

Теория и реализация языков программирования.

Задание 9: преобразование контекстно-свободных языков

Сергей Володин, 272 гр.

задано 2013.10.30

Упражнение 1

Упражнение 2

N -автомат $M = (\Sigma, \Gamma, Q, q_0, Z_0, \delta, \emptyset)$. $G = (N, \Sigma, P, S)$ — построена по алгоритму. Докажем, что $L(G) \subseteq L(M)$. Будем рассматривать только левые выводы.

1. $\forall w: S \xRightarrow[\text{вывод}]{\text{левый}} w$, $w \in (\Sigma \cup N)^*$, $w \notin \Sigma^* \hookrightarrow w = u[r_0 Y_1 r_1][r_1 Y_2 r_2] \dots [r_{n-1} Y_n r_n]$, $u \in \Sigma^*$, $r_i \in Q$, $Y_i \in \Gamma$ — доказывается индукцией по длине левого вывода из свойств добавленных правил (слева всегда, возможно, нетерминалы, затем, возможно, терминалы. $[qZp] \rightarrow \dots [\dots Yp]$, поэтому соседние состояния, отделенные скобками совпадают: $\dots r][r \dots)$.

2. $P(k) \stackrel{\text{def}}{=} [\forall n \forall w: S \xRightarrow[k \text{ шагов}]{\text{левый вывод}} w \equiv u[r_0 Y_1 r_1] \dots [r_{n-1} Y_n r_n] \hookrightarrow (q_0, u, Z_0) \vdash^* (r_0, \varepsilon, Y_1 \dots Y_n)]$

(a) $n = 1$. Из определения P могут быть только правила $S \rightarrow [q_0 Z_0 p]$, и $(q_0, \varepsilon, Z_0) \equiv (q_0, \varepsilon, Z_0) \Rightarrow P(1)$

(b) Пусть $P(k)$. Рассмотрим левый вывод длины $k + 1$: $S \Rightarrow^* y \equiv u[r_0 Y_1 r_1][r_1 Y_2 r_2] \dots [r_{n-1} Y_n r_n]$. Пусть начальная часть этого вывода длины k имеет вид $S \Rightarrow^* x \equiv u_l[s_0 Z_1 s_1][s_1 Z_2 s_2] \dots [s_{m-1} Z_m s_m]$. На последнем, $k + 1$ -м шаге был раскрыт первый нетерминал $[s_0 Z_1 s_1] \rightarrow z \equiv u_r[t_0 W_1 t_1][t_1 W_2 t_2] \dots [t_{l-1} W_l t_l]$.

Тогда $y = \underbrace{u_l}_{\text{префикс } x} \underbrace{u_r[t_0 W_1 t_1][t_1 W_2 t_2] \dots [t_{l-1} W_l t_l]}_z \underbrace{[s_1 Z_2 s_2] \dots [s_{m-1} Z_m s_m]}_{\text{суффикс } x}$.

Отсюда $W_1 \dots W_l Z_2 \dots Z_m = Y_1 \dots Y_n$, $u = u_l u_r$, $t_0 = r_0$.

$P(k) \Rightarrow (q_0, u_l, Z_0) \vdash^* (s_0, \varepsilon, Z_1 \dots Z_m)$. Применено правило $[s_0 Z_1 s_1] \rightarrow z \Rightarrow (s_0, u_r, Z_1) \vdash (t_0, \varepsilon, W_1 \dots W_l)$.

Тогда $(q_0, u, Z_0) \equiv (q_0, u_l u_r, Z_0) \vdash^* (s_0, u_r, Z_1 \dots Z_m) \vdash (t_0, \varepsilon, W_1 \dots W_l Z_2 \dots Z_m) \equiv (r_0, \varepsilon, Y_1 \dots Y_n)$.

3. Пусть $w \in L(G) \Rightarrow \exists k: S \xRightarrow[n \text{ шагов}]{\text{левый вывод}} w \in \Sigma^* \xRightarrow{P(n)} (q_0, w, Z_0) \vdash^* (q_0, \varepsilon, \varepsilon) \Rightarrow w \in L(M)$

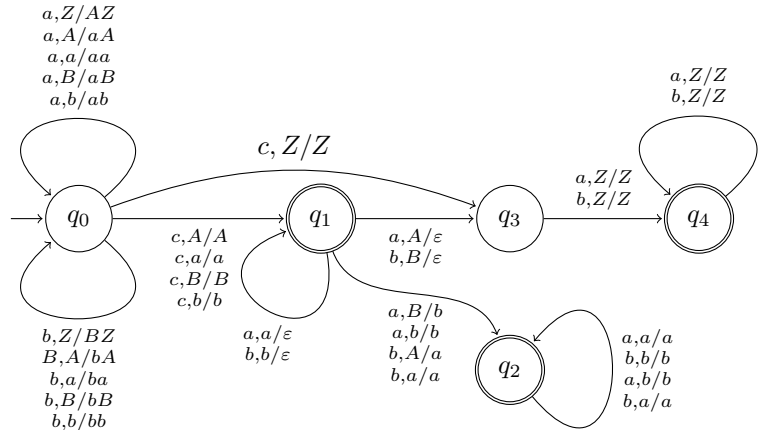
Упражнение 3

Упражнение 4

Задача 1

$L \stackrel{\text{def}}{=} \{xcy \mid x, y \in \{a, b\}^*, x \neq y^R\} \subset \Sigma^*$, $\Sigma \stackrel{\text{def}}{=} \{a, b, c\}$.

1. Определим МП-автомат $\mathcal{A} \stackrel{\text{def}}{=} (\Sigma, \Gamma, Q, q_0, Z, \delta, F)$, допускающий по принимающему состоянию:



1. $\Gamma \stackrel{\text{def}}{=} \{a, b, A, B, Z\}$

2. $Q \stackrel{\text{def}}{=} \{q_0, q_1, q_2, q_3, q_4\}$

3. δ изображена справа

4. $F \stackrel{\text{def}}{=} \{q_1, q_2, q_4\}$

2. \mathcal{A} — детерминированный, так как из каждой конфигурации (q, w, γ) переход определен однозначно, и ε -переходов нет.

3. Определим $U: \{a, b\}^* \rightarrow \{A, B\}^*$: $U(a) = A$, $U(b) = B$. Определим $U_r: \{a, b\}^* \rightarrow \{a, b, A, B\}^*$:

$$U_r(w) = \begin{cases} \varepsilon, & w = \varepsilon \\ w_1 \dots w_{n-1} U(w_n), & w = w_1 \dots w_n, \forall i \in \overline{1, n} \hookrightarrow w_i \in \{a, b\} \end{cases} \text{ — заменяет последний символ на заглавный.}$$

4. Докажем, что $L \subseteq L(\mathcal{A})$:

(a) Пусть $w \in \{a, b\}^*$. Докажем, что $(q_0, w, Z) \vdash^* (q_0, \varepsilon, U_r(w^R)Z)$ индукцией по $|w|$:

$$P(n) \stackrel{\text{def}}{=} [\forall w \in \{a, b\}^*: |w| = n \hookrightarrow (q_0, w, Z) \vdash^* (q_0, \varepsilon, U_r(w^R)Z)]$$

i. $n = 0 \Rightarrow |w| = 0 \Rightarrow w = \varepsilon$. Тогда $U_r(w^R) \equiv \varepsilon$, и $(q_0, w, Z) \equiv (q_0, \varepsilon, Z) \equiv (q_0, \varepsilon, U_r(w^R)Z) \Rightarrow P(0)$.

ii. $n = 1 \Rightarrow w = \sigma \in \Sigma$. Рассмотрим переходы из (q_0, σ, Z) . В стек будет добавлен $U_r(\sigma) \Rightarrow (q_0, w, Z) \equiv (q_0, \sigma, Z) \vdash (q_0, \varepsilon, U_r(\sigma)Z) \equiv (q_0, \varepsilon, U_r(w^R)Z) \Rightarrow P(1)$.

iii. Фиксируем $n \geq 1$, пусть $P(n)$. Пусть $w \in \{a, b\}^*$, $|w| = n + 1$. Тогда $w = w_0 \sigma$, $|w_0| = n > 0$.

$$P(n) \Rightarrow (q_0, w_0, Z) \vdash^* (q_0, \varepsilon, U_r(w_0^R)Z). \text{ Тогда } (q_0, w, Z) \equiv (q_0, w_0 \sigma, Z) \vdash^* (q_0, \sigma, U_r(w_0^R)Z).$$

\nrightarrow переходы из $(q_0, \sigma, U_r(w_0^R)Z)$. На верхушке стека $\gamma \in \{a, b, A, B\}$ — первый символ $U_r(w_0^R)$,

входной символ $\sigma \in \{a, b\}$. Во всех случаях он будет добавлен в стек (см. определение δ), значит, $(q_0, \sigma, U_r(w_0^R)Z) \vdash (q_0, \varepsilon, \sigma U_r(w_0^R)Z) \stackrel{|w_0| > 0}{=} (q_0, \varepsilon, U_r(w^R)Z) \Rightarrow P(n+1)$.

(b) Из определения δ имеем $(q_0, c w, \gamma) \vdash^* (q_1, w, \gamma)$, $|\gamma| > 0, \gamma \neq Z$.

(c) Докажем $(q_1, x, xZ) \vdash^* (q_1, \varepsilon, Z)$ индукцией по $|x|$: $P(n) \stackrel{\text{def}}{=} [\forall x \in \{a, b\}^*: |x| = n \hookrightarrow (q_1, x, xZ) \vdash^* (q_1, \varepsilon, Z)]$

i. $n = 0 \Rightarrow |x| = 0 \Rightarrow x = \varepsilon$. Тогда $(q_1, x, xZ) \equiv (q_1, \varepsilon, Z) \Rightarrow P(0)$

ii. Фиксируем $n \geq 0$. Пусть $P(n)$. Пусть $x \in \{a, b\}^*$: $|x| = n + 1 \Rightarrow x = x_0 \sigma$, $|x_0| = n \stackrel{P(n)}{\Rightarrow} (q_1, x_0, x_0 Z) \vdash^* (q_1, \varepsilon, Z)$. Тогда $(q_1, x, xZ) \equiv (q_1, x_0 \sigma, x_0 \sigma Z) \vdash^* (q_1, \sigma, \sigma Z)$. Входной символ совпадает с символом на верхушке стека, из определения δ получаем, что символ будет удален из стека: $(q_1, \sigma, \sigma Z) \vdash (q_1, \varepsilon, Z) \Rightarrow P(n)$.

(d) Пусть $\sigma_1, \sigma_2 \in \{a, b\}$, $\sigma_1 \neq \sigma_2$. Тогда $(q_1, \sigma_1, U_r(\sigma_2)\gamma) \vdash (q_2, \varepsilon, \sigma_2 \gamma)$ и $(q_1, \sigma_1, \sigma_2 \gamma) \vdash (q_2, \varepsilon, \sigma_2 \gamma)$ — из определения δ .

(e) Пусть $x \in \{a, b\}^*$, $\gamma \in \{a, b\}$. Тогда $(q_2, x, \gamma \kappa) \vdash^* (q_2, \varepsilon, \gamma \kappa)$ — доказывается очевидно по индукции (переходы из q_2 в q_2 определены для всех символов a, b на входе и в стеке и не изменяют стек).

(f) Пусть $\sigma \in \{a, b\}$. Тогда $(q_1, \sigma, U_r(\sigma)\gamma) \vdash (q_3, \varepsilon, \gamma)$ — из определения δ .

(g) Пусть $\sigma \in \{a, b\}$. Тогда $(q_3, \sigma, Z) \vdash (q_4, \varepsilon, Z)$ — из определения δ

(h) Пусть $x \in \{a, b\}^*$. Тогда $(q_4, x, Z) \vdash^* (q_4, \varepsilon, Z)$ — доказывается очевидно по индукции (из q_4 есть переходы в q_4 по a и b при Z на верхушке стека)

(i) Из определения δ имеем $(q_0, c, Z) \vdash (q_3, \varepsilon, Z)$.

(j) Пусть $\underline{w} \in L \Rightarrow w = xcy, x \neq y^R, x, y \in \{a, b\}^*$. $x \neq y^R \Leftrightarrow x^R \neq y$. Выделим максимальную по длине общую часть τ длины i у слов x^R и y : $x^R = \tau x_1, y = \tau y_1, x_1 \neq y_1$. Тогда $x = x_1^R \tau^R, w = xcy = x_1^R \tau^R c \tau y_1$.

i. Пусть $|x_1| > 0$. $(q_0, w, Z) \equiv (q_0, x_1^R \tau^R c \tau y_1, Z) \stackrel{4a}{\vdash^*}_{|x_1| > 0} (q_0, c \tau y_1, U_r(\tau x_1)Z) \vdash (q_1, \tau y_1, U_r(\tau x_1)Z) \stackrel{|x_1| > 0}{=} \equiv (q_1, \tau y_1, \tau U_r(x_1)Z) \stackrel{4c}{\vdash^*} (q_1, y_1, U_r(x_1)Z)$.

А. Пусть $|y_1| > 0$, $x_1[1] \neq y_1[1]$. Обозначим $y_1 = y^1 \dots y^l, \forall i \in \overline{1, l} \hookrightarrow y^i \in \{a, b\}^*$. Тогда $(q_1, y_1, U_r(x_1)Z) \equiv (q_1, y^1 \dots y^l, U_r(x_1)Z) \stackrel{4d}{\vdash} (q_2, y^2 \dots y^l, U_r(x_1)Z) \stackrel{4e}{\vdash^*} (q_2, \varepsilon, U_r(x_1)Z)$. $q_2 \in F \Rightarrow \underline{w} \in L(\mathcal{A})$.

В. Пусть $|y_1| = 0$. Тогда $w = x_1^R \tau^R c \tau y_1 \equiv x_1^R \tau^R c \tau \Rightarrow (q_0, w, Z) \equiv (q_0, x_1^R \tau^R c \tau, Z) \stackrel{4a}{\vdash^*}_{|x_1| > 0} (q_0, c \tau, \tau U_r(x_1)Z) \stackrel{4b}{\vdash}_{|x_1| > 0}$

$$(q_1, \tau, \tau U_r(x_1)Z) \stackrel{4e}{\vdash^*} (q_1, \varepsilon, U_r(x_1)Z). q_1 \in F \Rightarrow \underline{w} \in L(\mathcal{A})$$

ii. Пусть $|x_1| = 0$. Тогда $w = \tau^R c \tau y_1, y_1 \in \{a, b\}^*$. $x^R \neq y \Rightarrow \tau \neq \tau y_1 \Rightarrow |y_1| > 0 \Rightarrow y_1 = \varkappa \Psi, \varkappa \in \{a, b\}$

А. $|\tau| > 0 \Rightarrow \tau = \tau_0 \sigma, \sigma \in \{a, b\}$. Получаем $(q_0, w, Z) \equiv (q_0, \tau^R c \tau y_1, Z) \stackrel{4a}{\vdash^*}_{|\tau| > 0} (q_0, c \tau y_1, U_r(\tau)Z) \stackrel{4b}{\vdash}_{|\tau| > 0} (q_1, \tau y_1, U_r(\tau)Z) \equiv$

$$(q_1, \tau_0 \sigma y_1, \tau_0 U_r(\sigma)Z) \stackrel{4c}{\vdash^*} (q_1, \sigma y_1, U_r(\sigma)Z) \stackrel{4f}{\vdash} (q_3, y_1, Z) \equiv (q_3, \varkappa \Psi, Z) \stackrel{4g}{\vdash} (q_4, \Psi, Z) \stackrel{4h}{\vdash^*} (q_4, \varepsilon, Z). q_4 \in F \Rightarrow \underline{w} \in L(\mathcal{A})$$

В. $|\tau| = 0 \Rightarrow w = x_1^R \tau^R c \tau y_1 \equiv c y_1 \Rightarrow (q_0, w, Z) \equiv (q_0, c y_1, Z) \stackrel{4i}{\vdash} (q_3, y_1, Z) \equiv (q_3, \varkappa \Psi, Z) \stackrel{4g}{\vdash} (q_4, \Psi, Z) \stackrel{4h}{\vdash^*} (q_4, \varepsilon, Z)$. $q_4 \in F \Rightarrow \underline{w} \in L(\mathcal{A})$

5. Докажем, что $L(\mathcal{A}) \subseteq L$. Пусть $w \in L(\mathcal{A}) \Rightarrow (q_0, w, Z) \vdash^* (q, \varepsilon, \gamma), q \in F$:

(a) $q = q_1$. В q_1 прочитываются a, b . Переходы в q_1 есть только из q_0 по c . В q_0 прочитываются символы a, b . Значит, $w = xcy, x, y \in \{a, b\}^*$. Если $x = \varepsilon$, то был совершен переход $q_0 \xrightarrow{c, Z/Z} q_3$ — противоречие. Автомат детерминированный, поэтому цепочка конфигураций при выводе w имеет вид $(q_0, w, Z) \equiv (q_0, xcy, Z) \stackrel{4a}{\vdash}_{|x| > 0} (q_0, cy, U_r(x^R)Z) \stackrel{4b}{\vdash}_{|x| > 0}$

$(q_1, y, U_r(x^R)Z) \stackrel{4c}{\vdash^*}$. Выделим максимальную общую часть от начала для слов x^R и y : $x^R = \tau x_1, y = \tau y_1, x_1 \neq y_1$.

i. $|\tau| = 0, |x_1| = 0 \Rightarrow |x| = 0$ — противоречие

ii. $|\tau| > 0, |x_1| = 0 \Rightarrow \tau = \tau_0 \sigma, \sigma \in \{a, b\}$. $\stackrel{4c}{\vdash^*}(q_1, \tau_0 \sigma y_1, \tau_0 U_r(\sigma)Z) \stackrel{4f}{\vdash} (q_1, \sigma y_1, U_r(\sigma)Z) \vdash (q_3, \dots)$ — противоречие, из q_3 нет переходов в q_1 .

iii. $|\tau| \geq 0, |x_1| > 0$. Тогда $\boxed{=}(q_1, \tau y_1, \tau U_r(x_1)Z) \vdash^* (q_1, y_1, U_r(x_1)Z) \boxed{=1}$.

a. $|y_1| = 0 \Rightarrow \boxed{=1}(q_1, \varepsilon, U_r(x_1))$. Тогда $w = \underbrace{x_1^R \tau^R}_x \underbrace{c \tau y_1^R}_y, x^R = \tau x_1 \neq \tau = y \Rightarrow w \in L$.

b. $|y_1| > 0$. Тогда $x_1[1] \neq y_1[1]$, и $\boxed{=1}(q_1, y_1, U_r(x_1)Z) \vdash^{4d}_{x_1[1] \neq y_1[1]} (q_3, \dots) - \text{противоречие}$, из q_3 нет переходов в q_1 .

(b) $q = q_2$. В q_2 есть переходы только из q_1 , в q_2 прочитываются a, b . $5a \Rightarrow w = xcy, |x| \neq 0, x, y \in \{a, b\}^*$. При переходе в q_2 прочитывается символ, поэтому $|y| > 0$. Аналогично $5a$ выделим общую часть $x^R = \tau x_1, y = \tau y_1$. Аналогично $5a$ ($|x| > 0$) получаем $(q_0, w, Z) \vdash^* (q_1, \tau y_1, U_r(\tau x_1)Z) \boxed{=}$. Рассмотрим случаи:

i. $|\tau| = 0, |x_1| = 0 \Rightarrow |x| = 0 - \text{противоречие}$

ii. $|\tau| > 0, |x_1| = 0 \Rightarrow \tau = \tau_0 \sigma, \sigma \in \{a, b\}$. $\boxed{=}(q_1, \tau_0 \sigma y_1, \tau_0 U_r(\sigma)Z) \vdash^* (q_1, \sigma y_1, U_r(\sigma)Z) \vdash^{4f} (q_3, \dots) - \text{противоречие}$, из q_3 нет переходов в q_2 .

iii. $|\tau| \geq 0, |x_1| > 0$. Тогда $\boxed{=}(q_1, \tau y_1, \tau U_r(x_1)Z) \vdash^* (q_1, y_1, U_r(x_1)Z) \boxed{=1}$.

a. $|y_1| = 0 \Rightarrow \boxed{=1}(q_1, \varepsilon, U_r(x_1))$. В $5a$ было показано, что автомат остановится в $q_1 - \text{противоречие}$.

b. $|y_1| > 0$. Тогда $x_1[1] \neq y_1[1]$. Обозначим $x_1 = \sigma_1 x_1^0, y_1 = \sigma_2 y_1^0$, и $\boxed{=1}(q_1, \sigma_1 y_1^0, U_r(\sigma_2 x_1^0)Z) \vdash^{4d}_{x_1[1] \neq y_1[1]} (q_3, y_1^0, U_r(x_1^0)Z) \vdash^* (q_3, \varepsilon, U_r(x_1^0)Z)$ (последние переходы возможны только при $x_1^0 \neq \varepsilon$).

Получаем $x_1 \neq y_1 \Rightarrow x^R \neq y \Rightarrow w \in L$.

(c) $q = q_4$. В q_4 прочитываются a, b ; в q_4 есть переходы только из $q_3 \xrightarrow{a, Z/Z, b, Z/Z} q_4$, в q_3 есть переходы только из $p \in \{q_0, q_1\}$. Рассмотрим случаи:

i. $p = q_0$. Если в q_0 были прочитаны символы из $\{a, b\}$, то на верхушке стека не $Z \Rightarrow$ переход в q_3 не мог быть совершен. Получаем, что $w = cy, y \in \{a, b\}^*$. Но при переходе в q_4 из q_3 прочитывается хотя бы один символ, поэтому $|y| > 0 \Rightarrow w \in L$

ii. $p = q_1$. $5b \Rightarrow w = xcy, |x| > 0, x, y \in \{a, b\}^*$. Аналогично $5b$ разобьем $x^R = \tau x_1, y = \tau y_1$, рассмотрим случаи:

A. $|\tau| = 0, |x_1| = 0 \Rightarrow |x| = 0 - \text{противоречие}$

B. $|\tau| > 0, |x_1| = 0 \Rightarrow \tau = \tau_0 \sigma, \sigma \in \{a, b\}$. Аналогично $5b$ получим $(q_0, w, Z) \vdash^* (q_1, \tau_0 \sigma y_1, \tau_0 U_r(\sigma)Z) \vdash^{4c} (q_1, \sigma y_1, U_r(\sigma)Z) \vdash^{4f} (q_3, y_1, Z)$. При переходе из q_3 в q_4 был прочитан символ, поэтому $|y_1| > 0$. Имеем $x^R \equiv \tau x_1 \equiv \tau \neq \tau y_1 \equiv y \Rightarrow w \in L$.

C. $|\tau| \geq 0, |x_1| > 0$.

a. $|y_1| = 0$. В $5a$ было показано, что автомат остановится в $q_1 - \text{противоречие}$.

b. $|y_1| > 0$. В $5b$ было показано, что автомат остановится в $q_2 - \text{противоречие}$.

Задача 2

Задача 3

$\Sigma \stackrel{\text{def}}{=} \{a, b\}, \Gamma \stackrel{\text{def}}{=} (N, \Sigma, P, S). N \stackrel{\text{def}}{=} \{A, B, C, D, E, F, G\} P :$

$S \rightarrow A|B|C|E|AG$

$A \rightarrow C|aABC|\varepsilon$

$B \rightarrow bABa|aCbDaGb|\varepsilon$

$C \rightarrow BaAbC|aGD|\varepsilon$

$F \rightarrow aBaaCbA|aGE$

$E \rightarrow A$

1. Удалим бесплодные символы (для упрощения):

(a) $V_0 \stackrel{\text{def}}{=} \{a, b\}$

(b) $V_1 = V_0 \cup \{A, B, C\} = \{a, b, A, B, C\}$

(c) $V_2 = V_1 \cup \{S, F, E\} = \{a, b, S, A, B, C, F, E\}$

(d) $V_3 = V_2 \cup \emptyset$

Тогда $V_3 \setminus \Sigma = \{S, A, B, C, F, E\}$. Удалим нетерминалы $N \setminus V_3 = \{D, G\}$ и правила, их содержащие: $N' \stackrel{\text{def}}{=} N \setminus V_3 = \{S, A, B, C, F, E\}, P' :$

$S \rightarrow A|B|C|E|AG$

$A \rightarrow C|aABC|\varepsilon$

$B \rightarrow bABa|aCbDaGb|\varepsilon$

$C \rightarrow BaAbC|aGD|\varepsilon$

$F \rightarrow aBaaCbA|aGE$

$E \rightarrow A$

2. Удалим недостижимые символы (для упрощения):

(a) $V_0 \stackrel{\text{def}}{=} \{S\}$

(b) $V_1 = V_0 \cup \{A, B, C, E\}$

(c) $V_2 = V_1 \cup \emptyset$

$$\begin{aligned}
S &\rightarrow A|B|C|E|AG \\
A &\rightarrow C|aABC|\varepsilon \\
B &\rightarrow bABa|aCbDaGb|\varepsilon
\end{aligned}$$

$$\begin{aligned}
C &\rightarrow BaAbC|aGD|\varepsilon \\
F &\rightarrow aBaaCbA|aGE \\
E &\rightarrow A
\end{aligned}$$

1,2. Имеем P'' :

$$\begin{aligned}
S &\rightarrow A|B|C|E \\
A &\rightarrow C|aABC|\varepsilon \\
B &\rightarrow bABa|\varepsilon
\end{aligned}$$

$$\begin{aligned}
C &\rightarrow BaAbC|\varepsilon \\
E &\rightarrow A
\end{aligned}$$

3. Удалим ε -правила:

- (a) A, B, C — ε -порождающие.
- (b) S, E — ε -порождающие ($S \rightarrow A, E \rightarrow A$)

Перепишем правила, содержащие ε -порождающие нетерминалы справа (2^k правил для каждого правила, содержащего k ε -порождающих нетерминалов). P''' :

$$\begin{aligned}
S &\rightarrow A|B|C|E \\
A &\rightarrow C|a|aC|aB|aBC|aA|aAC|aAB|aABC \\
B &\rightarrow ba|bBa|bAa|bABa
\end{aligned}$$

$$\begin{aligned}
C &\rightarrow ab|abC|aAb|aAbC|Bab|BabC|BaAbC \\
E &\rightarrow A
\end{aligned}$$

Грамматика с такими правилами порождает язык $L(\Gamma) \setminus \{\varepsilon\}$.

4. Найдем цепные пары (множества пар соответствуют добавлениям на шагах алгоритма):

- (a) $(S, S), (A, A), (B, B), (C, C), (E, E)$
- (b) $(S, A), (S, B), (S, C), (S, E); (A, C); (E, A)$
- (c) $(S, C); (\cancel{S, A}); (E, C)$

5. Выпишем новое множество правил P'''' :

Цепная пара	Правила
(S, S)	\emptyset
(A, A)	$A \rightarrow a aC aB aBC aA aAC aAB aABC$
(B, B)	$B \rightarrow ba bBa bAa bABa$
(C, C)	$C \rightarrow ab abC aAb aAbC Bab BabC BaAbC$
(E, E)	\emptyset
(S, A)	$S \rightarrow a aC aB aBC aA aAC aAB aABC$
(S, B)	$S \rightarrow ba bBa bAa bABa$
(S, C)	$S \rightarrow ab abC aAb aAbC Bab BabC BaAbC$
(S, E)	\emptyset
(A, C)	$A \rightarrow ab abC aAb aAbC Bab BabC BaAbC$
(E, A)	$E \rightarrow a aC aB aBC aA aAC aAB aABC$
(S, C)	$S \rightarrow ab abC aAb aAbC Bab BabC BaAbC$
(E, C)	$E \rightarrow ab abC aAb aAbC Bab BabC BaAbC$

6. Нетерминалы A, B, C, E, S не являются бесплодными: $A \rightarrow a, B \rightarrow ba, C \rightarrow ab, E \rightarrow a, S \rightarrow ab$.

7. Удалим недостижимые:

- (a) $V_0 \stackrel{\text{def}}{=} \{S\}$
- (b) $V_1 \stackrel{\text{def}}{=} \{S, A, B, C\}$
- (c) $V_2 = V_1$

Удаляем E . $P^{(5)}$:

$$\begin{aligned}
A &\rightarrow a|aC|aB|aBC|aA|aAC|aAB|aABC|ab|abC|aAb|aAbC|Bab|BabC|BaAbC \\
B &\rightarrow ba|bBa|bAa|bABa \\
C &\rightarrow ab|abC|aAb|aAbC|Bab|BabC|BaAbC \\
S &\rightarrow a|aC|aB|aBC|aA|aAC|aAB|aABC|ba|bBa|bAa|bABa|ab|abC|aAb|aAbC|Bab|BabC|BaAbC
\end{aligned}$$

8. Приведем к нормальной форме Хомского. Добавим нетерминалы $A', B', A' \rightarrow a, B' \rightarrow b$. Заменим в правилах a на A' , b на B' . Подчеркнем слова из нетерминалов длины 2 в правых частях правил, которые заменим на новые нетерминалы:

$$\begin{aligned}
A &\rightarrow a|A'C|A'B|A'BC|A'A|A'AC|A'AB|A'AB|A'BC|A'B'|A'B'C|A'AB'|A'A'B'C|BA'B'|BA'B'C|BA'AB'C \\
B &\rightarrow B'A'|B'BA'|B'AA'|B'ABA' \\
C &\rightarrow A'B'|A'B'C|A'AB'|A'AB'|A'AB'|BA'B'|BA'B'C|BA'AB'C \\
S &\rightarrow a|A'C|A'B|A'BC|A'A|A'AC|A'AB|A'AB|A'BC|B'A'|B'BA'| \\
S &\rightarrow B'AA'|B'ABA'|A'B'|A'B'C|A'AB'|A'A'B'C|BA'B'|BA'B'C|BA'AB'C
\end{aligned}$$

$$A' \rightarrow a$$

$$B' \rightarrow b$$

Заменим подчеркнутые слова на новые нетерминалы, получим

Ответ:

$$A \rightarrow a|A'C|A'B|X_0C|A'A|X_1C|X_1B|X_1X_2|A'B'|X_3C|X_1B'|X_1X_4|X_5B'|X_5X_4|X_9C$$

$$B \rightarrow B'A'|X_7A'|X_8A'|X_8X_5$$

$$C \rightarrow A'B'|X_3C|X_1B'|X_1X_4|X_5B'|X_5X_4|X_9C$$

$$S \rightarrow a|A'C|A'B|X_0C|A'A|X_1C|X_1B|X_1X_2|B'A'|X_7A'|X_8A'|X_8X_5|A'B'|X_3C|X_1B'|X_1X_4|X_5B'|X_5X_4|X_9C$$

$$A' \rightarrow a$$

$$X_2 \rightarrow BC$$

$$X_6 \rightarrow AB'$$

$$B' \rightarrow b$$

$$X_3 \rightarrow A'B'$$

$$X_7 \rightarrow B'B$$

$$X_0 \rightarrow A'B$$

$$X_4 \rightarrow B'C$$

$$X_8 \rightarrow B'A$$

$$X_1 \rightarrow A'A$$

$$X_5 \rightarrow BA'$$

$$X_9 \rightarrow X_5X_6$$

Задача 4

Задача 5

$\Sigma_2 \stackrel{\text{def}}{=} \{[1], [2]\}$, $\bar{\Sigma}_2 \stackrel{\text{def}}{=} \{[1], [2]\}$. $D_2 \stackrel{\text{def}}{=}$ язык ПСП над $\Sigma \stackrel{\text{def}}{=} \Sigma_2 \cup \bar{\Sigma}_2$. $\Delta \stackrel{\text{def}}{=} \{a, b\}$. $\varphi: \Sigma^* \rightarrow \Delta^*$, $\varphi([1]) \stackrel{\text{def}}{=} a$, $\varphi([2]) \stackrel{\text{def}}{=} b$, $\varphi([1]) \stackrel{\text{def}}{=} b$, $\varphi([2]) \stackrel{\text{def}}{=} a$. Доопределим φ до морфизма (см. решение упр. 2 из задания 3). $L \stackrel{\text{def}}{=} \varphi(D_2 \cap \Sigma^*) \equiv \varphi(D_2)$. $L' \stackrel{\text{def}}{=} \{w \in \Delta^* \mid |w|_a = |w|_b\}$.

1. Докажем, что $L \subseteq L'$. Пусть $y \in L \equiv \varphi(D_2)$. Тогда $\exists x \in D_2: y = \varphi(x)$. x — ПСП $\Rightarrow \forall i \in \bar{1}, \bar{2} \hookrightarrow |x|_i = |x|_{\bar{i}}$. Сложим равенства, получим: $|x|_{[1]} + |x|_{[2]} = |x|_{[1]} + |x|_{[2]}$. Пусть $x = x_1 \dots x_m$, $\forall i \in \bar{1}, \bar{m} \hookrightarrow x_i \in \Sigma$. Тогда $y = \varphi(x) = \varphi(x_1) \dots \varphi(x_m) = y_1 \dots y_m$, $\forall i \in \bar{1}, \bar{m} \hookrightarrow y_i = \varphi(x_i) \in \Delta$. Но из определения φ имеем $[1], [2] \xrightarrow{\varphi} a; [1], [2] \xrightarrow{\varphi} b$. Тогда $|y|_a = |x|_{[1]} + |x|_{[2]} \equiv |x|_{[1]} + |x|_{[2]} = |y|_b \Rightarrow y \in L'$ ■

2. Докажем, что $L' \subseteq L$ индукцией по длине $y \in L'$: $P(n) \stackrel{\text{def}}{=} [\forall y \in L': |y| \leq n \hookrightarrow y \in L]$.

Заметим, что $y \in L \Leftrightarrow y \in \varphi(D_2) \Leftrightarrow \varphi^{-1}(y) \cap D_2 \neq \emptyset$. Поэтому будем искать прообраз слова y , принадлежащий D_2 .

(а) $n = 0 \Rightarrow |y| = 0 \Rightarrow y = \varepsilon \in L'$. Пусть $x \stackrel{\text{def}}{=} \varepsilon \in D_2$ (так как пустое слово — ПСП). Тогда $y = \varepsilon \equiv \varphi(x) \Rightarrow y \in \varphi(D_2) \equiv L \Rightarrow P(0)$

(б) Фиксируем $n > 0$. Пусть $P(n-1)$. Пусть $y \in L': |y| = n$. Поскольку $|y| = n > 0$, и $|y|$ — чётно (см. решение задачи 3 из задания 6), то $|y| \geq 2$. Рассмотрим первый и последний символы σ_l и σ_r слова $y \equiv \sigma_l y_1 \sigma_r$:

- $\sigma_l = a, \sigma_r = b$. Тогда $y = ay_1b$. $|y_1| = n - 2 \leq n - 1 \xrightarrow{P(n-1)} \exists x_1 \in D_2: \varphi(x_1) = y_1$. Определим $x = [1x_1]_1$. $x_1 \in D_2 \Rightarrow x_1$ — ПСП $\Rightarrow x$ — ПСП, так как получен из ПСП добавлением скобок типа 1 слева и справа $\Rightarrow x \in D_2$. Но $\varphi(x) \equiv \varphi([1x_1]_1) = \varphi([1])\varphi(x_1)\varphi([1]) = ay_1b \equiv y$. Получаем $\varphi^{-1}(y) \cap D_2 \ni x \Rightarrow \varphi^{-1}(y) \cap D_2 \neq \emptyset$.
- $\sigma_l = b, \sigma_r = b$. Тогда $y = by_1a$. $|y_1| = n - 2 \leq n - 1 \xrightarrow{P(n-1)} \exists x_1 \in D_2: \varphi(x_1) = y_1$. Определим $x = [2x_1]_2$. $x_1 \in D_2 \Rightarrow x_1$ — ПСП $\Rightarrow x$ — ПСП, так как получен из ПСП добавлением скобок типа 2 слева и справа $\Rightarrow x \in D_2$. Но $\varphi(x) \equiv \varphi([2x_1]_2) = \varphi([2])\varphi(x_1)\varphi([2]) = by_1a \equiv y$. Получаем $\varphi^{-1}(y) \cap D_2 \ni x \Rightarrow \varphi^{-1}(y) \cap D_2 \neq \emptyset$.
- $\sigma_l = \sigma_r$. Тогда $y = \sigma y_1 \sigma \in L'$. Воспользуемся утверждением в рамочке из решения задачи 3 задания 6:

$$y = \sigma y_1 \sigma \in L' \Rightarrow \exists y_l, y_r: y = y_l y_r, |y_l|, |y_r| \in \bar{1}, |y| - 2, y_l, y_r \in L'$$

Но $|y_l|, |y_r| \leq |y| - 2 = n - 2 \leq n - 1 \xrightarrow{P(n-1)} \exists x_l, x_r \in D_2: y_l = \varphi(x_l), y_r = \varphi(x_r)$. Определим $x \stackrel{\text{def}}{=} x_l x_r$. Тогда $x \in D_2$ (конкатенация ПСП — ПСП), и $\varphi(x) = \varphi(x_l x_r) = \varphi(x_l)\varphi(x_r) = y_l y_r = y \Rightarrow \varphi^{-1}(y) \cap D_2 \ni x \Rightarrow \varphi^{-1}(y) \cap D_2 \neq \emptyset$

■

Ответ: Верно, что $L = \{w \mid |w|_a = |w|_b\}$.

Задача 6

Автомат $\mathcal{A} = (\Sigma, \Gamma, Q, q_0, Z_0, \delta, \emptyset)$ из 7-го задания:

1. $\Sigma = \{a, b\}$

2. $\Gamma = \{a, Z_0\}$

3. $Q = \{q_0, q_1\}$

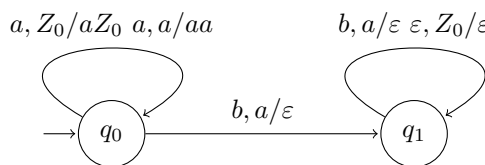
4. δ изображена справа

Определим грамматику $G = (N, \Sigma, P, S)$. $N = \{S\} \cup \{[qZp] \mid q, p \in Q, Z \in \Gamma\}$

1. Добавим правила $S \rightarrow [q_0 Z_0 q_0] [q_0 Z_0 q_1]$

2. Рассмотрим переходы из δ , добавим правила

$$(a) \delta \ni q_0 \xrightarrow{a, Z_0/aZ_0} q_0: [q_0 Z_0 q_0] \rightarrow a[q_0 a q_0][q_0 Z_0 q_0] [a[q_0 a q_1][q_1 Z_0 q_0], [q_0 Z_0 q_1] \rightarrow a[q_0 a q_0][q_0 Z_0 q_1] [a[q_0 a q_1][q_1 Z_0 q_1]$$



- (b) $\delta \ni q_0 \xrightarrow{a, a/aa} q_0: [q_0 a q_0] \rightarrow a[q_0 a q_0][q_0 a q_0] \mid a[q_0 a q_1][q_1 a q_0], [q_0 a q_1] \rightarrow a[q_0 a q_0][q_0 a q_1] \mid a[q_0 a q_1][q_1 a q_1]$
- (c) $\delta \ni q_0 \xrightarrow{b, a/\varepsilon} q_1: [q_0 a q_1] \rightarrow b$
- (d) $\delta \ni q_1 \xrightarrow{b, a/\varepsilon} q_1: [q_1 a q_1] \rightarrow b$
- (e) $\delta \ni q_0 \xrightarrow{\varepsilon, Z_0/\varepsilon} q_1: [q_1 Z_0 q_1] \rightarrow \varepsilon$

3. Удалим бесплодные нетерминалы:

- (a) $V_0 = \{a, b\}$
- (b) $V_1 = V_0 \cup \{[q_0 a q_1], [q_1 a q_1], [q_1 Z_0 q_1]\}$
- (c) $V_2 = V_1 \cup \{[q_0 Z_0 q_1]\}$
- (d) $V_3 = V_2 \cup \{S\}$
- (e) $V_4 = V_3$.

Имеем правила $S \rightarrow [q_0 Z_0 q_1]$, $[q_0 Z_0 q_1] \rightarrow a[q_0 a q_1][q_1 Z_0 q_1]$, $[q_0 a q_1] \rightarrow a[q_0 a q_1][q_1 a q_1] \mid b$, $[q_1 a q_1] \rightarrow b$, $[q_1 Z_0 q_1] \rightarrow \varepsilon$

4. Удалим недостижимые нетерминалы:

- (a) $V_0 = \{S\}$
- (b) $V_1 = V_0 \cup \{[q_0 Z_0 q_1]\}$
- (c) $V_2 = V_1 \cup \{[q_0 a q_1], [q_1 Z_0 q_1]\}$
- (d) $V_3 = V_2 \cup \{[q_1 a q_1]\}$
- (e) $V_4 = V_3$

(все достижимы)

5. Переобозначим:

$$S \rightarrow \underbrace{[q_0 Z_0 q_1]}_A, \underbrace{[q_0 Z_0 q_1]}_A \rightarrow a \underbrace{[q_0 a q_1]}_B \underbrace{[q_1 Z_0 q_1]}_C, \underbrace{[q_0 a q_1]}_B \rightarrow a \underbrace{[q_0 a q_1]}_B \underbrace{[q_1 a q_1]}_D \mid b, \underbrace{[q_1 a q_1]}_D \rightarrow b, \underbrace{[q_1 Z_0 q_1]}_C \rightarrow \varepsilon,$$

получим

$$S \rightarrow A, A \rightarrow aBC, B \rightarrow aBD \mid b, D \rightarrow b, C \rightarrow \varepsilon$$

6. Из D, C есть только правила $D \rightarrow b, C \rightarrow \varepsilon$, поэтому они раскрываются единственным образом. Уберем их, получим грамматику G' , причем G' — однозначная $\Leftrightarrow G$ — однозначная:

$$S \rightarrow A, A \rightarrow aB, B \rightarrow aBb \mid b$$

Аналогично для $S \rightarrow A$ (раскрывается единственным образом). Получим G'' : G'' — однозначная $\Leftrightarrow G'$ — однозначная:

$$S \rightarrow aB, B \rightarrow aBb \mid b$$

После применения правила $B \rightarrow b$ нельзя применить правило $B \rightarrow aBb$, и каждое применение $B \rightarrow aBb$ увеличивает количество символов a и b на 1. Поэтому количество его применений фиксировано для каждого $w \in L(G'')$. Отсюда получаем, что грамматика G'' — однозначная $\Rightarrow G'$ — однозначная $\Rightarrow G$ — однозначная.

Ответ:

1. $G: S \rightarrow \underbrace{[q_0 Z_0 q_1]}_A, \underbrace{[q_0 Z_0 q_1]}_A \rightarrow a \underbrace{[q_0 a q_1]}_B \underbrace{[q_1 Z_0 q_1]}_C, \underbrace{[q_0 a q_1]}_B \rightarrow a \underbrace{[q_0 a q_1]}_B \underbrace{[q_1 a q_1]}_D \mid b, \underbrace{[q_1 a q_1]}_D \rightarrow b, \underbrace{[q_1 Z_0 q_1]}_C \rightarrow \varepsilon$
2. После переобозначения $S \rightarrow A, A \rightarrow aBC, B \rightarrow aBD \mid b, D \rightarrow b, C \rightarrow \varepsilon$
3. G — однозначная.