

Отчет по лабораторной работе №2

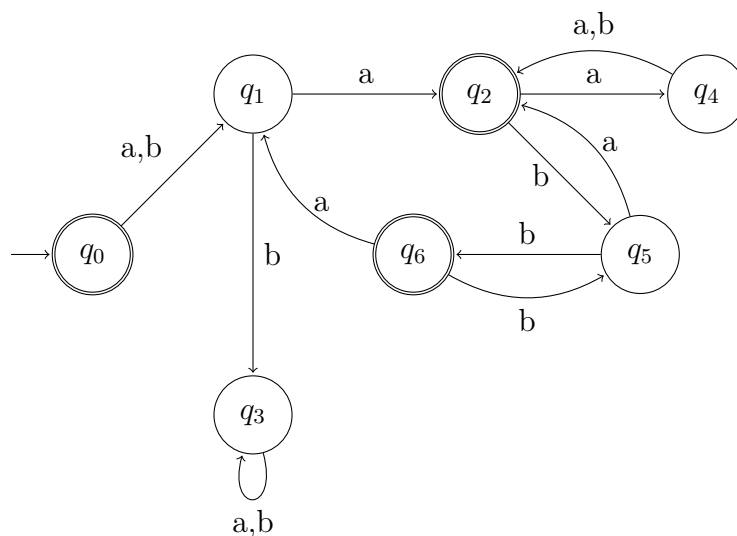
Гуров Вячеслав ИУ9-52Б

Вариант 8

Исходное регулярное выражение

$$((aa|ba)(ab)^*(bb)^*)^*$$

Детерминированный конечный автомат



Описание автомата:

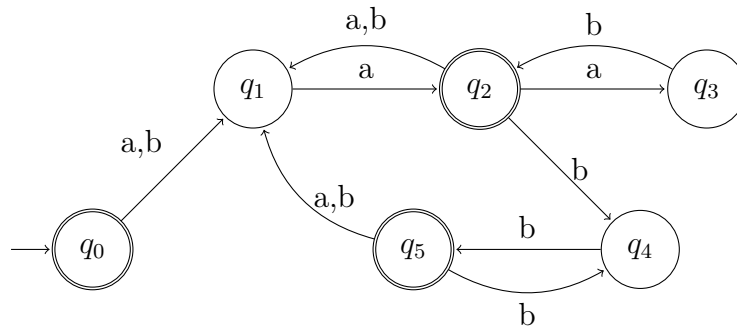
- Начальное состояние: q_0
- Конечные состояния: q_0, q_2, q_6
- Алфавит: $\{a, b\}$

Проверка минимальности ДКА

Состояние/Слово	ε	a	b	bab	bb	ab
$(q_0)\varepsilon$	+	-	-	-	-	-
$(q_1)a$	-	+	-	-	-	-
$(q_2)aa$	+	-	-	-	+	+
$(q_3)ab$	-	-	-	-	-	-
$(q_4)aaa$	-	+	+	+	-	-
$(q_5)aab$	-	+	+	-	-	-
$(q_6)aabb$	+	-	-	-	+	-

Все строки попарно различны, следовательно автомат минимален.

Малый НКА



Описание автомата:

- Начальное состояние: q_0
- Конечные состояния: q_0, q_2, q_5
- Алфавит: $\{a, b\}$

Таблица множеств классов эквивалентности:

слово/суффикс	ab	bb	bab	b	a	ε
aa	+	+	-	-	-	+
aabb	-	+	-	-	-	+
aaa	-	-	+	+	+	-
aab	-	-	-	+	+	-
a	-	-	-	-	+	-
ε	-	-	-	-	-	+

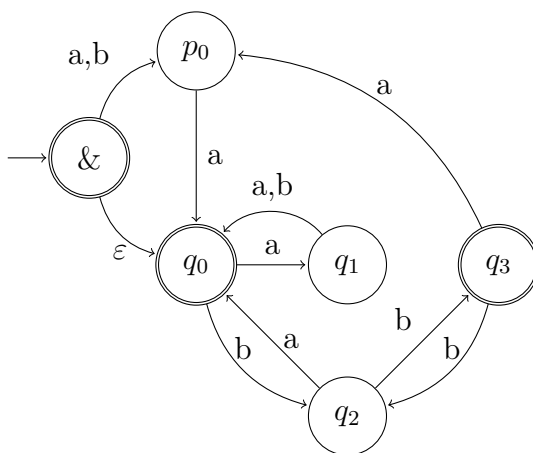
Малый ПКА

Более сложные пересечения языков приводят к автомату с большим числом состояний. Минимальный из рассмотренных автоматов был построен для следующего случая.

Исходный язык, заданный регулярным выражением, можно представить как пересечение двух языков:

1. Исходный язык.
2. Исходный язык, который не обязан начинаться с aa или ba (т. е. язык со словами четной длины, в котором после четного количества букв не может идти $bbab$).

Можно также пересечь 2 пункт с языком, в котором слова начинаются на ab или aa , но это даст $+1$ состояние.



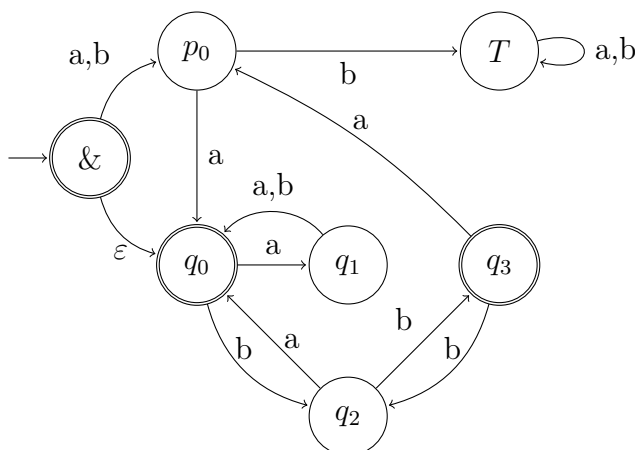
Описание автомата:

- Начальное состояние: $\&$
- Конечные состояния: $\&$, q_0 , q_3
- Алфавит: $\{a, b\}$

Таблица множеств классов эквивалентности:

слово/суффикс	ε	bab	b	a
aaa	-	+	+	+
aab	-	-	+	+
a	-	-	-	+
ab	-	-	-	-

Если необходим детерменизм всех состояний кроме состояния пересечения, то нужно добавить состояние ловушку:



Расширенное регулярное выражение

Заметим, что $(aa|ba)$ эквивалентно $(a|b)a$, а $(a|b)$ можно записать как $[ab]$. Тогда расширенное регулярное выражение равно:

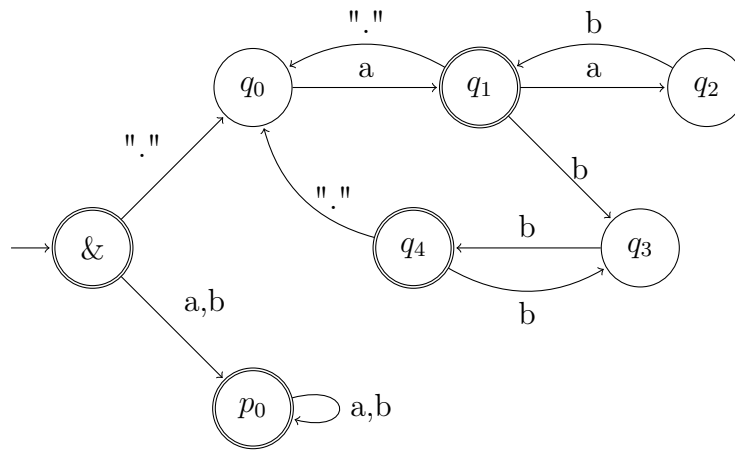
$^([ab]a(ab)^*(bb)^*)^*\$$

Можно пойти дальше и заменить $[ab]$ на wildcard-операцию $"."$, но тогда необходима операция предпросмотра, чтобы гарантировать, что всё слово состоит только из символов a и b :

$^(?=[ab]^*\$)(.a(ab)^*(bb)^*)^*\$$

Данная запись использует положительный lookahead $(?=[ab]^*\$)$ для проверки, что все символы в строке принадлежат множеству $\{a, b\}$, и wildcard операцию $"."$ для сокращения записи.

Новое регулярное выражение можно описать автоматом:



Описание автомата:

- Начальное состояние: $&$
- Конечные состояния: $&$, q_1 , q_4 , p_0
- $"."$ - любой символ