

Теоретическая информатика III

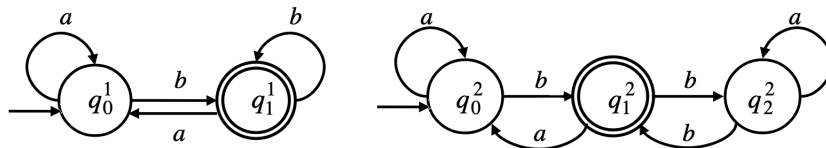
Задачи для семинара 11/14 сентября

1 Построение регулярных выражений

1. Построить регулярное выражение для множества всех
 - (a) строк над алфавитом $\{a\}$, длина которых не делится ни на 2, ни на 3.
 - (b) строк над алфавитом $\{a, b\}$, содержащих не более одной пары последовательно идущих символов a .
 - (c) построить регулярные выражения для языков $K \cup L$, $K \cap L$ и $K \setminus L$, где языки K и L задаются регулярными выражениями a^*ba^* и b^*ab^* (для краткости пишут $K = a^*ba^*$ и $L = b^*ab^*$).
2. Пусть $\Sigma = \{0, \dots, 9, e, ., -\}$ — алфавит символов, используемых для представления чисел в языках программирования. Например, строкой $-1.2e-4$ записывается число $-1.2 \cdot 10^{-4}$. Запись числа может начинаться со знака минус (а может с него не начинаться), и потом идёт мантисса, в которой есть хотя бы одна цифра и может быть не более одной точки. За мантиссой, возможно, следует экспонента следующего вида: сперва буква e , потом, возможно, знак минус, и наконец целое число. Построить регулярное выражение, задающее множество всех строк, являющихся правильными записями чисел.

2 Построение NFA

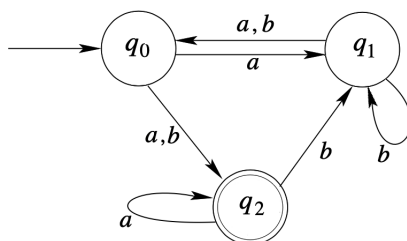
1. Построить NFA как можно меньшего размера (под размером понимается число состояний, доказывать минимальность не требуется), распознающий язык
 - (a) $\{a^m b^n \mid \text{существует число } k \in \{2, 3, 5\}, \text{ на которое делятся и } m, \text{ и } n\}$.
 - (b) всех строк над алфавитом $\{a, b\}$, в которых каждые 4 или каждые 5 символов встречается символ a : то есть, существует период $p \in \{4, 5\}$, и смещение i , так что в каждой позиции с номером, эквивалентным i по модулю p , находится символ a . Короткие строки длины менее чем 5 считать принадлежащими языку.
Это задача распознавания периодичных сигналов с периодом, известным только приблизительно. Например, строка $abaababbbabab$ удовлетворяет условию для $p = 4$ и $i = 4$. Для $p = 5$ условие не выполнится, но это неважно: строка принадлежит языку.
2. Пусть языки K и L распознаются автоматами, представленными ниже. Построить NFA, распознающий конкатенацию этих языков.



3 Основные преобразования

Для решения задач этой секции используйте формальные построения.

1. Применить построение подмножеств к автомату, распознающему множество всех строк над алфавитом $\{a, b\}$, в которых есть пара символов a , между которыми находятся ровно 1 символ.
2. Для регулярного выражения $b(a^*b \mid c)^*$ записать равносильный ему ε -NFA. Затем удалить из него ε -переходы.
3. Для регулярного выражения $(a \mid ab)^*b$ построить равносильный ему ε -NFA, перевести его в NFA и затем в DFA.
4. Постройте регулярное выражение для языка, распознаваемого данным автоматом.



5. Построить DFA, распознающий множество всех строк над алфавитом $\Sigma = \{a, b\}$, в которых число символов a и число символов b дают разные остатки при делении на 3. Затем перевести его в RE-NFA и преобразовать в регулярное выражение.

4 Замкнутость

Верно ли, что множество регулярных языков замкнуто относительно:

1. обращения всех строк $L^R = \{w \mid w^R = w_n \cdots w_1 \in L\}$
2. циклического замыкания $\text{cycle}(L) = \{w \mid w = xy, xy \in L\}$
Пример: $\text{cycle}(\{01, 010\}) = \{01, 10, 010, 001, 100\}$.
- 3.* операции "взятия половинок языка":

$$\frac{1}{2}L = \{x \in \Sigma^* \mid \exists y \in \Sigma^* : xy \in L, |x| = |y|\}$$

Пример: $\frac{1}{2}\{0, 11, 00111, 011001\} = \{1, 011\}$.

5* Синхронизация

DFA $(\Sigma, Q, q_0, \delta, F)$ называется *синхронизируемым*, если существует такая синхронизирующая строка $w \in \Sigma^*$ и такое состояние $q \in Q$, что, начав вычисление на w из любого состояния $q \in Q$, автомат закончит чтение этой строки в состоянии \hat{q} .

Доказать, что если в синхронизируемом автомате n состояний, то для него есть синхронизирующая строка длины $O(n^3)$ — желательно, с константой поменьше.