Регулярни езици — част 2

(по-сложни примери)

В следващите примери нека $\Sigma = \{a, b\}$.

1) $L_1 = \{w \in \Sigma^* \mid w$ започва с $\underline{a}\}$.

$w \in L_1 \iff \exists v\,(v \in \Sigma^* \;\&\; w = a \cdot v)$

(ако $w = L_1$, то w започва с $\underline{a}$; нека v е частта от w след първото $\underline{a}$; тогава $w = av$ като $v \in \Sigma^*$ като част от $w \in \Sigma^*$.

обратно, ако $v \in \Sigma^*$ и $a \cdot v = w$, то понеже $a \in \Sigma$ имаме, че $w \in \Sigma^*$ и започва с $\underline{a}$; така $w \in L_1$)

Тогава $L_1 = \{a\} \cdot \Sigma^* = \{a\} \cdot (\{a\} \cup \{b\})^*$

и, следователно, L_1 е регулярен език над Σ.

Рег. израз, описващ L_1 е $a \cdot (a \cup b)^*$,

или по-просто: $a\Sigma^*$

<u>Упр. 1)</u>. Докажете, че $L_1' = \{w \in \Sigma^* \mid w$ започва с 'b'$\}$ е регулярен над Σ.

2) $L_2 = \{w \mid w \in \Sigma^* \;\&\; w$ завършва с 'b'$\}$

$w \in L_2 \iff \exists v\,(v \in \Sigma^* \;\&\; w = v \cdot b)$

(защо?)

Следователно $L_2 = \Sigma^*.\{b\} = (\{a\}\cup\{b\})^*.\{b\}$, което е рег. език над Σ; описва се с рег. израз $\Sigma^*.b$

<u>Упр. 2.</u> Намерете рег. израз описващ езика

$L_2' = \{ w \in \Sigma^* \mid w$ завършва на 'ba'$\}$

3) $L_3 = \{ w \in \Sigma^* \mid w$ съдържа 'aa'$\}$

$w \in L_3 \Leftrightarrow \exists v \exists u \, (v \in \Sigma^* \ \& \ u \in \Sigma^* \ \& \ w = v.aa.u)$

(ако $w \in L_3$, то w съдържа aa: $\underbrace{\underbrace{\quad}_{v}\, aa \,\underbrace{\quad}_{u}}_{w}$

нека например да разгледаме първото срещане на 2 последователни a-та в w; нека частта преди тях е v, а след тях — u. Понеже $w \in \Sigma^*$, а v и u са части на w, то $u, v \in \Sigma^*$ и $w = v.aa.u$;

обратно, ако $w = v.aa.u$, $v, u \in \Sigma^*$, то $w \in \Sigma^*$ и w съдържа aa, т.е. $w \in L_3$).

Следователно, $L_3 = \Sigma^*.\{aa\}.\Sigma^*$, откъдето е регулярен над Σ. Описва се с израза

$(a \cup b)^*.aa(a\cup b)^*$, или по-просто

$\Sigma^* aa \Sigma^*$

4) $L_4 = \{ w \in \Sigma^* \setminus \{\varepsilon\} \mid$ първият и последният символ на w са равни $\}$

Да забележим, че $L_4 = L_a \cup L_b$, където

$L_a = \{ w \in \Sigma^+ \mid$ първият и последният символ на w е 'a' $\}$

и

$L_b = \{ w \in W^+ \mid$ първият и последният символ на w е 'b' $\}$

Освен това $w \in L_a \Leftrightarrow w = a$ или $\exists v (v \in \Sigma^* \& w = ava)$

Следователно, $L_a = \{a\} \cup \{a\} . \Sigma^* . \{a\}$.

Напълно подобно, $L_b = \{b\} \cup \{b\} . \Sigma^* . \{b\}$.

Но L_a и L_b са рег. езици, откъдето като тяхно обединение L_4 също ще бъде рег. над Σ. Описва се от рег. израз

$$a \cup a\Sigma^* a \cup b \cup b\Sigma^* b$$

<u>Упр. 4)</u> Док., че езикът $L_4' = \{ w \in \Sigma^* \mid$ първият и последният символ на w са различни $\}$ е регулярен. Намерете рег. израз за L_4'.

5) $L_5 = \{ w \in \Sigma^+ \mid$ последният символ на w се е срещал по-рано $\}$

Пак разграничаваме 2 възможности, в зависимост кой е последният символ a или b: $L_5 = L_a \cup L_b$ където $L_x = \{ w \in L_5 \mid$ последният символ на w е $x \}$, $x \in \{a, b\}$.

Тогава L_a ~~се~~ характеризира с:

$w \in L_a \iff w$ завършва на a, което се е срещало преди това $\iff$

$\iff \exists u, v\ (u, v \in \Sigma^* \ \&\ w = u.a.v.a)$

// както $\underline{u}$, така и $\underline{v}$ могат да са ε!)

След., $L_a = \Sigma^*.\{a\}.\Sigma^*.\{a\}$.

Подобно, $L_b = \Sigma^*.\{b\}.\Sigma^*.\{b\}$

и като регулярни, тяхното обединение L_5 също ще бъде регулярен. Описва се от рег. изр.

$$\Sigma^* a \Sigma^* a \ \cup\ \Sigma^* b \Sigma^* b$$

6) $L_6 = \{w \in \Sigma^* \mid \exists s \in \Sigma$ (w започва с s и съдържа 'ss')$\}$

Отново разграничаваме два случая, в зависимост кой от двата символа a или b е s:

$L_6 = L_a \cup L_b$, където

$L_a = \{w \in \Sigma^* \mid w$ започва с $\underline{a}$ и съдържа 'aa'$\}$

и L_b дефинирано аналогично.

Ако $w \in L_a$, то има две възможности за w: вторият символ е a, т.е. $w = aav$ за някое $v \in \Sigma^*$; или пък вторият символ на w не е $\underline{a}$ (и значи е $\underline{b}$), т.е. $w = abv$

Тогава за да бъде от L_6 трябва v да съдържа 'aa' — така ~~w~~

$$w = ab\,v_1\,aa\,v_2$$

за подходящи $v_1, v_2 \in \Sigma^*$.

Не е трудно да се съобрази, че ако w има някое от тези изразявания, то $w \in L_a$.

Следователно, $L_a = \{aa\}\cdot\Sigma^* \cup \{ab\}\cdot\Sigma^*\{aa\}\Sigma^*$

Напълно аналогично, $L_b = \{bb\}\Sigma^* \cup \{ba\}\cdot\Sigma^*\{bb\}\Sigma^*$.

Така $L_6 = L_a \cup L_b$ също е регулярен, като се описва от израза

$$aa\Sigma^* \cup ab\Sigma^* aa\Sigma^* \cup bb\Sigma^* \cup ba\Sigma^* bb\Sigma^*.$$

7) $L_7 = \{w \in \Sigma^* \mid |w|_a \text{ е четно}\}$

За да опишем по-просто L_7 ще се нуждаем от следните помощни езици:

$\Lambda_0 = \{w \in \Sigma^* \mid |w|_a = 0\}$ и // w няма a-та

$\Lambda_2 = \{w \in \Sigma^* \mid |w|_a = 2\}$ // w има точно 2 a-та

Имаме, че: $w \in \Lambda_0 \Leftrightarrow w$ не съдържа $\underline{a} \Leftrightarrow$

$\Leftrightarrow w \in (\Sigma \setminus \{a\})^* \Leftrightarrow w \in \{b\}^*$, т.е.

$\Lambda_0 = \{b\}^*$.

$w \in \Lambda_2 \iff w$ е от вида $b\ldots bab\ldots bab\ldots b$

$\iff \exists v_1, v_2, v_3\ (v_1 \in \{b\}^* \,\&\, v_2 \in \{b\}^* \,\&\, v_3 \in \{b\}^* \,\&\, w = v_1 a v_2 a v_3)$

Следователно, $\Lambda_2 = \{b\}^* \cdot \{a\} \cdot \{b\}^* \cdot \{a\} \cdot \{b\}^*$

Сега да се върнем към L_7. Да забележим, че

$w \in L_7 \iff w \in \Lambda_0$ или $\exists u, v\ (u \in \Lambda_2 \,\&\, v \in L_7 \,\&\, w = u \cdot v)$.

Наистина, посоката $\Leftarrow$ е ясна: ако $w \in \Lambda_0$, то $|w|_a = 0$ е четно; ако пък $u \in \Lambda_2$, $v \in L_7$ и $w = uv$, то понеже $|u|_a = 2$, $|v|_a$ – четно и $|w|_a = |u|_a + |v|_a$, ще получим, че $|w|_a$ е четно, т.е. $w \in L_7$.

За посоката $(\Rightarrow)$: Нека $w \in L_7$ и $w \notin \Lambda_0$; тогава w съдържа поне 1 a; но понеже броят на a-та в w е четен, то трябва да има поне 2 a-та в w. Нека $\underline{u}$ е началото на w, което свършва на второто $\underline{a}$ в w:

$$\underbrace{\underbrace{b\ldots bab\ldots ba}_{u}\,v}_{w}$$

а $\underline{v}$ е остатъкът от w;

Така в $\underline{u}$ има 2 a-та, т.е. $u \in \Lambda_2$. Понеже $|w|_a$ е четно, $|u|_a = 2$ и $|v|_a = |w|_a - |u|_a$, то $|v|_a$ е четно и, следователно, $v \in L_7$.

Сега ще докажем, че $L_7 = \Lambda_2^* \cdot \Lambda_0$.

$(\supseteq)$ Нека първо $w \in \Lambda_2^* \cdot \Lambda_0$. Следователно,

$w = v_1 \cdot \ldots v_n \cdot u$, където $v_1, \ldots, v_n \in \Lambda_2$, $u \in \Lambda_0$ или $w = u \in \Lambda_0$. И в двата случая $|w|_a$ е четно число, откъдето $w \in L_7$.

($\subseteq$) Нека $w \in L_7$; ако $|w|_a = 0$, то $w \in \Lambda_0 \subseteq \Lambda_2^* \cdot \Lambda_0$ (защото $\varepsilon \in \Lambda_2^*$). Нека сега $|w|_a > 0$ и за всяко $\tilde{w} \in L_7$, за което $|\tilde{w}|_a < |w|_a$, е в сила, че $\tilde{w} \in \Lambda_2^* \cdot \Lambda_0$. Щом $|w|_a > 0$ и $w \in L_7$, то в w има поне 2 а-та. Тогава w се представя като конкатенацията

$$w = b \ldots b\, a\, b \ldots b\, a \cdot \tilde{w},$$

където $b \ldots b a b \ldots b a \in \Lambda_2$, а $\tilde{w}$ е дума с четен брой а-та (т.е. $\tilde{w} \in L_7$) като $|\tilde{w}|_a < |w|_a$ (махнали сме 2 а-та от w). Следователно, $\tilde{w} \in \Lambda_2^* \cdot \Lambda_0$. Тогава: $w \in \Lambda_2 \cdot \Lambda_2^* \cdot \Lambda_0 \subseteq \Lambda_2^* \cdot \Lambda_0$

Окончателно, $L_7 = \Lambda_2^* \cdot \Lambda_0$. Понеже Λ_2 и Λ_0 са рег., то и L_7 ще бъде регулярен. Описва се от рег. израз:

$$(b^* a b^* a b^*)^* b^*.$$