

Теория и реализация языков программирования.

Задание 6: Грамматики

Сергей Володин, 272 гр.

задано 2013.10.09

Задача 1

Задача 2

$\Sigma \stackrel{\text{def}}{=} \{a, b\}$, $\Sigma^* \supset L \stackrel{\text{def}}{=} \{w \in \Sigma^* \mid w = w^R\}$ — язык палиндромов из a, b .

1. Определим КС-грамматику $\Gamma \stackrel{\text{def}}{=} (\{S\}, \Sigma, P, S)$, $P \stackrel{\text{def}}{=} \left\{ \underbrace{S \longrightarrow \varepsilon}_{(1)}, \underbrace{S \longrightarrow aSa}_{(2)}, \underbrace{S \longrightarrow bSb}_{(3)}, \underbrace{S \longrightarrow a}_{(4)}, \underbrace{S \longrightarrow b}_{(5)} \right\}$.

Докажем, что $L(\Gamma) = L$:

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

1. $k = 0 \Rightarrow w_k \equiv w_0 = S$. Поэтому $\exists! i = 1: w_0[i] = S$. 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, $P(n+1)$. $P(n) \Rightarrow w_n = w_n^R, \forall i: w_n[i] = S \hookrightarrow i = \frac{|w_n|+1}{2}$.

Предположим, что $\nexists i: w_n[i] = S \Rightarrow w \in \Sigma^*$. 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, $\exists i: w_n[i] = S$. Но $P(n) \Rightarrow \forall i: w_n[i] = S \hookrightarrow i = \frac{|w_n|+1}{2}$. Поэтому $\exists! i = \frac{|w_n|+1}{2}: w_n[i] = S$. 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: $(xx^R)^R = x^R x^R = xx^R$. $w_{n+1} = xy \in \Sigma^* \Rightarrow \forall i \hookrightarrow w_{n+1}[i] \neq S \Rightarrow P(n+1)$ ■

(2). $w_n = xSx^R \xrightarrow{(2)} xSaSx^R = w_{n+1}$. $w_{n+1}^R = x^R a^R S^R a^R x^R = xSaSx^R \equiv w_{n+1}$. $a \neq S \Rightarrow \exists! i: w_{n+1} = S$, $i = \frac{|w_n|+1}{2} + 1 = \frac{|w_n|+3}{2} \equiv \frac{|w_{n+1}|+1}{2} \Rightarrow P(n+1)$ ■.

(3). $w_n = xSx^R \xrightarrow{(3)} xSbSx^R = w_{n+1}$. $w_{n+1}^R = x^R b^R S^R b^R x^R = xSbSx^R \equiv w_{n+1}$. $b \neq S \Rightarrow \exists! i: w_{n+1} = S$, $i = \frac{|w_n|+1}{2} + 1 = \frac{|w_n|+3}{2} \equiv \frac{|w_{n+1}|+1}{2} \Rightarrow P(n+1)$ ■.

(4). $w_n = xSx^R \xrightarrow{(4)} xax^R = w_{n+1}$. $w_{n+1}^R = x^R a^R x^R = xax^R \equiv w_{n+1}$. $w_{n+1} = xax^R \in \Sigma^* \Rightarrow \forall i \hookrightarrow w_{n+1}[i] \neq S \Rightarrow P(n+1)$ ■

(5). $w_n = xSx^R \xrightarrow{(5)} xbx^R = w_{n+1}$. $w_{n+1}^R = x^R b^R x^R = xbx^R \equiv w_{n+1}$. $w_{n+1} = xbx^R \in \Sigma^* \Rightarrow \forall i \hookrightarrow w_{n+1}[i] \neq S \Rightarrow P(n+1)$ ■

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

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

0. $n \bmod 2 = 0 \Rightarrow w = xy$, $|x| = |y|$. $w = w^R \Rightarrow xy = y^R x^R$. Поскольку $|x| = |y|$, $y = x^R \Rightarrow \boxed{w = xx^R}$

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

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

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

а. Пусть $x = \varepsilon$. $S \xrightarrow{(1)} \varepsilon = \varepsilon\varepsilon^R = w \Rightarrow w \in L(\Gamma)$ ■.

б. Иначе $x = x_1 \dots x_m$, $\forall i \in \overline{1, m} \hookrightarrow x_i \in \Sigma$. Рассмотрим символы x_m, \dots, x_1 . Применим правило (2), если $x_i = a$ и

(3) иначе. Примененное правило обозначим за $R(i)$. Получим $S \xrightarrow{(R(m))} x_m S x_m \Rightarrow \dots \xrightarrow{(R(1))} x_1 \dots x_m S x_m \dots x_1$.

Теперь покажем, как получить w :

1. $w = xx^R$. Было получено $S \Rightarrow^* xSx^R$. Тогда $S \Rightarrow^* xSx^R \xrightarrow{(1)} xx^R$ ■

2. $w = xax^R$. Было получено $S \Rightarrow^* xSx^R$. Тогда $S \Rightarrow^* xSx^R \xrightarrow{(4)} xax^R$ ■

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

Получаем $w \in L(\Gamma)$.

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

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

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 $D \Rightarrow^* w \in \Sigma \Leftrightarrow w \in \Sigma^*$, что равносильно $w \in \Sigma^* \Rightarrow D \Rightarrow^* w$. Если $w = \varepsilon$, то применим $D \xRightarrow{(3)} \varepsilon \equiv w$.
Иначе $w = w_1 \dots w_n$, $\forall i \in \overline{1, n} \hookrightarrow w_i \in \Sigma$. Рассмотрим символы w_1, \dots, w_n , если $w_i = a$, применим (1), иначе применим (2). Примененное правило обозначим за $P(i)$. Тогда $D \xRightarrow{P(1)} w_1 D \xRightarrow{P(2)} \dots \xRightarrow{P(n)} w_1 \dots w_n D \xRightarrow{(3)} w_1 \dots w_n \equiv w$.

Задача 3

$$\Sigma \stackrel{\text{def}}{=} \{a, b\}. \Sigma^* \supset L \stackrel{\text{def}}{=} L^- \stackrel{\text{def}}{=} \{w \in \Sigma^* \mid |w|_a = |w|_b\}. \text{ КС-грамматика } \Gamma = \{N, \Sigma, P, S\},$$

$$P \stackrel{\text{def}}{=} \left\{ \underbrace{S \rightarrow SS}_{(1)}, \underbrace{S \rightarrow aSb}_{(2)}, \underbrace{S \rightarrow bSa}_{(3)}, \underbrace{S \rightarrow \varepsilon}_{(4)} \right\}.$$

Докажем, что $L(\Gamma) = L^-$:

- $L(\Gamma) \subset L$. $w \in L(\Gamma) \Rightarrow S \Rightarrow^* w$. Пусть $\{w_k\}_{k=0}^I \subset (N \cup \Sigma)^*$ — последовательность слов при выводе.
 $P(k) \stackrel{\text{def}}{=} [|w_k|_a = |w_k|_b]$. Докажем, что $\forall k \in \overline{0, I} \hookrightarrow P(k)$:
 - $k = 0 \Rightarrow w_k \equiv w_0 = S$. $|w_k|_a = |S|_a = 0 = |S|_b = |w_k|_b \Rightarrow P(0)$.
 - $P(n) \Rightarrow |w_n|_a = |w_n|_b$. $n < I$. Пусть $w_n \xRightarrow{(i)} w_{n+1}$. 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 $P(I) \Rightarrow |w|_a \equiv |w|_a \stackrel{P(I)}{=} |w|_b \equiv |w|_b \Rightarrow w \in L^-$.

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

Пусть $L \ni w = axa$. Тогда $0 = S(w) = |axa|_a - |axa|_b = 2 + S(x) \Rightarrow S(x) = -2$. Отсюда следует, что $|x| \geq 0$. Пусть $|x| = t$, $x = x_1 \dots x_t$, $\forall i \in \overline{1, t} \hookrightarrow x_i \in \Sigma$. Обозначим $f(t): \overline{1, t} \rightarrow \mathbb{Z}: f(i) = S(ax_1 \dots x_i)$. Тогда $f(0) \equiv S(a) = 1$, $f(t) \equiv S(ax_1 \dots x_t) = 1 + S(x) = 1 - 2 = -1$. Заметим, что $|f(t+1) - f(t)| = 1$ («аналог непрерывности»). Поэтому $\exists i \in \overline{1, t}: f(i) = 0$ «принимает промежуточное значение». Получаем, что $w = ax_1 \dots x_i x_{i+1} \dots x_t a = w_l w_r$. 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 $|w_l|, |w_r| \geq |a| = 1$, $|w_l|, |w_r| \leq |w| - 1$. Но $w, w_l, w_r \in L \Rightarrow |w|, |w_l|, |w_r| - \text{четные}$. значит, $|w_l|, |w_r| \leq |w| - 2$. 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

$$w = \sigma x \sigma \in L \Rightarrow w = w_l w_r, |w_l|, |w_r| \in \overline{1, |w| - 2}, w_l, w_r \in L.$$

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

- $m = 0 \Rightarrow w = \varepsilon$. $S \xRightarrow{(4)} \varepsilon = w \Rightarrow w \in L(\Gamma)$
- 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) :
 - $\sigma_1 = a, \sigma_2 = b$. 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, $S \xRightarrow{P(m)} x$. Но $S \xRightarrow{(2)} aSb \xRightarrow{P(m)} x \Rightarrow w \in L(\Gamma)$
 - $\sigma_1 = b, \sigma_2 = a$. Аналогично получаем $w = bxa$, $x \in L(\Gamma) \Rightarrow S \xRightarrow{(3)} bSa \xRightarrow{P(m)} bxa \Rightarrow w \in L(\Gamma)$
 - $\sigma \stackrel{\text{def}}{=} \sigma_1 = \sigma_2$. Разобьем слово $w = \sigma x \sigma$ на подслова $w = w_1 \dots w_k$, $\forall i \in \overline{1, k} \hookrightarrow w_i \in \Sigma^* \cap L, |w_i| \leq |w| - 2, w_i[1] \neq w_i[|w_i|]$.

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 $w_l, w_r \leq |w| - 2$. Значит, $w_i \leq |w| - 2$. Поэтому $S \xRightarrow{P(m)} w_l, S \xRightarrow{P(m)} w_r$ — по предположению индукции. Покажем, как вывести w из S : воспользуемся правилом (1) $k - 1$ раз:

$S \xRightarrow{(1)} SS \xRightarrow{(1)} \dots \xRightarrow{(1)} S^k$. Далее воспользуемся выводами w_i : $S^k \xRightarrow{\text{вывод } w_1} w_1 S^{k-1} \xRightarrow{*} \dots \xRightarrow{\text{вывод } w_k} w_1 \dots w_k \equiv w \Rightarrow w \in L(\Gamma)$

Задача 4

$\Sigma \stackrel{\text{def}}{=} \{a, b\}, \Sigma^* \supset L \stackrel{\text{def}}{=} \{w \in \Sigma^* \mid |w|_b \leq |w|_a\} \equiv \{w \in \Sigma^* \mid S(w) \geq 0\}$ (определение $S(w)$ см. в задаче 3).
Определим КС-грамматику $\Gamma = \{\{S\}, \Sigma, P, S\}$,

$$P \stackrel{\text{def}}{=} \left\{ \underbrace{S \longrightarrow SS}_{(1)}, \underbrace{S \longrightarrow aSb}_{(2)}, \underbrace{S \longrightarrow bSa}_{(3)}, \underbrace{S \longrightarrow \varepsilon}_{(4)}, \underbrace{S \longrightarrow aS}_{(5)} \right\}$$

(добавим в грамматику из предыдущей задачи правила $S \longrightarrow Sa$ и $S \longrightarrow aS$).

Докажем, что $L(\Gamma) = L$:

- $L(\Gamma) \subset L$. $w \in L(\Gamma) \Rightarrow S \Rightarrow^* w$. Пусть $\{w_k\}_{k=0}^I \subset (N \cup \Sigma)^*$ — последовательность слов при выводе.

$P(k) \stackrel{\text{def}}{=} [S(w_k) \geq 0]$. 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. $P(n) \Rightarrow S(w_n) \geq 0$. $n < I$. Пусть $w_n \xrightarrow{(i)} w_{n+1}$. 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 $\Rightarrow S(w_{n+1}) \geq 0 \Rightarrow P(n+1)$

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

- $L \subset L(\Gamma)$. 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. $|w^-| = n^-$, $w^- \in L^- \Rightarrow \exists \{x_i\}_{i=1}^I \subset \{S\} \cup \Sigma^*$ — последовательность слов при выводе слова w^- в грамматике Γ^- из *предыдущей задачи*, $\{p_i\}_{i=1}^I$ — примененные правила.

Определим $f: \overline{1, n^-} \rightarrow \mathbb{N} \cup \{0\}$: $f(i)$ — количество букв a , которые нужно добавить после i -го символа w^- , чтобы получить после всех добавлений w , то есть, $w = w_1^- a^{f(1)} w_2^- a^{f(2)} \dots w_n^- a^{f(n^-)}$.

Модифицируем вывод, добавив буквы a .

Заметим, что если в Γ было применено правило $\alpha_1 \alpha_2 S \alpha_3 \Rightarrow \alpha_1 \alpha_2 \beta \alpha_3$, то после добавления символа a то же правило из Γ также может быть применено: $\alpha_1 a \alpha_2 S \alpha_3 \Rightarrow \alpha_1 a \alpha_2 \beta \alpha_3$, 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 Γ^- (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 $x_{j(i)-1} \equiv \alpha S \beta \xrightarrow{(2)} \alpha \underline{a} S b \beta \equiv x_{j(i)}$. Заменим это на $\alpha S \beta \xrightarrow{(2)} \alpha \underline{a} S b \beta \xrightarrow{(5)} \alpha \underline{a} a S b \beta \xrightarrow{(5)} \dots \xrightarrow{(5)} \alpha \underline{a} a^{f(i)} S b \beta$ — верное применение правил в Γ .
 $f(i)$ раз
2. $w_i^- = a$, $p_{j(i)} = (3)$. То есть $x_{j(i)-1} \equiv \alpha S \beta \xrightarrow{(3)} \alpha b S \underline{a} \beta \equiv x_{j(i)}$. Заменим это на $\alpha S \beta \xrightarrow{(1)} \alpha S S \beta \xrightarrow{(2)} \alpha b S \underline{a} S \beta \xrightarrow{(5)} \alpha b S \underline{a} a S \beta \xrightarrow{(5)} \dots \xrightarrow{(5)} \alpha b S \underline{a} a^{f(i)} S \beta \xrightarrow{(4)} \alpha b S \underline{a} a^{f(i)} \beta$ — верное применение правил в Γ .
 $f(i)$ раз
3. $w_i^- = b$, $p_{j(i)} = (3)$. То есть $x_{j(i)-1} \equiv \alpha S \beta \xrightarrow{(3)} \alpha \underline{b} S a \beta \equiv x_{j(i)}$. Заменим это на $\alpha S \beta \xrightarrow{(3)} \alpha \underline{b} S a \beta \xrightarrow{(5)} \alpha \underline{b} a S a \beta \xrightarrow{(5)} \dots \xrightarrow{(5)} \alpha \underline{b} a^{f(i)} S a \beta$ — верное применение правил в Γ .
 $f(i)$ раз
4. $w_i^- = b$, $p_{j(i)} = (2)$. То есть $x_{j(i)-1} \equiv \alpha S \beta \xrightarrow{(2)} \alpha a S \underline{b} \beta \equiv x_{j(i)}$. Заменим это на $\alpha S \beta \xrightarrow{(1)} \alpha S S \beta \xrightarrow{(3)} \alpha a S \underline{b} S \beta \xrightarrow{(5)} \alpha a S \underline{b} a S \beta \xrightarrow{(5)} \dots \xrightarrow{(5)} \alpha a S \underline{b} a^{f(i)} S \beta \xrightarrow{(4)} \alpha a S \underline{b} a^{f(i)} \beta$ — верное применение правил в Γ .
 $f(i)$ раз

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 $w = w_1^- \dots w_i^- a^{f(i)} \dots w_n^-$, получен его вывод в Γ . Применим такие же рассуждения для остальных символов, получим вывод w в $\Gamma \Rightarrow w \in L(\Gamma)$ ■

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

Задача 5