Теория и реализация языков программирования.
Задание 9: преобразование контекстно-свободных языков

Сергей Володин, 272 гр.

задано 2013.10.23

Упражнение 1

Упражнение 2

Упражнение 3

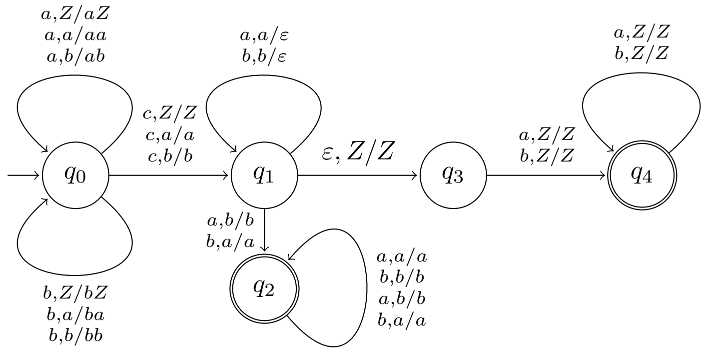
Упражнение 4

Задача 1

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

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

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



2. $\mathcal{A}$ — детерминированный, так как из каждой конфигурации (q, w, γ) переход определен однозначно. ε-переход $q_1 \overset{\varepsilon, Z/Z}{\longrightarrow} q_3$ — единственный переход из q_1 при Z на верхушке стека.

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

    1. Пусть $w \in \{a,b\}^*$. Докажем, что $(q_0, w, Z) \vdash^* (q_0, \varepsilon, w^R Z)$ по индукции по $|w|$:
    $P(n) \overset{\text{def}}{=} \left[\forall w \in \{a,b\}^* \colon |w| = n \hookrightarrow (q_0, w, Z) \vdash^* (q_0, \varepsilon, w^R Z)\right]$
        i. $n = 0 \Rightarrow |w| = 0 \Rightarrow w = \varepsilon$. Тогда $w^R \equiv \varepsilon$, и $(q_0, w, Z) \equiv (q_0, \varepsilon, Z) \equiv (q_0, \varepsilon, w^R Z) \Rightarrow P(0)$.
        ii. Фиксируем $n \geqslant 0$, пусть $\underline{P(n)}$. Пусть $w \in \{a,b\}^*, |w| = n+1$. Тогда $w = w_0\sigma, |w_0| = n$. $P(n) \Rightarrow (q_0, w_0, Z) \vdash^* (q_0, \varepsilon, w_0^R Z)$. Тогда $(q_0, w, \underline{Z)} \equiv (q_0, w_0\sigma, Z) \vdash^* (q_0, \sigma, w_0^R Z)$.
        $\sphericalangle$ переходы из $(q_0, \sigma, w_0^R Z)$. На верхушке стека $\gamma \in \Gamma$, входной символ $\sigma \in \{a,b\}$. Во всех случаях он будет добавлен в стек (см. определение δ), значит, $(q_0, \sigma, w_0^R Z) \vdash (q_0, \varepsilon, \sigma w_0^R Z) \equiv (q_0, \varepsilon, w^R Z) \Rightarrow \underline{P(n+1)}$
    2. Из определения δ $(q_0, cw, \gamma) \vdash^* (q_1, w, \gamma)$, $|\gamma| > 0$.
    3. Докажем $(q_1, x, xZ) \vdash^* (q_1, \varepsilon, Z)$ по индукции по $|x|$: $P(n) \overset{\text{def}}{=} \left[\forall w \in \{a,b\}^* \colon |w| = n \hookrightarrow (q_1, x, xZ) \vdash^* (q_1, \varepsilon, Z)\right]$
        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 $x \in \{a,b\}^* \colon |x| = n+1 \Rightarrow x = x_0\sigma, |x_0| = n \overset{P(n)}{\Rightarrow} (q_1, x_0, x_0 Z) \vdash^* (q_1, \varepsilon, Z)$. Тогда $(q_1, x, xZ) \equiv (q_1, x_0\sigma, x_0\sigma \overline{Z) \vdash^*} (q_1, \sigma, \sigma Z)$. Входной символ совпадает с символом на верхушке стека, из определения δ получаем, что символ будет удален из стека: $(q_1, \sigma, \sigma Z) \vdash (q_1, \varepsilon, Z) \Rightarrow P(n)$.
    4. Из определения δ имеем $(q_1, \sigma_1 x, \sigma_2\gamma) \vdash (q_2, x, \sigma_2\gamma)$ при $\sigma_1 \neq \sigma_2$, $\sigma_1, \sigma_2 \in \{a,b\}^*$.
    5. Из определения δ имеем $(q_2, x, \gamma) \vdash^* (q_2, \varepsilon, \gamma)$, $x \in \{a,b\}^*$ (доказывается очевидно по индукции).
    6. Из определения δ имеем $(q_1, x, Z) \vdash (q_3, x, Z)$
    7. Из определения δ имеем $(q_3, \sigma x, Z) \vdash (q_4, x, Z)$, $\sigma \in \{a,b\}$.

8. Из определения δ имеем $(q_4, x, Z) \vdash^* (q_4, \varepsilon, Z)$, $x \in \{a,b\}^*$ (доказывается очевидно по индукции).
9. Пусть $w \in L \Rightarrow w = xcy, x \neq y^R; x, y \in \{a,b\}^*$. $x \neq y^R \Leftrightarrow x^R \neq y$. Рассмотрим случаи:
   i. Выделим максимальную по длине общую часть w длины i у слов x^R и y: $x^R = wx_1, y = wy_1$, $x_1 \neq y_1$. Тогда $x = x_1^R w^R, w = xcy = x_1^R w^R c w y_1$. Цепочка конфигураций:
   $(q_0, w, Z) \equiv (q_0, x_1^R w^R c w y_1, Z) \overset{31}{\vdash^*} (q_0, cwy_1, wx_1Z) \overset{32}{\vdash} (q_1, wy_1, wx_1Z) \overset{33}{\vdash^*} (q_1, y_1, x_1Z)$.
      A. $|x_1| > 0, |y_1| > 0$, $x_1[1] \neq y_1[1]$. Обозначим $y_1 = y^1...y^l$, $\forall i \in \overline{1,l} \hookrightarrow y^i \in \{a,b\}^*$. Тогда $(q_1, y_1, x_1Z) \equiv (q_1, y^1...y^l, x_1Z) \overset{34}{\vdash} (q_2, y^2...y^l, x_1Z) \overset{35}{\vdash^*} (q_2, \varepsilon, x_1Z)$. $q_2 \in F \Rightarrow \underline{w \in L(\mathcal{A})}$.
      B. $|x_1| = 0, |y_1| > 0$. $y_1 = \sigma y_0, \sigma \in \{a,b\}$. $(q_1, y_1, x_1Z) \equiv (q_1, \sigma y_0, Z) \overset{36}{\vdash} (q_3, \sigma y_0, Z) \overset{37}{\vdash} (q_4, y_0, Z) \overset{38}{\vdash^*} (q_4, \varepsilon, Z)$. $q_4 \in F \Rightarrow \underline{w \in L(\mathcal{A})}$.
      C. $|x_1| > 0, |y_1| = 0$. Тогда $(q_1, y_1, x_1Z) \equiv (q_1, \varepsilon, x_1Z)$.

df

Задача 2

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 :

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

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':

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

2. Удалим недостижимые символы (для упрощения):
   (a) $V_0 \overset{\text{def}}{=} \{S\}$
   (b) $V_1 = V_0 \cup \{A,B,C,E\}$
   (c) $V_2 = V_1 \cup \varnothing$

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

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

1,2. Имеем P'':

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

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''':

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

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

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) $V_0 \overset{\text{def}}{=} \{S\}$

   (b) $V_1 \overset{\text{def}}{=} \{S,A,B,C\}$

   (c) $V_2 = V_1$

   Удаляем E. $P^{(5)}$:
   $A \to a|aC|aB|aBC|aA|aAC|aAB|aABC|ab|abC|aAb|aAbC|Bab|BabC|BaAbC$

   $B \to ba|bBa|bAa|bABa$

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

   $S \to a|aC|aB|aBC|aA|aAC|aAB|aABC|ba|bBa|bAa|bABa|ab|abC|aAb|aAbC|Bab|BabC|BaAbC$

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

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

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

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

   $A' \to a$

   $B' \to b$

   $X_0 \to A'B$

   $X_1 \to A'A$

   $X_2 \to BC$

   $X_3 \to A'B'$

   $X_4 \to B'C$

   $X_5 \to BA'$

   $X_6 \to AB'$

   $X_7 \to B'B$

   $X_8 \to B'A$

   $X_9 \to X_5X_6$

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). $L \overset{\text{def}}{=} \varphi(D_2 \cap \Sigma^*) \equiv \varphi(D_2)$.

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 $y \in L'$: $P(n) \overset{\text{def}}{=} \left[\forall y \in L'\colon |y| \leqslant n \hookrightarrow y \in L\right]$.
   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 $\varphi(x) = \varphi(x_l x_r) = \varphi(x_l)\varphi(x_r) = y_l y_r = y \Rightarrow \varphi^{-1}(y) \cap D_2 \ni x \Rightarrow \varphi^{-1}(y) \cap D_2 \neq \varnothing$

■

Задача 6