§7. Регулярные языки. Теорема Клини

Опр. Язык называется регулярным, если он получается из конечных языков применением операций объединения, произведения, итерации.

Обозначим $\Omega_{\text{рег}}$ класс всех регулярных языков над фиксированным алфавитом Σ.

Теорема (Клини).

$\Omega_{\text{рег}} = \Omega_{\text{авт}}$.

Доказательство $\Omega_{\text{рег}} \subseteq \Omega_{\text{авт}}$ следует из теоремы о замкнутости класса $\Omega_{\text{авт}}$ и утверждения о конечных языках из предыдущего параграфа.

Доказательство $\Omega_{\text{рег}} \supseteq \Omega_{\text{авт}}$:
Покажем, что всякий язык, допускаемый НКА, является регулярным. Приведем доказательство индукцией по *n*, равному количеству всех «команд» в НКА, т.е. по общему количеству переходов из (q, a) в q′, где $q' \in \delta(q, a)$.

Б. И. *n* = 1. Пусть язык L допускается НКА (Q, Σ, δ, Q_0 , Q_F), имеющим только одну «команду» δ(q, a) = q′.

1) Если $q \notin Q_0$, $Q_0 \cap Q_F = \varnothing$, то L = ∅. Такой язык является регулярным.

2) Если $q \notin Q_0$, $Q_0 \cap Q_F \neq \varnothing$, то L = {ε}. Такой язык является регулярным.

3) Если $q \in Q_0$, $q' \notin Q_F$, $q \notin Q_F$ то L = ∅. Такой язык является регулярным.

4) Если $q \in Q_0$, $q' \notin Q_F$, $q \in Q_F$, то L = {ε}. Такой язык является регулярным.

5) Если $q \in Q_0$, $q' \in Q_F$, то допускаются следующие языки:

L = {a}	$q \neq q'$, $q \notin Q_F$
L = {ε, a}	$q \neq q'$, $q \in Q_F$
L = {a}*	$q = q'$

Все такие языки являются регулярными.

База индукции доказана.

Ш.И. n > 1. Пусть язык L допускается НКА (Q, Σ, δ, Q_0, Q_F), имеющим n «команд».
Пусть теорема верна для любого языка, допускаемого НКА с числом «команд» равным (n–1).
Зафиксируем одну команду-переход из (q, a) в q′.
Обозначим δ' функцию переходов, полученную из δ удалением зафиксированной команды.

Рассмотрим следующие автоматы:

$A' = (Q, \Sigma, \delta', Q_0, Q_F)$,

$A_1 = (Q, \Sigma, \delta', Q_0, \{q\})$,

$A_2 = (Q, \Sigma, \delta', \{q'\}, \{q\})$,

$A_3 = (Q, \Sigma, \delta, \{q'\}, Q_F)$.

Выберем слово $w \in L$.
Если при обработке w исходный НКА ни разу не применил зафиксированную команду, то $w \in L(A')$.

Если при обработке w зафиксированная команда применялась, то
$w \in L(A_1) \cdot (a \cdot L(A_2)^* \cdot a \cdot L(A_3))$.

первая буква			a				a			a			последняя буква	
нач. сост.			q	q′			q	q′		q	q′			заключ. сост.
			первое применение							последнее применение				

Следовательно, $w \in L(A') \cup L(A_1) \cdot (a \cdot L(A_2)^* \cdot a \cdot L(A_3))$.

Верно и включение в обратную сторону, если $w \in L(A') \cup L(A_1) \cdot (a \cdot L(A_2)^* \cdot a \cdot L(A_3))$, то $w \in L$.

Таким образом, $L = L(A') \cup L(A_1) \cdot (a \cdot L(A_2)^* \cdot a \cdot L(A_3))$.

По предположению индукции, языки $L(A'), L(A_1), L(A_2), L(A_3)$ являются регулярными. Следовательно, язык L тоже регулярный.

Шаг индукции доказан.
Теорема доказана.

Замечание:
Для описания регулярного языка используется регулярное выражение без фигурных скобок.

Например. Для $L = \{a\}^* \cdot (\{abba\} \cup \{aa\}) \cdot (\{b\} \cup \{b\}^*)$ используется $L = a^* \cdot (ab^2a \cup a^2) \cdot b^n$ или $L = a^*(ab^2a + a^2)b^n$.