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

1. $\forall w\colon S \underset{\text{вывод}}{\overset{\text{левый}}{\Rightarrow^*}} w$, $w \in (\Sigma \cup N)^*$, $w \notin \Sigma^* \hookrightarrow w = u[r_0Y_1r_1][r_1Y_2r_2]...[r_{n-1}Y_nr_n]$, $u \in \Sigma^*$, $r_i \in Q$, $Y_i \in \Gamma$ — доказывается индукцией по длине левого вывода из свойств добавленных правил (слева всегда, возможно, нетерминалы, затем, возможно, терминалы. $[qZp] \to ...[...Yp]$, поэтому соседние состояния, отделенные скобками совпадают: $...r][r...$).

2. $P(k) \overset{\text{def}}{=} \left[\forall n \, \forall w\colon S \underset{\text{k шагов}}{\overset{\text{левый вывод}}{\Rightarrow^*}} w \hookrightarrow w = u[r_0Y_1r_1]...[r_{n-1}Y_nr_n] \Rightarrow (q_0, u, Z_0) \vdash^* (r_0, \varepsilon, Y_1...Y_n)\right]$.

   (a) $n = 1$. Из определения P могут быть только правила $S \to [q_0Z_0p]$, и $(q_0, \varepsilon, Z_0) \equiv (q_0, \varepsilon, Z_0) \Rightarrow P(1)$

   (b) Пусть $P(k)$. Рассмотрим левый вывод длины $k+1$: $S \Rightarrow^* y \equiv u[r_0Y_1r_1][r_1Y_2r_2]...[r_{n-1}Y_nr_n]$. Пусть начальная часть этого вывода длины k имеет вид $S \Rightarrow^* x \equiv u_l[s_0Z_1s_1][s_1Z_2s_2]...[s_{m-1}Z_ms_m]$. На последнем, $k+1$-м шаге был раскрыт первый нетерминал $[s_0Z_1s_1] \to z$:

      i. $[s_0Z_1s_1] \to z \equiv u_r[t_0W_1t_1][t_1W_2t_2]...[t_{l-1}W_lt_l]$.
      Тогда $y = \underbrace{u_l}_{\text{префикс } x} \underbrace{u_r[t_0W_1t_1][t_1W_2t_2]...[t_{l-1}W_lt_l]}_{z} \underbrace{[s_1Z_2s_2]...[s_{m-1}Z_ms_m]}_{\text{суффикс } x}$.
      Отсюда $W_1...W_lZ_2...Z_m = Y_1...Y_n$, $u = u_lu_r$, $t_0 = r_0$.
      $P(k) \Rightarrow (q_0, u_l, Z_0) \vdash^* (s_0, \varepsilon, Z_1...Z_m)$. Применено правило $[s_0Z_1s_1] \to z \Rightarrow (s_0, u_r, Z_1) \vdash (t_0, \varepsilon, W_1...W_l)$.
      Тогда $(q_0, u, Z_0) \equiv (q_0, u_lu_r, Z_0) \vdash^* (s_0, u_r, Z_1...Z_m) \vdash (t_0, \varepsilon, W_1...W_lZ_2...Z_m) \equiv (r_0, \varepsilon, Y_1...Y_n)$.

      ii. $[s_0Z_1s_1] \to z \equiv u_r \in \Sigma^*$. Тогда $y = \underbrace{u_l}_{\text{префикс } x} \underbrace{u_r}_{z} \underbrace{[s_1Z_2s_2]...[s_{m-1}Z_ms_m]}_{\text{суффикс } x}$. Отсюда $Z_2...Z_m = Y_1...Y_n$, $s_1 = r_0$, $u = u_lu_r$. $[s_0Z_1s_1] \to u_r \in P \Rightarrow (s_0, u_r, Z_1) \vdash (s_1, \varepsilon, \varepsilon)$ — из определения P.
      $(q_0, u, Z_0) \equiv (q_0, u_lu_r, Z_0) \overset{P(k)}{\vdash^*} (s_0, u_r, Z_1...Z_m) \vdash (s_1, \varepsilon, Z_2...Z_m) \equiv (r_0, \varepsilon, Y_1...Y_n)$.

3. Пусть $\underline{w \in L(G)} \Rightarrow \exists n\colon S \underset{\text{n шагов}}{\overset{\text{левый вывод}}{\Rightarrow^*}} w \in \Sigma^*$. На последнем, $n-1$ шаге был раскрыт нетерминал $[qZp] \to w_r$, поэтому этот левый вывод имеет вид $S \Rightarrow^* w_l[qZp] \Rightarrow w_lw_r$. Имеем $w = w_lw_r$. $[qZp] \to w_r \in P \Rightarrow (q, w_r, Z) \to (p, \varepsilon, \varepsilon)$.
   $(q_0, w, Z_0) \equiv (q_0, w_lw_r, Z_0) \overset{P(n-1)}{\vdash^*} (q, w_r, Z) \vdash (p, \varepsilon, \varepsilon) \Rightarrow \underline{w \in L(M)}$ ■

Упражнение 3

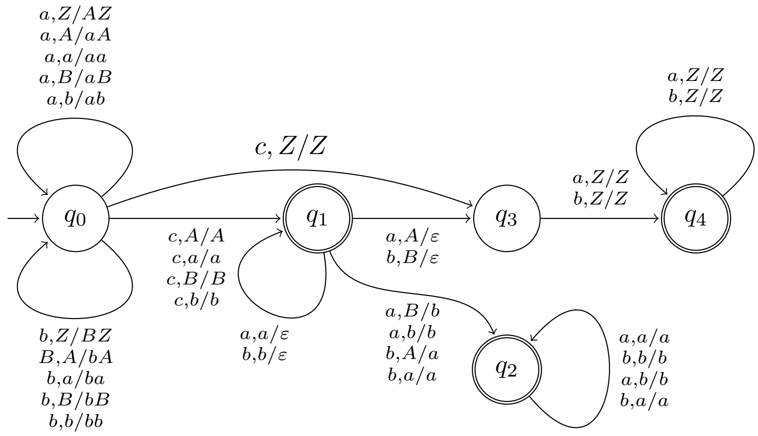
Упражнение 4

Задача 1

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

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

   1. $\Gamma \overset{\text{def}}{=} \{a, b, A, B, Z\}$
   2. $Q \overset{\text{def}}{=} \{q_0, q_1, q_2, q_3, q_4\}$
   3. δ изображена справа
   4. $F \overset{\text{def}}{=} \{q_1, q_2, q_4\}$



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

3. Определим $U\colon \{a,b\} \to \{A, B\}\colon U(a) = A$, $U(b) = B$. Определим $U_r\colon \{a,b\}^* \to \{a,b,A,B\}^*$:
$U_r(w) = \begin{cases} \varepsilon, & w = \varepsilon \\ w_1 ... w_{n-1} U(w_n), & w = w_1 ... w_n, \forall i \in \overline{1,n} \hookrightarrow w_i \in \{a,b\} \end{cases}$ — 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 $|w|$:
   $P(n) \overset{\text{def}}{=} \left[\forall w \in \{a,b\}^*\colon |w| = n \hookrightarrow (q_0, w, Z) \vdash^* (q_0, \varepsilon, U_r(w^R)Z)\right]$

   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 $n \geqslant 1$, пусть $\underline{P(n)}$. Пусть $w \in \{a,b\}^*, |w| = n+1$. Тогда $w = w_0\sigma, |w_0| = n > 0$.
   $P(n) \Rightarrow (q_0, w_0, Z) \vdash^* (q_0, \varepsilon, U_r(w_0^R)Z)$. Тогда $(q_0, w, Z) \equiv (q_0, w_0\sigma, Z) \vdash^* (q_0, \sigma, U_r(w_0^R)Z)$.
   $\measuredangle$ переходы из $(q_0, \sigma, U(w_0^R)Z)$. 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, $(q_0, \sigma, U_r(w_0^R)Z) \vdash (q_0, \varepsilon, \sigma U_r(w_0^R)Z) \overset{|w_0|>0}{=} (q_0, \varepsilon, U_r(w^R)Z) \Rightarrow \underline{P(n+1)}$.

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

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

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

   ii. Фиксируем $n \geqslant 0$. Пусть $\underline{P(n)}$. Пусть $x \in \{a,b\}^*\colon |x| = n+1 \Rightarrow x = x_0\sigma, |x_0| = n \overset{P(n)}{\Rightarrow} (q_1, x_0, x_0Z) \vdash^* (q_1, \varepsilon, Z)$. 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 $|x_1| > 0$. $(q_0, w, Z) \equiv (q_0, x_1^R \tau^R c \tau y_1, Z) \underset{|x_1|>0}{\overset{4a}{\vdash^*}} (q_0, c\tau y_1, U_r(\tau x_1)Z) \overset{4b}{\vdash} (q_1, \tau y_1, U_r(\tau x_1)Z) \overset{|x_1|>0}{\equiv}$
   $\equiv (q_1, \tau y_1, \tau U_r(x_1)Z) \overset{4c}{\vdash^*} (q_1, y_1, U_r(x_1)Z)$.

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

B. Пусть $|y_1| = 0$. Тогда $w = x_1^R\tau^R c\tau y_1 \equiv x_1^R\tau^R c\tau \Rightarrow (q_0, w, Z) \equiv (q_0, x_1^R\tau^R c\tau, Z) \overset{4a}{\underset{|x_1|>0}{\vdash^*}} (q_0, c\tau, \tau U_r(x_1)Z) \overset{4b}{\underset{|x_1|>0}{\vdash}} (q_1, \tau, \tau U_r(x_1)Z) \overset{4e}{\vdash^*} (q_1, \varepsilon, U_r(x_1)Z)$. $q_1 \in F \Rightarrow \underline{w \in L(\mathcal{A})}$

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

A. $|\tau| > 0 \Rightarrow \tau = \tau_0\sigma$, $\sigma \in \{a,b\}$. Получаем $(q_0, w, Z) \equiv (q_0, \tau^R c\tau y_1, Z) \overset{4a}{\underset{|\tau|>0}{\vdash^*}} (q_0, c\tau y_1, U_r(\tau)Z) \overset{4b}{\underset{|\tau|>0}{\vdash}} (q_1, \tau y_1, U_r(\tau)Z) \equiv (q_1, \tau_0\sigma y_1, \tau_0 U_r(\sigma)Z) \overset{4c}{\vdash^*} (q_1, \sigma y_1, U_r(\sigma)Z) \overset{4f}{\vdash} (q_3, y_1, Z) \equiv (q_3, \varkappa\Psi, Z) \overset{4g}{\vdash} (q_4, \Psi, Z) \overset{4h}{\vdash^*} (q_4, \varepsilon, Z)$. $q_4 \in F \Rightarrow \underline{w \in L(\mathcal{A})}$

B. $|\tau| = 0 \Rightarrow w = x_1^R\tau^R c\tau y_1 \equiv cy_1 \Rightarrow (q_0, w, Z) \equiv (q_0, cy_1, Z) \overset{4i}{\vdash} (q_3, y_1, Z) \equiv (q_3, \varkappa\Psi, Z) \overset{4g}{\vdash} (q_4, \Psi, Z) \overset{4h}{\vdash^*} (q_4, \varepsilon, Z)$. $q_4 \in F \Rightarrow \underline{w \in L(\mathcal{A})}$

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 $(q_0, w, Z) \equiv (q_0, xcy, Z) \overset{4a}{\underset{|x|>0}{\vdash}} (q_0, cy, U_r(x^R)Z) \overset{4b}{\underset{|x|>0}{\vdash}} (q_1, y, U_r(x^R)Z)\boxed{=}$. 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. $|\tau| = 0, |x_1| = 0 \Rightarrow |x| = 0$ — <u>противоречие</u>

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

iii. $|\tau| \geqslant 0, |x_1| > 0$. Тогда $\boxed{=}(q_1, \tau y_1, \tau U_r(x_1)Z) \overset{4c}{\vdash^*} (q_1, y_1, U_r(x_1)Z)\boxed{=_1}$.

a. $|y_1| = 0 \Rightarrow \boxed{=_1}(q_1, \varepsilon, U_r(x_1))$. Тогда $w = \underbrace{x_1^R\tau^R}_{x} c \underbrace{\tau y_1^R}_{y}$, $x^R = \tau x_1 \underset{|x_1|>0}{\neq} \tau = y \Rightarrow \underline{w \in L}$.

b. $|y_1| > 0$. Тогда $x_1[1] \neq y_1[1]$, и $\boxed{=_1}(q_1, y_1, U_r(x_1)Z) \overset{4d}{\underset{x_1[1]\neq y_1[1]}{\vdash}} (q_3, \ldots)$ — <u>противоречие</u>, 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. $|\tau| = 0, |x_1| = 0 \Rightarrow |x| = 0$ — <u>противоречие</u>

ii. $|\tau| > 0, |x_1| = 0 \Rightarrow \tau = \tau_0\sigma$, $\sigma \in \{a,b\}$. $\boxed{=}(q_1, \tau_0\sigma y_1, \tau_0 U_r(\sigma)Z) \overset{4c}{\vdash^*} (q_1, \sigma y_1, U_r(\sigma)Z) \overset{4f}{\vdash} (q_3, \ldots)$ — <u>противоречие</u>, из q_3 нет переходов в q_2.

iii. $|\tau| \geqslant 0, |x_1| > 0$. Тогда $\boxed{=}(q_1, \tau y_1, \tau U_r(x_1)Z) \overset{4c}{\vdash^*} (q_1, y_1, U_r(x_1)Z)\boxed{=_1}$.

a. $|y_1| = 0 \Rightarrow \boxed{=_1}(q_1, \varepsilon, U_r(x_1))$. В $5a$ было показано, что автомат остановится в q_1 — <u>противоречие</u>.

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 $\boxed{=_1}(q_1, \sigma_1 y_1^0, U_r(\sigma_2 x_1^0)Z) \overset{4d}{\underset{x_1[1]\neq y_1[1]}{\vdash}} (q_3, y_1^0, U_r(x_1^0)Z) \overset{4e}{\vdash^*} (q_3, \varepsilon, U_r(x_1^0)Z)$ *(последние переходы возможны только при $x_1^0 \neq \varepsilon$)*.
Получаем $x_1 \neq y_1 \Rightarrow x^R \neq y \Rightarrow \underline{w \in L}$.

(c) $q = q_4$. В q_4 прочитываются a, b; в q_4 есть переходы только из $q_3 \xrightarrow{\substack{a,Z/Z \\ b,Z/Z}} q_4$, 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 $|y| > 0 \Rightarrow \underline{w \in L}$

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. $|\tau| = 0, |x_1| = 0 \Rightarrow |x| = 0$ — <u>противоречие</u>

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

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

a. $|y_1| = 0$. В $5a$ было показано, что автомат остановится в q_1 — <u>противоречие</u>.

b. $|y_1| > 0$. В $5b$ было показано, что автомат остановится в q_2 — <u>противоречие</u>.

Задача 2

Схема доказательства:$L \in \mathsf{CFL} \Rightarrow \exists n, \varphi, R \colon R \in \mathsf{REG}$, φ — морфизм: $\Sigma^* \longrightarrow \Delta^*$, $L = \varphi(D_n \cap R)$. Из доказательства теоремы XI
(не дописано).

Задача 3

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

$$S \to A|B|C|E|AG$$
$$A \to C|aABC|\varepsilon$$
$$B \to bABa|aCbDaGb|\varepsilon$$
$$C \to BaAbC|aGD|\varepsilon$$
$$F \to aBaaCbA|aGE$$
$$E \to A$$

1. Удалим бесплодные символы (для упрощения):

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

   (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) $V_3 = V_2 \cup \varnothing$

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

$S \to A|B|C|E|\cancel{AG}$
$A \to C|aABC|\varepsilon$
$B \to bABa|\cancel{aCbDaGb}|\varepsilon$

$C \to BaAbC|\cancel{aGD}|\varepsilon$
$F \to aBaaCbA|\cancel{aGE}$
$E \to A$

2. Удалим недостижимые символы (для упрощения):

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

   $N'' \stackrel{\text{def}}{=} \{A, B, C, E, S\}$, P'':

   $S \to A|B|C|E|\cancel{AG}$
   $A \to C|aABC|\varepsilon$
   $B \to bABa|\cancel{aCbDaGb}|\varepsilon$

   $C \to BaAbC|\cancel{aGD}|\varepsilon$
   $F \to \cancel{aBaaCbA|aGE}$
   $E \to A$

1,2. Имеем P'':

   $S \to A|B|C|E$
   $A \to C|aABC|\varepsilon$
   $B \to bABa|\varepsilon$

   $C \to BaAbC|\varepsilon$
   $E \to A$

3. Удалим ε-правила:

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

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

   $S \to A|B|C|E$
   $A \to C|a|aC|aB|aBC|aA|aAC|aAB|aABC$
   $B \to ba|bBa|bAa|bABa$

   $C \to ab|abC|aAb|aAbC|Bab|BabC|BaAbC$
   $E \to A$

   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)	$\varnothing$
(A, A)	$A \to a\|aC\|aB\|aBC\|aA\|aAC\|aAB\|aABC$
(B, B)	$B \to ba\|bBa\|bAa\|bABa$
(C, C)	$C \to ab\|abC\|aAb\|aAbC\|Bab\|BabC\|BaAbC$
(E, E)	$\varnothing$
(S, A)	$S \to a\|aC\|aB\|aBC\|aA\|aAC\|aAB\|aABC$
(S, B)	$S \to ba\|bBa\|bAa\|bABa$
(S, C)	$S \to ab\|abC\|aAb\|aAbC\|Bab\|BabC\|BaAbC$
(S, E)	$\varnothing$
(A, C)	$A \to ab\|abC\|aAb\|aAbC\|Bab\|BabC\|BaAbC$
(E, A)	$E \to a\|aC\|aB\|aBC\|aA\|aAC\|aAB\|aABC$
(S, C)	$S \to ab\|abC\|aAb\|aAbC\|Bab\|BabC\|BaAbC$
(E, C)	$E \to ab\|abC\|aAb\|aAbC\|Bab\|BabC\|BaAbC$

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

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. $P^{(5)}$:
   $A \to a|aC|aB|aBC|aA|aAC|aAB|aABC|ab|abC|aAb|aAbC|Bab|BabC|BaAbC$
   $B \to ba|bBa|bAa|bABa$
   $C \to ab|abC|aAb|aAbC|Bab|BabC|BaAbC$

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

   $A \to a|A'C|A'B|\underline{A'B}C|A'A|\underline{A'A}C|\underline{A'A}B|\underline{A'A}\,\underline{BC}|A'B'|\underline{A'B'}C|\underline{A'A}B'|\underline{A'A}\,\underline{B'C}|\underline{BA'}B'|\underline{BA'}\,\underline{B'C}|\underline{BA'}\,\underline{AB'}C$

   $B \to B'A'|\underline{B'B}A'|\underline{B'A}A'|\underline{B'A}\,\underline{BA'}$

   $C \to A'B'|\underline{A'B'}C|\underline{A'A}B'|\underline{A'A}\,\underline{B'C}|\underline{BA'}B'|\underline{BA'}\,\underline{B'C}|\underline{BA'}\,\underline{AB'}\,C$

   $S \to a|A'C|A'B|\underline{A'B}C|A'A|\underline{A'A}C|\underline{A'A}B|\underline{A'A}\,\underline{BC}|B'A'|\underline{B'B}A'$

   $S \to \underline{B'A}A'|\underline{B'A}\,\underline{BA'}|A'B'|\underline{A'B'}\,C|\underline{A'A}B'|\underline{A'A}\,\underline{B'C}|\underline{BA'}B'|\underline{BA'}\,\underline{B'C}|\underline{BA'}\,\underline{AB'}C$

   $A' \to a$

   $B' \to b$

   Заменим подчеркнутые слова на новые нетерминалы, получим

Ответ:

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

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

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

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

$A' \to a$

$B' \to b$

$X_0 \to A'B$

$X_1 \to A'A$

$X_2 \to BC$

$X_3 \to A'B'$

$X_4 \to B'C$

$X_5 \to BA'$

$X_6 \to AB'$

$X_7 \to B'B$

$X_8 \to B'A$

$X_9 \to X_5X_6$

Задача 4

Задача 5

$\Sigma_2 \overset{\text{def}}{=} \{[_1, [_2\}$, $\overline{\Sigma_2} \overset{\text{def}}{=} \{]_1,]_2\}$. $D_2 \overset{\text{def}}{-}$ язык ПСП над $\Sigma \overset{\text{def}}{=} \Sigma_2 \cup \overline{\Sigma_2}$. $\Delta \overset{\text{def}}{=} \{a, b\}$. $\varphi\colon \Sigma^* \longrightarrow \Delta^*$, $\varphi([_1) \overset{\text{def}}{=} a$, $\varphi([_2) \overset{\text{def}}{=} b$, $\varphi(]_1) \overset{\text{def}}{=} b$, $\varphi(]_2) \overset{\text{def}}{=} a$. Доопределим φ до морфизма (см. решение упр. 2 из задания 3). $L \overset{\text{def}}{=} \varphi(D_2 \cap \Sigma^*) \equiv \varphi(D_2)$. $L' \overset{\text{def}}{=} \{w \in \Delta^* \big| |w|_a = |w|_b\}$.

1. Докажем, что $L \subseteq L'$. Пусть $\underline{y \in L} \equiv \varphi(D_2)$. Тогда $\exists x \in D_2\colon y = \varphi(x)$. x — ПСП $\Rightarrow \forall i \in \overline{1,2} \hookrightarrow |x|_{[_i} = |x|_{]_i}$. Сложим равенства, получим: $|x|_{[_1} + |x|_{]_2} = |x|_{]_1} + |x|_{[_2}$. Пусть $x = x_1 ... x_m$, $\forall i \in \overline{1,m} \hookrightarrow x_i \in \Sigma$. Тогда $y = \varphi(x) = \varphi(x_1)...\varphi(x_m) = y_1...y_m$, $\forall i \in \overline{1,m} \hookrightarrow y_i = \varphi(x_i) \in \Delta$. Но из определения φ имеем $[_1,]_2 \overset{\varphi}{\to} a$; $]_1, [_2 \overset{\varphi}{\to} b$. Тогда $|y|_a = |x|_{[_1} + |x|_{]_2} \equiv |x|_{]_1} + |x|_{[_2} = |y|_b \Rightarrow \underline{y \in L'}$ ■

2. Докажем, что $L' \subseteq L$ индукцией по длине $y \in L'$: $P(n) \overset{\text{def}}{=} \big[\forall y \in L'\colon |y| \leqslant n \hookrightarrow y \in L\big]$.
Заметим, что $y \in L \Leftrightarrow y \in \varphi(D_2) \Leftrightarrow \varphi^{-1}(y) \cap D_2 \neq \varnothing$. Поэтому будем искать прообраз слова y, принадлежащий D_2.
   (a) $n = 0 \Rightarrow |y| = 0 \Rightarrow y = \varepsilon \in L'$. Пусть $x \overset{\text{def}}{=} \varepsilon \in D_2$ (так как пустое слово — ПСП). Тогда $y = \varepsilon \equiv \varphi(x) \Rightarrow y \in \varphi(D_2) \equiv L \Rightarrow P(0)$
   (b) Фиксируем $n > 0$. Пусть $P(n-1)$. Пусть $y \in L'\colon |y| = n$. Поскольку $|y| = n > 0$, и $|y|$ — четно (см. решение задачи 3 из задания 6), то $|y| \geqslant 2$. Рассмотрим первый и последний символы σ_l и σ_r слова $y \equiv \sigma_l y_1 \sigma_r$:
      i. $\sigma_l = a$, $\sigma_r = b$. Тогда $y = a y_1 b$. $|y_1| = n - 2 \leqslant n - 1 \overset{P(n-1)}{\Rightarrow} \exists x_1 \in D_2\colon \varphi(x_1) = y_1$. Определим $x = [_1 x_1]_1$. $x_1 \in D_2 \Rightarrow x_1$ — ПСП $\Rightarrow x$ — ПСП, так как получен из ПСП добавлением скобок типа 1 слева и справа $\Rightarrow x \in D_2$. Но $\varphi(x) \equiv \varphi([_1 x_1]_1) = \varphi([_1)\varphi(x_1)\varphi(]_1) = a y_1 b \equiv y$. Получаем $\varphi^{-1}(y) \cap D_2 \ni x \Rightarrow \varphi^{-1}(y) \cap D_2 \neq \varnothing$.
      ii. $\sigma_l = b$, $\sigma_r = b$. Тогда $y = b y_1 a$. $|y_1| = n - 2 \leqslant n - 1 \overset{P(n-1)}{\Rightarrow} \exists x_1 \in D_2\colon \varphi(x_1) = y_1$. Определим $x = [_2 x_1]_2$. $x_1 \in D_2 \Rightarrow x_1$ — ПСП $\Rightarrow x$ — ПСП, так как получен из ПСП добавлением скобок типа 2 слева и справа $\Rightarrow x \in D_2$. Но $\varphi(x) \equiv \varphi([_2 x_1]_2) = \varphi([_2)\varphi(x_1)\varphi(]_2) = b y_1 a \equiv y$. Получаем $\varphi^{-1}(y) \cap D_2 \ni x \Rightarrow \varphi^{-1}(y) \cap D_2 \neq \varnothing$.
      iii. $\sigma_l = \sigma_r$. Тогда $y = \sigma y_1 \sigma \in L'$. Воспользуемся утверждением в рамочке из решения задачи 3 задания 6:

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

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

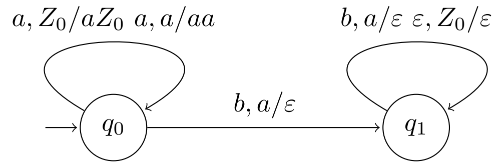
■

<u>Ответ</u>: Верно, что $L = \{w \big| |w|_a = |w|_b\}$.

Задача 6

Автомат $\mathcal{A} = (\Sigma, \Gamma, Q, q_0, Z_0, \delta, \varnothing)$ 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 $G = (N, \Sigma, P, S)$. $N = \{S\} \cup \{[qZp] \big| q, p \in Q, Z \in \Gamma\}$

1. Добавим правила $S \to [q_0 Z_0 q_0] \big| [q_0 Z_0 q_1]$
2. Рассмотрим переходы из δ, добавим правила
   (a) $\delta \ni q_0 \overset{a, Z_0/aZ_0}{\longrightarrow} q_0$: $[q_0 Z_0 q_0] \to a[q_0 a q_0][q_0 Z_0 q_0] \big| a[q_0 a q_1][q_1 Z_0 q_0]$, $[q_0 Z_0 q_1] \to a[q_0 a q_0][q_0 Z_0 q_1] \big| a[q_0 a q_1][q_1 Z_0 q_1]$
   (b) $\delta \ni q_0 \overset{a, a/aa}{\longrightarrow} q_0$: $[q_0 a q_0] \to a[q_0 a q_0][q_0 a q_0] \big| a[q_0 a q_1][q_1 a q_0]$, $[q_0 a q_1] \to a[q_0 a q_0][q_0 a q_1] \big| a[q_0 a q_1][q_1 a q_1]$
   (c) $\delta \ni q_0 \overset{b, a/\varepsilon}{\longrightarrow} q_1$: $[q_0 a q_1] \to b$
   (d) $\delta \ni q_1 \overset{b, a/\varepsilon}{\longrightarrow} q_1$: $[q_1 a q_1] \to b$
   (e) $\delta \ni q_0 \overset{\varepsilon, Z_0/\varepsilon}{\longrightarrow} q_1$: $[q_1 Z_0 q_1] \to \varepsilon$
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 $S \to [q_0 Z_0 q_1]$, $[q_0 Z_0 q_1] \to a[q_0 a q_1][q_1 Z_0 q_1]$, $[q_0 a q_1] \to a[q_0 a q_1][q_1 a q_1] \big| b$, $[q_1 a q_1] \to b$, $[q_1 Z_0 q_1] \to \varepsilon$

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:

$$S \to \underbrace{[q_0 Z_0 q_1]}_{A}, \underbrace{[q_0 Z_0 q_1]}_{A} \to a \underbrace{[q_0 a q_1]}_{B} \underbrace{[q_1 Z_0 q_1]}_{C}, \underbrace{[q_0 a q_1]}_{B} \to a \underbrace{[q_0 a q_1]}_{B} \underbrace{[q_1 a q_1]}_{D} | b, \underbrace{[q_1 a q_1]}_{D} \to b, \underbrace{[q_1 Z_0 q_1]}_{C} \to \varepsilon,$$

   получим

$$S \to A, A \to aBC, B \to aBD|b, D \to b, C \to \varepsilon$$

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

$$S \to A, A \to aB, B \to aBb|b$$

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

$$S \to aB, B \to aBb|b$$

   После примения правила $B \to b$ нельзя применить правило $B \to aBb$, и каждое применение $B \to aBb$ 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: $S \to \underbrace{[q_0 Z_0 q_1]}_{A}, \underbrace{[q_0 Z_0 q_1]}_{A} \to a \underbrace{[q_0 a q_1]}_{B} \underbrace{[q_1 Z_0 q_1]}_{C}, \underbrace{[q_0 a q_1]}_{B} \to a \underbrace{[q_0 a q_1]}_{B} \underbrace{[q_1 a q_1]}_{D} | b, \underbrace{[q_1 a q_1]}_{D} \to b, \underbrace{[q_1 Z_0 q_1]}_{C} \to \varepsilon$
2. После переобозначения $S \to A, A \to aBC, B \to aBD|b, D \to b, C \to \varepsilon$
3. G — однозначная.