

# Теория и реализация языков программирования.

## Задание 2: НКА и алгоритмы поиска подстрок

Сергей Володин, 272 гр.

задано 2013.09.18

### Упражнение 1

Пусть  $\sim \subset X \times X$ .  $C(x) = \{z \in X | x \sim z\}$ ,  $C(y) = \{w \in X | y \sim w\}$ .

Пусть  $\exists z \in C(x) \cap C(y)$ . Тогда  $x \sim z, y \sim z$ , и  $w \in C(x) \stackrel{\text{def}}{\Leftrightarrow} x \sim w \stackrel{\substack{\text{транз.} \\ \text{симм.}}}{\Leftrightarrow} z \sim w \stackrel{y \sim z}{\Leftrightarrow} y \sim w \stackrel{\text{def}}{\Leftrightarrow} w \in C(y)$ , то есть,  $C(x) = C(y)$ .

В противном случае  $\neg(\exists z \in C(x) \cap C(y)) \Leftrightarrow C(x) \cap C(y) = \emptyset$ . Получаем, что возможны два случая:

1.  $C(x) \cap C(y) = \emptyset$  (не пересекаются)
2.  $C(x) = C(y)$  (совпадают)

### Упражнение 2

Пусть  $\varphi: \Sigma^* \supseteq X \longrightarrow \Delta^*$ .  $\varphi(\sigma_i) \stackrel{\text{def}}{=} \delta_i \in \Delta^*$ ,  $|\sigma_i| = 1$ .

1. (единственность) Предположим, что существует такое  $\varphi$  — морфизм. Тогда  $\forall w = w_1 \dots w_n \in X, |w_i| = 1 \hookrightarrow \varphi(w) \equiv \varphi(w_1 \dots w_n) = \varphi(w_1) \cdot \varphi(w_2 \dots w_n) = \dots = \varphi(w_1) \cdot \dots \cdot \varphi(w_n) \in \Delta^*$ . Для  $w = \varepsilon$  получаем  $\varphi(\varepsilon) = \varepsilon$ , так как  $\varphi$  — морфизм:  $w_0 \stackrel{\text{def}}{=} \varphi(\varepsilon) = \varepsilon$ .  $\varphi(\varepsilon) \equiv \varphi(\varepsilon\varepsilon) = \varphi(\varepsilon)\varphi(\varepsilon) = w_0w_0 \Rightarrow w_0 = w_0w_0 \Rightarrow |w_0| = |w_0||w_0| \Rightarrow w_0 = \varphi(\varepsilon) = \varepsilon$ .

Таким образом, получаем, что такой морфизм единственный (если существует).

2. (существование) Докажем, что определенное выше отображение  $\varphi$  — морфизм: пусть  $x, y \in X$ . Рассмотрим случаи:
  - a.  $|x| = 0, |y| = 0 \Rightarrow \varphi(xy) = \varphi(\varepsilon\varepsilon) = \varphi(\varepsilon) = \varepsilon = \varepsilon\varepsilon = \varphi(\varepsilon)\varphi(\varepsilon)$
  - b.  $|x| = 0, |y| > 0 \Rightarrow \varphi(xy) = \varphi(y) = \varepsilon\varphi(y) = \varphi(x)\varphi(y)$
  - c.  $|x| > 0, |y| = 0 \Rightarrow \varphi(xy) = \varphi(x) = \varphi(x)\varepsilon = \varphi(x)\varphi(y)$
  - d.  $|x| > 0, |y| > 0 \Rightarrow \varphi(xy) = \varphi(x_1 \dots x_m y_1 \dots y_n) = \underbrace{\varphi(x_1) \dots \varphi(x_m)}_{\varphi(x)} \underbrace{\varphi(y_1) \dots \varphi(y_n)}_{\varphi(y)} = \varphi(x)\varphi(y)$ .

Таким образом, если заданы значения  $\varphi(\sigma_i), \sigma_i \in X \subset \Sigma$ , то морфизм  $\varphi: \Sigma^* \supseteq X \longrightarrow \Delta^*$  с этими значениями существует и единственен.

### Задача 1

Определим  $R_3: \text{REG} \ni X \longrightarrow \mathbb{N} \cup \{0\}$  — количество применений правила 3 из определения регулярности  $X$ . В случае  $X = AB$  или  $X = A|B$ ,  $A, B \in \text{REG}$   $R_3(X) \stackrel{\text{def}}{=} 1 + R_3(A) + R_3(B)$ . В случае  $X = A^*$ ,  $A \in \text{REG}$ , определим  $R_3(X) \stackrel{\text{def}}{=} 1 + R_3(A)$ . В случае  $X = \emptyset$  или  $X = \{\sigma\}$  определим  $R_3(X) \stackrel{\text{def}}{=} 0$ . Функция  $R_3(X)$  определена корректно, так как определение регулярного языка корректное.

Пусть  $\varphi: \Sigma^* \supset X \longrightarrow Y \subset \Delta^*$  — морфизм,  $X \in \text{REG}$ . Докажем, что  $Y \equiv \varphi(X) \in \text{REG}$  индукцией по  $R_3(X)$ :

$P(i) = (\forall X \in \text{REG}: R_3(X) \leq i \forall \varphi \text{ — морфизм} \hookrightarrow \varphi(X) \in \text{REG})$ .

1. Докажем  $P(0)$ : пусть  $X \in \text{REG}: R_3(X) = 0$ . Тогда  $X$  получен без применения третьего правила. Значит,  $\forall \varphi$  — морфизм либо  $X = \emptyset \Rightarrow \varphi(X) = \emptyset$ , либо  $X = \{\sigma\} \Rightarrow \varphi(X) = \{\varphi(\sigma)\} = \{w\}, w \in \Delta^*$ .

Докажем, что  $\Delta^* \supset \{w\} \in \text{REG}$ .  $\{w\} \equiv \{\sigma_1 \dots \sigma_n\} \equiv \{\sigma_1\} \cdot \dots \cdot \{\sigma_n\}$ . Поскольку  $\{\sigma_i\} \in \text{REG}$ , и регулярные языки замкнуты относительно конкатенации (по определению), получаем требуемое.

Итак,  $\varphi(X) \in \text{REG}$  ■

2. Пусть  $P(n)$ . Докажем  $P(n+1)$ . Пусть  $\text{REG} \ni X: R_3(X) \leq n+1$ . Если  $R_3(X) < n+1$ ,  $P(n) \Rightarrow X \in \text{REG}$ .

✂  $X: R_3(X) = n+1$ . Возможны случаи:

- a.  $X = WZ$ ,  $W, Z \in \text{REG}$ . Тогда  $\varphi(X) \equiv \varphi(WZ) = \{\varphi(wz) | w \in W, z \in Z\} = \{\varphi(w)\varphi(z) | w \in W, z \in Z\} = \{\varphi(w) | w \in W\} \cdot \{\varphi(z) | z \in Z\} = \varphi(W)\varphi(Z)$ .  $R_3(X) = 1 + R_3(W) + R_3(Z) = n+1 \Rightarrow R_3(W), R_3(Z) \leq n \xrightarrow{P(n)} \varphi(W), \varphi(Z) \in \text{REG} \Rightarrow \varphi(X) = \varphi(W)\varphi(Z) \in \text{REG}$ .

- b.  $X = W|Z$ ,  $W, Z \in \text{REG}$ . Тогда  $\varphi(X) \equiv \varphi(W|Z) \equiv \varphi(W)|\varphi(Z)$ . Аналогично  $R_3(W), R_3(Z) \leq n \xrightarrow{P(n)} \varphi(W), \varphi(Z) \in \text{REG} \Rightarrow \varphi(X) = \varphi(W)|\varphi(Z) \in \text{REG}$ .
- c.  $X = W^*$ ,  $W \in \text{REG}$ . Тогда  $R_3(X) = 1 + R_3(W) = n + 1 \Rightarrow R_3(W) = n \xrightarrow{P(n)} \varphi(W) \in \text{REG} \Rightarrow \varphi(W^*) = \varphi(\varepsilon|W|WW|\dots) = \varphi(\varepsilon)|\varphi(W)|\varphi(WW)\dots \stackrel{\varphi(\varepsilon)=\varepsilon}{=} \varepsilon|\varphi(W)|\varphi(WW)\dots = \varphi(W)^* \in \text{REG}$ .

Получаем  $\forall i \geq 0 \hookrightarrow P(i) \Rightarrow \forall X \in \text{REG} \forall \varphi - \text{морфизм} \hookrightarrow \varphi(X) \in \text{REG} \blacksquare$

## Задача 2

- Нет. Пусть  $\Sigma = \{0, 1\}$ ,  $L = \Sigma^*$ . Определим  $\varphi: L \longrightarrow L: \forall w \in L \hookrightarrow \varphi(w) = \varepsilon$ . В этом случае  $\varphi$  — морфизм, так как  $\forall x \in L \forall y \in L \hookrightarrow \varphi(xy) = \varepsilon = \varepsilon\varepsilon = \varphi(x)\varphi(y)$ . Тогда  $\forall \varnothing \neq X \subset L \hookrightarrow \varphi(X) = \{\varepsilon\}$ , так как  $\forall w \in L \hookrightarrow \varphi(w) = \varepsilon$ . Поскольку  $\varphi(\varepsilon) = \varepsilon \in L$ ,  $\varphi^{-1}(\varepsilon) \ni \varepsilon \Rightarrow \varphi^{-1}(L) \supset \{\varepsilon\} \neq \varnothing \Rightarrow \varphi^{-1}(L) \neq \varnothing \Rightarrow \varphi(\varphi^{-1}(L)) = \{\varepsilon\} \neq L$ . Таким образом,  $\exists L \subseteq \Sigma^* \exists \varphi - \text{морфизм}: \varphi(\varphi^{-1}(L)) \neq L$ .
- Нет. Пусть  $\Sigma = \{a, b\}$ ,  $L = \{b\}^*$ ,  $\varphi(a) \stackrel{\text{def}}{=} \varphi(b) \stackrel{\text{def}}{=} a$ . Доопределим  $\varphi$  так, чтобы оно было морфизмом (это возможно, см. упражнение 2). Тогда  $\varphi(L) \equiv \varphi(\{b^*\}) \ni \varphi(b) = a \Rightarrow \varphi^{-1}(\varphi(L)) \supset \varphi^{-1}(a) \ni a \notin L \Rightarrow \varphi^{-1}(\varphi(L)) \not\subseteq L \Rightarrow \varphi^{-1}(\varphi(L)) \neq L$ . Таким образом,  $\exists L \subseteq \Sigma^* \exists \varphi - \text{морфизм}: \varphi^{-1}(\varphi(L)) \neq L$ .
- Нет. Пусть  $\Sigma = \{a, b\}$ ,  $L = \{ab\}$ , морфизм  $\varphi: \Sigma^* \longrightarrow \Sigma^*$  — из предыдущего пункта. Тогда  $\varphi(L) = \{\varphi(ab)\} = \{\varphi(a)\varphi(b)\} = \{aa\}$ ,  $\varphi^{-1}(L) = \{x \in \Sigma^* | \varphi(x) \in \{ab\}\} = \{x \in \Sigma^* | \varphi(x) = ab\} = \varnothing$ , так как  $\varphi(\Sigma^*) = \varphi((a|b)^*) \stackrel{1,2,c}{=} (\varphi(a|b))^* = \{\varphi(a), \varphi(b)\}^* = \{a\}^* = a^* \not\ni ab$ . Тогда  $\varphi(\varphi^{-1}(L)) = \varphi(\varnothing) = \varnothing \not\ni aa \in \varphi^{-1}(aa) = \varphi^{-1}(\varphi(L))$ . Таким образом,  $\exists L \subseteq \Sigma^* \exists \varphi - \text{морфизм}: \varphi(\varphi^{-1}(L)) \neq \varphi^{-1}(\varphi(L))$ .

## Задача 3

$L \in \text{REG} \Rightarrow \exists \mathcal{A} - \text{ДКА}: L(\mathcal{A}) = L$ . Построим ДКА  $\mathcal{A}'$  для  $L^{-1} \stackrel{\text{def}}{=} \varphi^{-1}(L)$ . Для этого каждый переход по  $\sigma$  в  $\mathcal{A}$  заменим на переход по  $\varphi^{-1}(\sigma)$ , а именно, переход по множеству слов понимается как множество переходов по словам, переход по слову — с введением дополнительных состояний.

## Задача 4

Пусть языки  $\Sigma^* \supset X, Y \in \text{REG}$ . Докажем, что

- $X \cup Y \in \text{REG}$ : из определения регулярности  $\forall X, Y \in \text{REG} \hookrightarrow X \cup Y \in \text{REG} \blacksquare$
- $\overline{X} \stackrel{\text{def}}{=} \Sigma^* \setminus X \in \text{REG}$ :  $X \in \text{REG} \Rightarrow \exists$  полный ДКА  $\mathcal{A}: L(\mathcal{A}) = X$ .  $F' \stackrel{\text{def}}{=} Q \setminus F$ ,  $\mathcal{A}'$  — автомат  $\mathcal{A}$  с множеством принимающих состояний  $F'$ . Докажем, что  $L(\mathcal{A}') = \Sigma^* \setminus X$ :  $w \in \Sigma^*$ ,  $(q_0, w) \vdash^* (q_w, \varepsilon)$  (здесь используется полнота).  $w \in X \Leftrightarrow w \in L(\mathcal{A}) \Leftrightarrow q_w \in F \Leftrightarrow \neg(q_w \in Q \setminus F) \Leftrightarrow \neg(q_w \in F') \Leftrightarrow \neg(w \in L(\mathcal{A}'))$ . Но  $w \in X \Leftrightarrow \neg(w \in \Sigma^* \setminus X)$ , откуда  $\neg(w \in \Sigma^* \setminus X) \Leftrightarrow \neg(w \in L(\mathcal{A}'))$  и Получаем ДКА  $\mathcal{A}': L(\mathcal{A}') = \Sigma^* \setminus X \xrightarrow[\text{семинаре}]{\text{на}} \Sigma^* \setminus X \in \text{REG} \blacksquare$
- $X \cap Y \in \text{REG}$ :  $X \cap Y = \overline{\overline{X \cap Y}}$ .  $X, Y \in \text{REG} \xrightarrow{(2)} \overline{X}, \overline{Y} \in \text{REG} \xrightarrow{(1)} \overline{X} \cup \overline{Y} \in \text{REG} \xrightarrow{(2)} \overline{\overline{X \cap Y}} \in \text{REG} \blacksquare$   
 $w \in X \cap Y \Leftrightarrow \begin{cases} w \in X \\ w \in Y \end{cases} \Leftrightarrow \begin{cases} \neg(w \in \overline{X}) \\ \neg(w \in \overline{Y}) \end{cases} \Leftrightarrow \neg \left[ \begin{matrix} w \in \overline{X} \\ w \in \overline{Y} \end{matrix} \right] \Leftrightarrow \neg(w \in \overline{X} \cup \overline{Y}) \Leftrightarrow w \in \overline{\overline{X \cap Y}}$  (подразумевается  $w \in \Sigma^*$ )  $\blacksquare$
- $X \setminus Y \in \text{REG}$ :  $X \setminus Y = X \cap \overline{Y}$ .  $Y \in \text{REG} \xrightarrow{(2)} \overline{Y} \in \text{REG} \xrightarrow{(3)} X \cap \overline{Y} \in \text{REG} \blacksquare$   
 $w \in X \cap \overline{Y} \Leftrightarrow \begin{cases} w \in X \\ w \in \overline{Y} \end{cases} \Leftrightarrow \begin{cases} w \in X \\ \neg(w \in Y) \end{cases} \Leftrightarrow w \in X \setminus Y$  (подразумевается  $w \in \Sigma^*$ )  $\blacksquare$

## Задача 5

$\Sigma \stackrel{\text{def}}{=} \{a\}$ . Предположим, что  $\Sigma^* \subset L = \{a^{2^n} | n \geq 0\} \in \text{REG} \xrightarrow[\text{о накачке}]{\text{по лемме}} \exists p = p_0 \geq 1: \forall w \in L \hookrightarrow (w = xyz, |y| \geq 1, |xy| \leq p, (\forall i \geq 0 \hookrightarrow xy^i z \in L))$ . Фиксируем  $n = n_0 = p, w_0 = a^{2^p} \in L$ . Получаем  $w_0 = x_0 y_0 z_0, |y_0| \geq 1, |x_0 y_0| \leq p$ . Поскольку  $L \subset a^*$ ,  $y \in a^*$ , откуда  $y = a^j, j \geq 1$ . Аналогично  $x = a^i, z = a^k \Rightarrow w_0 = a^{2^p} = xyz = a^{i+j+k} \Rightarrow i+j+k = 2^p$ . По лемме должно выполняться  $xy^2 z = a^{i+2j+k} \in L \Rightarrow a^{i+2j+k} = a^{2^q}$ , откуда  $i+2j+k = 2^q \Rightarrow j = 2^q - 2^p \geq 2^{p+1} - 2^p = 2^p(2-1) = 2^p$ . Но  $|x_0 y_0| \leq p \Rightarrow |y_0| \leq p$ . Получаем  $p \geq |y_0| = j \geq 2^p$  — противоречие, т.к.  $\forall p \geq 1 \hookrightarrow p < 2^p$ .

Значит, предположение неверно, и  $L \notin \text{REG} \blacksquare$