightarrow S(x) = -2$. Отсюда следует, что $|x| \geqslant 0$. Пусть $|x| = t$, $x = x_1 ... x_t, \forall i \in \overline{1, t} \hookrightarrow x_i \in \Sigma$. Обозначим $f(t) \colon \overline{1, t} \longrightarrow \mathbb{Z}$: $f(i) = S(ax_1 ... x_i)$. Тогда $f(0) \equiv S(a) = 1$, $f(t) \equiv S(ax_1 ... x_t) = 1 + S(x) = 1 - 2 = -1$. Заметим, что $|f(t+1) - f(t)| = 1$ *(«аналог непрерывности»)*. Поэтому $\exists i \in \overline{1, t}$: $f(t) = 0$ *«принимает промежуточное значение»*. Получаем, что $w = ax_1 ... x_i x_{i+1} ... x_t a = w_l w_r$. Поскольку $0 = S(w) = S(w_l) + S(w_r)$ и $S(w_l) \equiv f(i) = 0$, $S(w_r) = 0$. $S(w_l) = S(w_r) = 0 \Rightarrow w_l, w_r \in L$. Поскольку $|w_l|, |w_r| \geqslant |a| = 1, |w_l|, |w_r| \leqslant |w| - 1$. Но $w, w_l, w_r \in L \Rightarrow |w|, |w_l|, |w_r|$ — четные. значит, $|w_l|, |w_r| \leqslant |w| - 2$. Итак, $w = axa \in L \Rightarrow w = w_l w_r, |w_l|, |w_r| \in \overline{1, |w| - 2}, w_l, w_r \in L$. Аналогично доказываем для $L \ni w = bxb$. Получаем $\boxed{w = \sigma x \sigma \in L \Rightarrow w = w_l w_r, |w_l|, |w_r| \in \overline{1, |w| - 2}, w_l, w_r \in L}$.

  $P(m) \overset{\text{def}}{=} \big[\forall w \in L \colon |w| = 2m \hookrightarrow w \in L(\Gamma)\big]$. Докажем $\forall i \geqslant 0 \hookrightarrow P(i)$:

  1. $m = 0 \Rightarrow w = \varepsilon$. $S \overset{(4)}{\Longrightarrow} \varepsilon = w \Rightarrow w \in L(\Gamma)$

2. Пусть $P(m)$. Докажем $P(m+1)$. Рассмотрим $w \in L$: $|w| = 2(m+1) > 2$. Значит, $w = \sigma_1 x \sigma_2$. Заметим, что $|x| = 2m$. Рассмотрим варианты для (σ_1, σ_2):

   1. $\sigma_1 = a, \sigma_2 = b$. Тогда $w \in L \Rightarrow 0 = S(w) = |axb|_a - |axb|_b = 1 + |x|_a - |x|_b - 1 = S(x)$. Как было замечено, $|x| = 2m$, поэтому, по предположению индукции, $S \overset{P(m)}{\Longrightarrow^*} x$. Но $S \overset{(2)}{\Longrightarrow} aSb \overset{P(m)}{\Longrightarrow^*} axb \Rightarrow w \in L(\Gamma)$

   2. $\sigma_1 = b, \sigma_2 = a$. Аналогично получаем $w = bxa$, $x \in L(\Gamma) \Rightarrow S \overset{(3)}{\Longrightarrow} bSa \overset{P(m)}{\Longrightarrow^*} bxa \Rightarrow w \in L(\Gamma)$

3, 4. $\sigma \overset{\text{def}}{=} \sigma_1 = \sigma_2$. Разобьем слово $w = \sigma x \sigma$ на *подслова* $w = w_1 ... w_k$, $\forall i \in \overline{1,k} \hookrightarrow w_i \in \Sigma^* \cap L$, $|w_i| \leqslant |w| - 2$, $w_i[1] \neq w_i[|w_i|]$.
Для этого воспользуемся утверждением в рамочке (см. выше): разобьем $w = w_l w_r$, потом, если первый и последний символы w_l совпадают, повторим для него (возможно, так как $w_l \in L$ по построению): $w_l = w_{ll} w_{lr}$, аналогично для w_r. Всего разбиений будет не больше $|w|$, так как части разбиения непустые (см. утверждение) $\Rightarrow$ алгоритм конечен. Каждое разбиение дает подслова из L — также см. утверждение. И части разбиения не длиннее исходного слова, а также $w_l, w_r \leqslant |w| - 2$. Значит, $w_i \leqslant |w| - 2$. Поэтому $S \overset{P(m)}{\Longrightarrow^*} w_l$, $S \overset{P(m)}{\Longrightarrow^*} w_r$ — по предположению индукции. Покажем, как вывести w из S: воспользуемся правилом (1) $k-1$ раз:
$S \overset{(1)}{\Longrightarrow} SS \overset{(1)}{\Longrightarrow} ... \overset{(1)}{\Longrightarrow} S^k$. Далее воспользуемся выводами w_i: $S^k \overset{\text{вывод } w_1}{\Longrightarrow^*} w_1 S^{k-1} \Longrightarrow^* ... \overset{\text{вывод } w_k}{\Longrightarrow^*} w_1 ... w_k \equiv w \Rightarrow w \in L(\Gamma)$

Задача 4

$\Sigma \overset{\text{def}}{=} \{a, b\}$, $\Sigma^* \supset L \overset{\text{def}}{=} \{w \in \Sigma^* \big| |w|_b \leqslant |w|_a\} \equiv \{w \in \Sigma^* \big| S(w) \geqslant 0\}$ *(определение $S(w)$ см. в задаче 3).*
Определим КС-грамматику $\Gamma = \{\{S\}, \Sigma, P, S\}$,
$P \overset{\text{def}}{=} \{ \underbrace{S \longrightarrow SS}_{(1)}, \underbrace{S \longrightarrow aSb}_{(2)}, \underbrace{S \longrightarrow bSa}_{(3)}, \underbrace{S \longrightarrow \varepsilon}_{(4)}, \underbrace{S \longrightarrow aS}_{(5)} \}$
(добавим в грамматику из предыдущей задачи правила $S \longrightarrow Sa$ и $S \longrightarrow aS$).
Докажем, что $L(\Gamma) = L$:

- $L(\Gamma) \subset L$. $w \in L(\Gamma) \Rightarrow S \Longrightarrow^* w$. Пусть $\{w_k\}_{k=0}^{I} \subset (N \cup \Sigma)^*$ — последовательность слов при выводе.
  $P(k) \overset{\text{def}}{=} \big[S(w_k) \geqslant 0\big]$. Докажем, что $\forall k \in \overline{0, I} \hookrightarrow P(k)$:

  1. $k = 0 \Rightarrow w_k \equiv w_0 = S$. $|w_k|_a = |S|_a = 0 = |S|_b = |w_k|_b \Rightarrow P(0)$.

  2. $P(n) \Rightarrow S(w_n) \geqslant 0$. $n < I$. Пусть $w_n \overset{(i)}{\Longrightarrow} w_{n+1}$. Каждое из правил (1)—(4) не уменьшает разницу $|w|_a - |w|_b \equiv S(w)$: (1) и (4) не изменяют операнды, (2) и (3) увеличивают каждое на 1, (5) увеличивает разницу на 1$\Rightarrow S(w_{n+1}) \geqslant 0 \Rightarrow P(n+1)$

  Получаем $P(I) \Rightarrow S(w) \equiv S(w_I) \geqslant 0 \Rightarrow w \in L$.

- $L \subset L(\Gamma)$. Докажем, что $\forall w \in L \, \exists w^= \in L^=$: w — слово w_0 с добавленными в некоторые позиции символами a. Действительно, рассмотрим w, удалим из него $S(w)$ любых символов a, получим w_0, $S(w_0) = 0 \Rightarrow w_0 \in L^=$, и w получается из w_0 добавлением символов a (в те же позиции, с которых они были удалены).

  Фиксируем $w \in L$, $|w| = n$; $w^=$ найдем из доказанного утверждения выше. $|w^=| = n^=$, $w^= \in L^= \Rightarrow \exists \{x_i\}_{i=0}^{I} \subset \{S\} \cup \Sigma^*$ — последовательность слов при выводе слова $w^=$ в грамматике $\Gamma^=$ *из предыдущей задачи*, $\{p_i\}_{i=1}^{I}$ — примененные правила.
  Определим $f \colon \overline{1, n^=} \longrightarrow \mathbb{N} \cup \{0\}$: $f(i)$ — количество букв a, которые нужно добавить после i-го символа $w^=$, чтобы получить после всех добавлений w, то есть, $w = w_1^= a^{f(1)} w_2^= a^{f(2)} ... w_n^= a^{f(n^=)}$.

  Модифицируем вывод, добавив буквы a.

  Заметим, что если в Γ было применено правило $\alpha_1 \alpha_2 S \alpha_3 \Longrightarrow \alpha_1 \alpha_2 \beta \alpha_3$, то после добавления символа a то же правило из Γ также может быть применено: $\alpha_1 a \alpha_2 S \alpha_3 \Longrightarrow \alpha_1 a \alpha_2 \beta \alpha_3$, аналогично в случае, где a добавлено после S. Иными словами, добавление букв a оставляет возможность применить те же правила к тем же символам S в последующих шагах вывода, причем результатом будет слово с добавленной буквой a. Такие же рассуждения применимы к добавлению многих букв a.

  Каждый символ $w_i^=$ был получен из правил (2) или (3) грамматики $\Gamma^=$ (остальные правила не добавляют терминалов). Пусть это произошло на $j(i)$-м шаге вывода. Покажем, как модифицировать этот шаг, чтобы после i-го символа $w^=$ добавить $f(i)$ букв a:

  1. $w_i^= = a$, $p_{j(i)} = (2)$. То есть $x_{j(i)-1} \equiv \alpha S \beta \overset{(2)}{\Longrightarrow} \alpha \underline{a} S b \beta \equiv x_{j(i)}$. Заменим это на
     $\alpha S \beta \overset{(2)}{\Longrightarrow} \alpha \underline{a} S b \beta \underbrace{\overset{(5)}{\Longrightarrow} \alpha \underline{a} a S b \beta \overset{(5)}{\Longrightarrow} ... \overset{(5)}{\Longrightarrow} \alpha \underline{a} a^{f(i)} S b \beta}_{f(i) \text{ раз}}$ — верное применение правил в Γ.

  2. $w_i^= = a$, $p_{j(i)} = (3)$. То есть $x_{j(i)-1} \equiv \alpha S \beta \overset{(3)}{\Longrightarrow} \alpha b S \underline{a} \beta \equiv x_{j(i)}$. Заменим это на
     $\alpha S \beta \overset{(1)}{\Longrightarrow} \alpha S S \beta \overset{(2)}{\Longrightarrow} \alpha b S \underline{a} S \beta \underbrace{\overset{(5)}{\Longrightarrow} \alpha b S \underline{a} a S \beta \overset{(5)}{\Longrightarrow} ... \overset{(5)}{\Longrightarrow} \alpha b S \underline{a} a^{f(i)} S \beta}_{f(i) \text{ раз}} \overset{(4)}{\Longrightarrow} \alpha b S \underline{a} a^{f(i)} \beta$ — верное применение правил в Γ.

3. $w_i^{=} = b$, $p_{j(i)} = (3)$. То есть $x_{j(i)-1} \equiv \alpha S\beta \xRightarrow{(3)} \alpha\underline{b}Sa\beta \equiv x_{j(i)}$. Заменим это на
$\alpha S\beta \xRightarrow{(3)} \alpha\underline{b}Sa\beta \underbrace{\xRightarrow{(5)} \alpha\underline{b}aSa\beta \xRightarrow{(5)} ... \xRightarrow{(5)} \alpha\underline{b}a^{f(i)}Sa\beta}_{f(i)\text{ раз}}$ — верное применение правил в Γ.

4. $w_i^{=} = b$, $p_{j(i)} = (2)$. То есть $x_{j(i)-1} \equiv \alpha S\beta \xRightarrow{(2)} \alpha aS\underline{b}\beta \equiv x_{j(i)}$. Заменим это на
$\alpha S\beta \xRightarrow{(1)} \alpha SS\beta \xRightarrow{(3)} \alpha aS\underline{b}S\beta \underbrace{\xRightarrow{(5)} \alpha aS\underline{b}aS\beta \xRightarrow{(5)} ... \xRightarrow{(5)} \alpha aS\underline{b}a^{f(i)}S\beta}_{f(i)\text{ раз}} \xRightarrow{(4)} \alpha aS\underline{b}a^{f(i)}\beta$ — верное применение правил в Γ.

Дальнейшее применение правил (после этой измененной части) останется возможным (см. выше), результатом будет «старый» результат, с $f(i)$ буквами a после соответствующего символа (также показано ранее).
Таким образом получено слово $\acute{w} = w_1^{=}...w_i^{=}a^{f(i)}...w_{n^=}^{=}$, получен его вывод в Γ. Применим такие же рассуждения для остальных символов, получим вывод w в $\Gamma \Rightarrow w \in L(\Gamma)$ ■

$\boxed{\text{Ответ: } \Gamma = \{\{S\}, \Sigma, P, S\},\ P = \{S \longrightarrow SS|aSb|bSa|\varepsilon|aS\}.}$

Задача 5