§4. Замкнутость класса регулярных языков относительно операций: реверс, морфизмы, деление (частное), замыкания относительно префиксов, суффиксов, подслов. Построение соответствующих автоматов.

Опр. Пусть $w = a_1 \dots a_n \in \Sigma^*$. Реверсом слова w называется слово $R(w) = a_n \dots a_1$.

Опр. Реверсом языка L называется язык $R(L) = \{R(w) \mid \forall w \in L\}$.

Лемма 1. Если L регулярный, то $R(L)$ тоже регулярный.

Доказательство:

Пусть язык L допускается ДКА $A = (Q, \Sigma, \varphi, q_0, Q_F)$.

Построим НКА $B = (Q, \Sigma, \delta, Q_F, q_0)$,

где $\delta(q', a) \ni q$, если $\varphi(q, a) = q'$.

(На диаграмме автомата B все стрелки разворачиваются в обратную сторону).

Очевидно, $L(B) = R(L(A))$. Лемма 1 доказана.

Опр. Пусть Σ_1 и Σ_2 – два алфавита,

всюду определенная функция подстановки h: $\Sigma_1 \to \Sigma_2^*$,

$w = a_1 \dots a_n \in \Sigma_1^*$.

Гомоморфизмом слова w называется слово

$\Gamma(w) = h(a_1) \dots h(a_n) \in \Sigma_2^*$.

Пример. $\Sigma_1 = \{a, b\}$, $\Sigma_2 = \{0, 1\}$.

$h(a) = 100$, $h(b) = \varepsilon$.

$w = abba$.

$\Gamma(w) = ?$

$\Gamma(w) = 100100$.

Опр. Гомоморфизмом языка L называется язык
$\Gamma(L) = \{\Gamma(w) \mid \forall w \in L\}$.

Лемма 2. Если L регулярный, то $\Gamma(L)$ тоже регулярный.

Доказательство:
Пусть язык L допускается неполным ДКА $A = (Q, \Sigma_1, \varphi, q_0, Q_F)$.
Построим ε-НКА $B = (\widetilde{Q}, \Sigma_2, \delta, q_0, Q_F)$, в котором каждая дуга из q в q' с пометкой a заменяется на путь из q в q' с пометками $h(a)$.

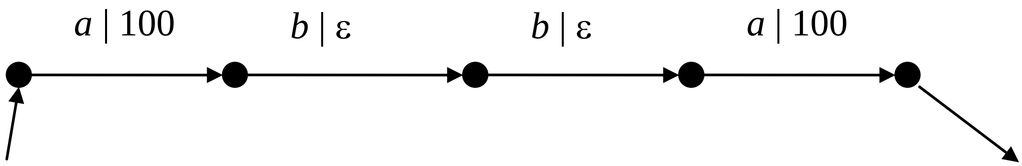
Очевидно, $L(B) = \Gamma(L(A))$. Лемма 2 доказана.

Замечание: для известной функции подстановки h существует трансдьюсер τ, такой, что $\tau(\mathrm{L}) = \Gamma(L)$.

Пример. $\Sigma_1 = \{a, b\}$, $\Sigma_2 = \{0, 1\}$.

$h(a) = 100$, $h(b) = \varepsilon$.

$L = \{abba\}$.



Опр. Пусть $L \subseteq \Sigma_2^*$. Обратным гомоморфизмом языка L называется язык $\Gamma^{-1}(L) = \{w \mid \Gamma(w) \in L\}$.

Пример. $\Sigma_1 = \{a, b\}$, $\Sigma_2 = \{0, 1\}$.
$h(a)$ =100, $h(b)$ = ε.
L = {100, 110, 111 }.

$\Gamma^{-1}(L)$ = ?

$\Gamma^{-1}(L) = \{b^* a b^*\}$.

Лемма 3. Если L регулярный, то $\Gamma^{-1}(L)$ тоже регулярный.

Доказательство:

Пусть язык L допускается неполным ДКА $A = (Q, \Sigma_2, \varphi, q_0, Q_F)$.

Построим НКА $B = (Q, \Sigma_1, \delta, q_0, Q_F)$,

где $\delta(q, x) = q' \Leftrightarrow \varphi(q, h(x)) = q'$ (1 этап);

если $h(x) = \varepsilon$, то $\delta(q, x) = q$ для каждого q (2 этап).

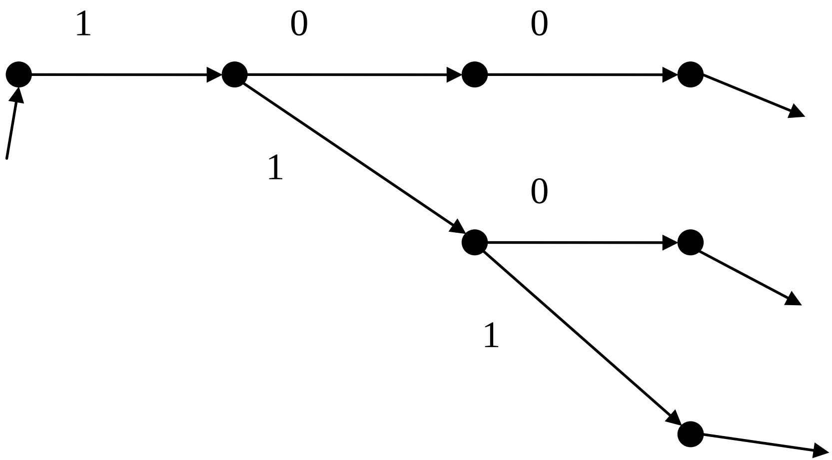
Очевидно, $L(B) = \Gamma^{-1}(L(A))$. Лемма 3 доказана.

Пример. $\Sigma_1 = \{a, b\}$, $\Sigma_2 = \{0, 1\}$.

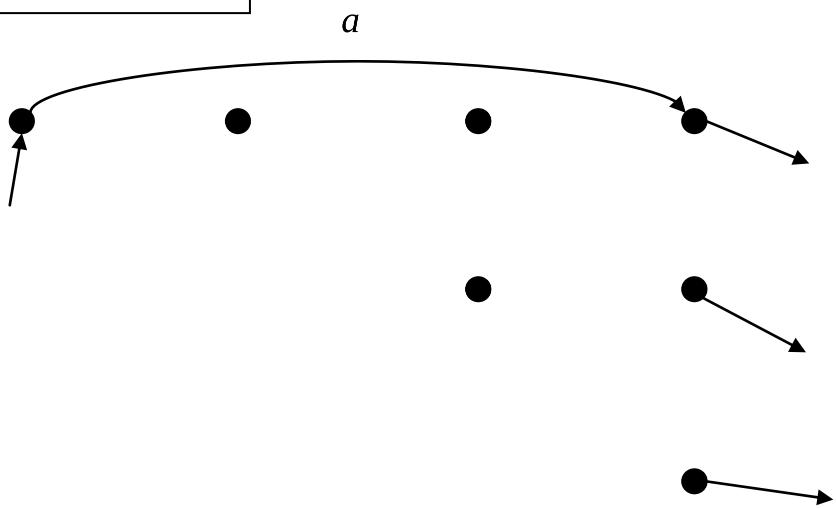
$h(a) = 100$, $h(b) = \varepsilon$.

$L = \{100, 110, 111\}$.

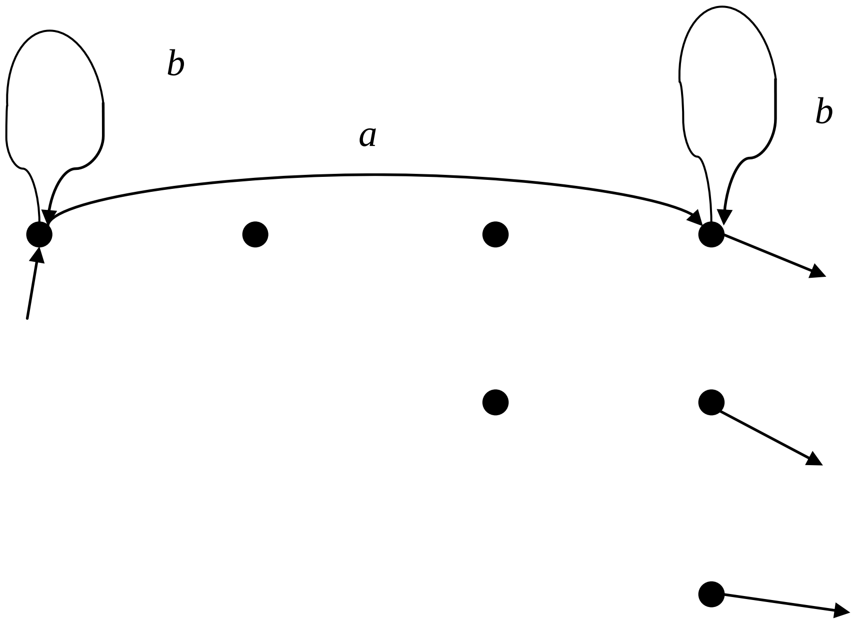
$\boxed{A}$



B – 1 этап



B – 2 этап



(повтор) Опр. Пусть L и K – языки над общим алфавитом Σ. Будем называть правое частное – язык $LK^{-1} = \{w \in \Sigma^* \mid \exists u \in K : wu \in L\}$.
«префиксы слов языка L, относительно K»

(повтор) Опр. Пусть L и K – языки над общим алфавитом Σ. Будем называть левое частное – язык $K^{-1}L = \{w \in \Sigma^* \mid \exists u \in K : uw \in L\}$.
«суффиксы слов языка L, относительно K»

Лемма 4. Если язык L регулярный а K произвольный, то LK^{-1} и $K^{-1}L$ регулярные.

Доказательство:

2) Докажем регулярность для $K^{-1}L$.

Рассмотрим бинарное отношение ~ на Σ^*:

$u \sim v \Leftrightarrow u^{-1}(K^{-1}L) = v^{-1}(K^{-1}L)$.

Обозначим $T_1, T_2, ..., T_n, ...$ – классы эквивалентности по отношению ~.

Для каждого T_n выполняется $T_n^{-1}(K^{-1}L) = u^{-1}(K^{-1}L)$ для каждого $u \in T_n$.

$$T_n^{-1}(K^{-1}L) = \{w \mid \exists t \in T_n : tw \in K^{-1}L\} =$$
$$= \{w \mid \exists t \in T_n : \exists k \in K : ktw \in L\} = (KT_n)^{-1}L\,.$$

Пусть $S_1, S_2, \dots, S_m$ – классы эквивалентности по отношению $\sim_L$.
Так как L – регулярный язык, количество этих классов конечное.
Этим классам соответствуют несовпадающие множества
$S_1^{-1}L, \dots, S_m^{-1}L$.

Выше показано, что для любого T_n существует $S_i :\ KT_n \subseteq S_i$.

Покажем, что количество классов $T_1, T_2, \ldots, T_n, \ldots$ не превышает количества $S_1, S_2, \ldots, S_m$.

(от противного) Предположим, что $T_1 \neq T_2$, $T_1 \cap T_2 = \varnothing$ и $KT_1 = KT_2$.
Это значит, $\forall t \in T_1, \forall k \in K : kt \in KT_2$.
Пусть $t_{\min} \in T_1$ имеет наименьшую длину;
$k_{\min} \in K$ имеет наименьшую длину.
$k_{\min} t_{\min} \in KT_2 \Rightarrow$ либо $t_{\min} \in T_2$ (противоречие с $T_1 \cap T_2 = \varnothing$), либо существует собственный суффикс $t_{суф}$ слова $t_{\min}$, такой, что $t_{суф} \in T_2$.

$$k_{\min} t_{суф} \in KT_2 \Rightarrow k_{\min} t_{суф} \in KT_1 \Rightarrow$$

либо $t_{суф} \in T_1$ (противоречие с минимальностью длины $t_{\min}$), либо существует собственный суффикс t_{c2} слова $t_{суф}$, такой, что $t_{c2} \in T_1$ (противоречие с минимальностью длины $t_{\min}$).

1) Для доказательства регулярности LK^{-1} используют аналогичные рассуждения для бинарных отношений:

$\sim$ на $\Sigma^*: u \sim v \Leftrightarrow (LK^{-1})u^{-1} = (LK^{-1})v^{-1}$;

$\sim_L$ на $\Sigma^*: u \sim_L v \Leftrightarrow Lu^{-1} = Lv^{-1}$.

Лемма 4 доказана.

Опр. Пусть L язык над алфавитом Σ. Будем называть префиксным замыканием языка L язык $\mathrm{Pref}(L) = \{w \in \Sigma^* \mid \exists u \in \Sigma^* : wu \in L\}$.
«все префиксы слов языка L»

Будем называть суффиксным замыканием языка L язык
$\mathrm{Suff}(L) = \{w \in \Sigma^* \mid \exists u \in \Sigma^* : uw \in L\}$.
«все суффиксы слов языка L»

Будем называть замыканием относительно подслов языка L язык
$Fact(L) = \{w \in \Sigma^* \mid \exists u_1, u_2 \in \Sigma^* : u_1 w u_2 \in L\}$.
«все подслова слов языка L»

Свойства:

1) $\mathrm{Pref}(L) = L(\Sigma^*)^{-1}$.

2) $\mathrm{Suff}(L) = (\Sigma^*)^{-1}L$.

3) $\mathrm{Fact}(L) = \mathrm{Pref}(\mathrm{Suff}(L))$.

4) $\mathrm{Fact}(L) \subset \mathrm{Suff}(\mathrm{Pref}(L))$.

Д/З – доказать, что $\mathrm{Fact}(L) \neq \mathrm{Suff}(\mathrm{Pref}(L))$.

Лемма 5. Если L – регулярный язык, то $\mathrm{Pref}(L)$, $\mathrm{Suff}(L)$, $\mathrm{Fact}(L)$ – также регулярные.

Доказательство следует из леммы 4 и свойств.

Построение конечного автомата для $\mathrm{Pref}\,(L)$ и $\mathrm{Suff}\;(L)$, где L – регулярный.

Пусть язык L допускается неполным ДКА $A = (Q, \Sigma, \varphi, q_0, Q_F)$, в котором все состояния достижимы и из любого состояния есть маршрут в какое-нибудь заключительное состояние (т.е. нет тупиковых состояний).

Построим неполный ДКА $P = (Q, \Sigma, \varphi, q_0, Q)$.

Очевидно, что $L(P) = \mathrm{Pref}\,(L)$.

Построим НКА $S = (Q, \Sigma, \varphi, Q, Q_F)$.

Очевидно, что $L(S) = \mathrm{Suff}\,(L)$.

Вопрос: как построить НКА для $\mathrm{Fact}(L)$?

Ответ: $W = (Q, \Sigma, \varphi, Q, Q)$.

При построении конечного автомата для $\mathrm{Suff}(L)$, где L – регулярный, можно также использовать свойство:

$$\mathrm{R}(\mathrm{Suff}(L)) = \mathrm{Pref}(R(L)).$$

Обоснование свойства:

$$\forall w \in \mathrm{Suff}(L) \to \exists u \in \Sigma^* : uw \in L \;\Rightarrow$$
$$R(uw) = R(w)R(u) \in R(L) \;\Leftrightarrow$$
$$R(w) = \mathrm{Pref}(R(L))$$

Д/З*. Построить конечный автомат для LK^{-1} и $K^{-1}L$, где L регулярный, а K – произвольный.