

Лабораторная работа 2 (вариант 1)

1. Исходное регулярное выражение

Алфавит: $\Sigma = \{a, b\}$.

Регулярное выражение:

$$R = ((a|b)^* abb (a|b)) \mid ((a|ba)^* bb (a^*b)^* aba (a|b)).$$

2. НКА

Построенный НКА изображён на рис. 1.

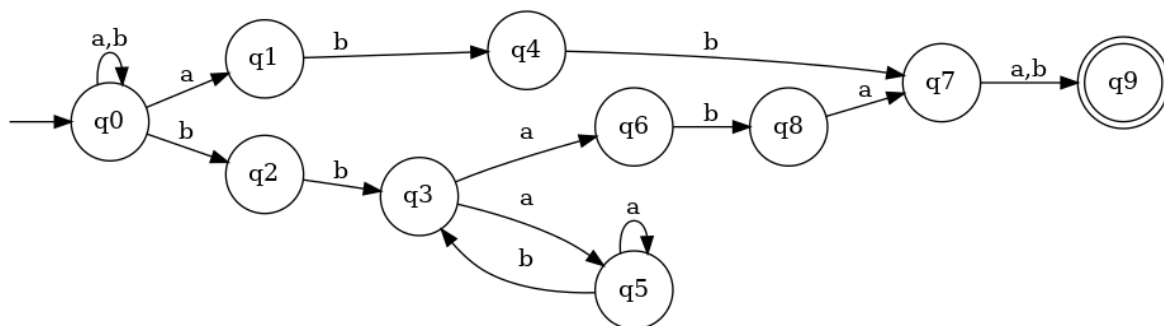


Рис. 1: НКА

2.1. Таблица префиксов и суффиксов НКА

	ba	ε	bba	b	abab	babaa	abba
bab	1	0	0	0	0	1	1
abba	0	1	1	0	0	1	1
a	0	0	1	0	0	0	1
abb	0	0	0	1	1	1	1
bb	0	0	0	0	1	1	1
b	0	0	0	0	0	1	1
ε	0	0	0	0	0	0	1

Таблица 1: Таблица префиксов и суффиксов НКА

3. ДКА и минимизация

На рис. 2 показан минимальный ДКА.

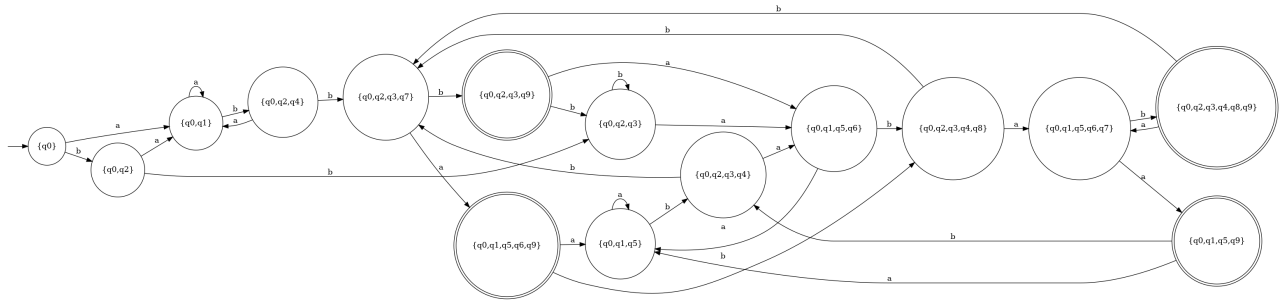


Рис. 2: минимальный ДКА

3.1. Классы эквивалентности

Отметим классы эквивалентности: рис. 3.

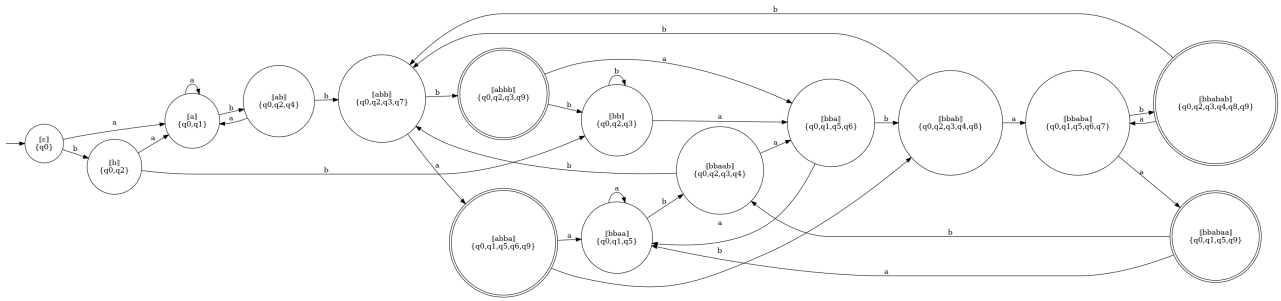


Рис. 3: Минимальный ДКА

3.2. Проверка минимальности

Построим таблицу эквивалентности. Если все строки различны, то состояния попарно различны, значит ДКА минимален.

	ε	a	aa	abaa	abba	ba	baa	babaa	bba
ε	0	0	0	0	1	0	0	0	0
a	0	0	0	0	1	0	0	0	1
b	0	0	0	0	1	0	0	1	0
ab	0	0	0	0	0	1	0	1	0
abb	0	1	0	1	1	0	0	1	0
abba	1	0	0	0	0	0	1	1	1
abbb	1	0	0	1	1	0	0	1	0
bb	0	0	0	1	1	0	0	1	0
bba	0	0	0	0	1	0	0	1	1
bbaa	0	0	0	0	1	0	1	1	1
bbaab	0	0	0	1	1	1	0	1	0
bbab	0	0	1	1	1	1	0	1	0
bbaba	0	1	0	0	1	0	1	1	1
bbabaa	1	0	0	0	1	0	0	1	1
bbabab	1	0	1	1	1	1	1	1	0

Таблица 2: Таблица различимости

Все строки различны, значит он минимален.

4. Расширенное регулярное выражение и ПКА

4.1. Расширенная регулярка

Исходное выражение:

$$R = ((a|b)^* abb (a|b)) \mid ((a|ba)^* bb (a^*b)^* aba (a|b)).$$

Перепишем его в расширенном виде:

$$R = ^{.}(.^{*}((?<=abb) \mid (?=bb)^{.} (?<=baba))^{.})^{.}$$

Внешняя структура имеет вид

$$.^{*} \alpha ^{.}$$

Первая часть $.^{*}$ задаёт произвольный префикс слова. Последний символ $^{.}$ соответствует одному произвольному символу алфавита.

Внутреннее выражение содержит две альтернативы.

Первая альтернатива: $(?<=abb)^{.}$. Она требует, чтобы непосредственно перед текущей позицией стояло подслово **abb**. Поскольку далее следует последний символ, всё слово оканчивается на **abbx**, где $x \in \{a, b\}$. То есть, его суффикс равен **abba** или **abbb**.

Вторая альтернатива: $(?=bb)^{.} (?<=baba)^{.}$.

Конструкция $(?=bb)$ требует, чтобы начиная с текущей позиции в слове стояло подслово **bb**. Тем самым фиксируется существование некоторого вхождения **bb**.

После этого допускается произвольная последовательность символов $.^{*}$, а затем $(?<=baba)^{.}$ требует, чтобы непосредственно перед последним символом стояло подслово **baba**.

Следовательно, во второй альтернативе слово обязательно содержит подслово **bb** и оканчивается на **babax**, где $x \in \{a, b\}$.

Таким образом, язык состоит из всех слов над $\{a, b\}$, которые либо имеют суффикс **abb(a|b)**, либо содержат подслово **bb** и одновременно имеют суффикс **baba(a|b)**.

4.2. ПКА

ПКА для того же языка изображён на рис. 4.

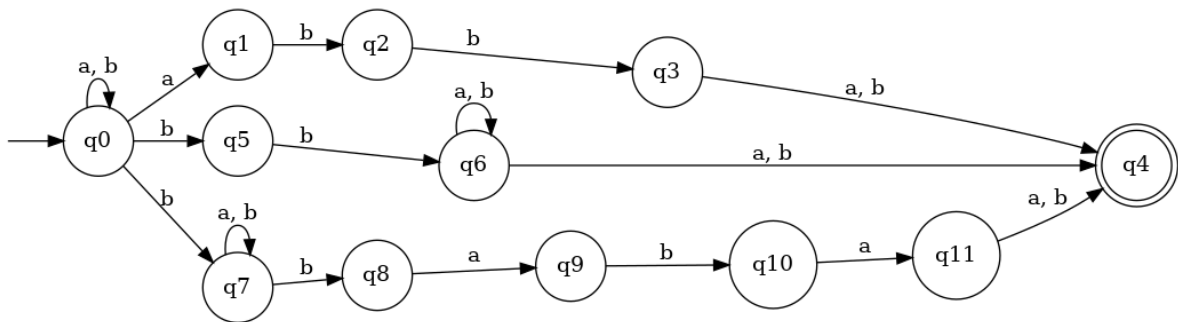


Рис. 4: ПКА

	ε	a	abaa	ba	baa	babaa
bb	0	1	1	1	1	1
bab	0	0	1	1	1	1
b	0	0	0	1	1	1
baba	0	0	0	0	1	1
ba	0	0	0	0	0	1
a	0	0	0	0	0	0

Таблица 3: Таблица префиксов и суффиксов для ПКА