Отделение лингвистики, 2013-14 уч. год
Дискретная математика
Конечные автоматы (23 апреля 2014)
Ю. Г. Кудряшов, И. В. Щуров, К. Г. Куюмжиян, Р. Я. Будылин

Определение 1. Рассмотрим конечный алфавит $\mathcal{A}$. Например, для простоты можно считать, что $\mathcal{A} = \{a, b\}$.

Словом над алфавитом $\mathcal{A}$ называется конечная цепочка букв. Например, $abba$ — слово над алфавитом $\{a, b\}$. Бывает «пустое слово», не содержащее ни одной буквы, оно обычно обозначается ε. Множество всех возможных слов над алфавитом $\mathcal{A}$ обозначается $\mathcal{A}^*$.

Языком над алфавитом $\mathcal{A}$ называется любое подмножество множества всех слов $\mathcal{A}^*$.

Определение 2. *Конечный автомат* — это такая машинка (точнее, математическая модель), которая имеет конечное число внутренних состояний, получает на вход слово из букв некоторого алфавита (буква за буквой) и сообщает в ответ, является ли это слово «допустимым».

Описать конечный автомат — значит задать алфавит, с которым он работает, множество всех его состояний, а также правило, которое по текущему состоянию автомата и поданной на вход букве определяет новое состояние автомата. Среди состояний выделяются одно стартовое (в нём автомат находится до того, как ему подали на вход первую букву), а также одно или несколько допускающих (завершающих) состояний. Если автомат после получения последней буквы слова оказался в одном из допускающих состояний, слово считается допустимым, иначе недопустимым.

Таким образом, конечный автомат задаёт некоторый язык.

Задача 1. Для каждого из конечных автоматов, изображённых на рис. 1, ответьте на следующие вопросы.

(a) В каком состоянии автомат закончит работу, если на вход подать слово «free»?
(b) В каком состоянии автомат закончит работу, если на вход подать слово «referrer»?
(c) Выпишите несколько слов, для которых автомат закончит работу в одном из допускающих состояний.
(d) Выпишите несколько слов, для которых автомат закончит работу в одном из недопускающих состояний.

Задача 2. Для каждого из следующих языков найдите задающий их конечный автомат.

(a) Непустые слова, содержащие только букву «`a`» (одну или много) и никаких других букв.
(b) Слова, содержащие хотя бы три буквы «`a`» и не содержащие никаких других букв.
(c) Слова, содержащие чётное количество букв «`a`» (и, возможно, какое-то количество букв «`b`»).
(d) Слова, длина которых делится на 3.
(e) Слова, начинающиеся и заканчивающиеся на букву «`b`».
(f) Слова, в которых встречаются буквы «`a`» и «`b`», причём первое вхождение «`a`» встречается раньше, чем первое вхождение «`b`».
(g) Слова, в которых запрещено ставить три буквы «`a`» подряд.
(h) Слова, в которых встречается хотя бы раз подслово «`aba`».

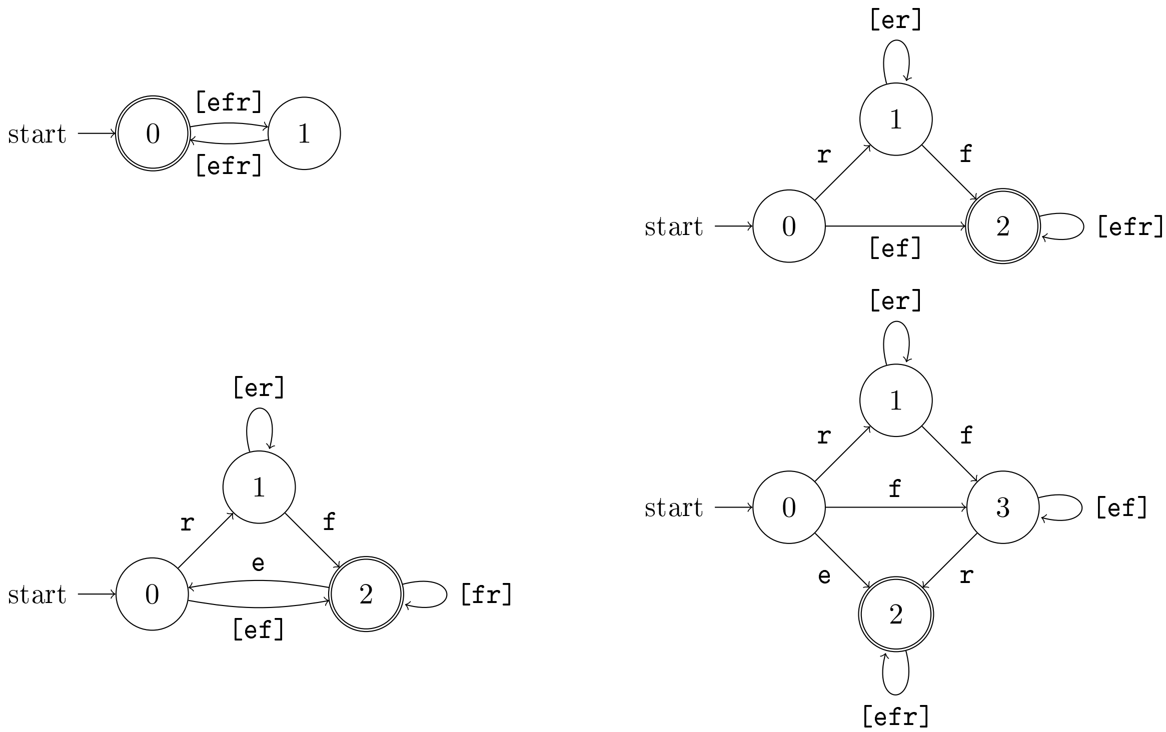

Рис. 1: К задаче 1

Обозначения:

start — начальное состояние автомата;

недопускающее состояние автомата: автомат возвращает `FALSE`;

допускающее состояние автомата: автомат возвращает `TRUE`.

`[abc]` если рядом с ребром написано несколько букв в квадратных скобках, значит, по этому ребру можно пройти по любой букве из перечисленных. Иными словами, можно считать, что мы так сокращённо рисуем несколько рёбер с одинаковыми стартовой и конечной вершинами.

(i) Слова в алфавите из трёх букв {`a`,`b`,`c`}, в которых встречается хотя бы раз подслово «`abac`».

Задача 3. Доказать, что язык, содержащий только слова, в котором букв «`a`» больше, чем букв «`b`», невозможно задать никаким конечным автоматом.