Формальные языки

домашнее задание до 23:59 05.03

1. Доказать или опровергнуть утверждение: произведение двух минимальных автоматов всегда дает минимальный автомат (рассмотреть случаи для пересечения, объединения и разности языков).

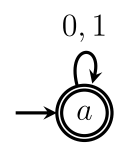
   Решение:

   Утверждение неверно во всех случаях.

   (a) Объединение

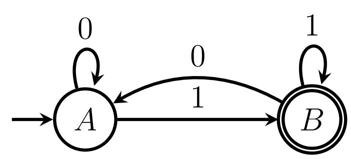
   Первый автомат задает весь язык

   (1)

   Второй автомат задает слова, заканчивающиеся на единичку

   (2)

   Произведение будет содержать две вершины и задавать автомат, который принимает весь язык. Он задается автоматом из одной вершины, противоречие.

   (b) Пересечение

   Возьмем два языка, которые не пересекаются.

   Например язык (2) и язык, который получается если в автомате (2) на всех ребрах инвертировать биты, получим два автомата.

   Автомат произведения будет содержать 2 вершины, а минимальный - одну, которая все отвергает.

   (c) Разность

   Возьмем два раза автомат (2). Разность таких языков очевидно задает пустой язык, который принимает автомат из одной вершины, но автомат произведения (2) * (2) содержит две.

   Итого рассмотрели все случаи. (*) везде выше автомат содержит вершины, если они достижимы.

2. Для регулярного выражения:

$$(a \mid b)^+(aa \mid bb \mid abab \mid baba)^*(a \mid b)^+$$

   Построить эквивалентные:

(a) Недетерминированный конечный автомат

(b) Недетерминированный конечный автомат без ε-переходов

(c) Минимальный полный детерминированный конечный автомат

Решение:

Докажем, что регулярка на самом деле сильно проще.

Утверждается, что $(a \mid b)^+(aa \mid bb \mid abab \mid baba)^*(a \mid b)^+ \equiv (a \mid b)^+(a \mid b)^+$

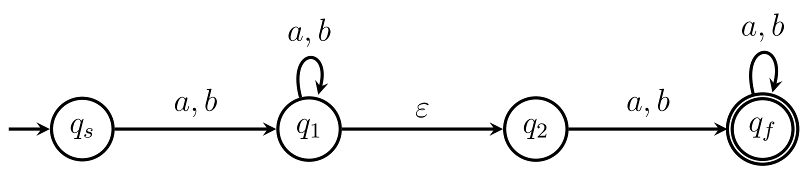
Очевидно $(a \mid b)^+(aa \mid bb \mid abab \mid baba)^*(a \mid b)^+ \supset (a \mid b)^+(a \mid b)^+$

В обратную сторону. Мы можем построить НКА для $(a \mid b)^+(aa \mid bb \mid abab \mid baba)^*(a \mid b)^+$, в котором отдельно будет первая черная часть, затем красная, затем вторая черная и связаны они будут только ε-переходами между соседними частями. Будем подкрашивать части слов, которое получаются прохождением через красную часть автомата в красный соответственно. Заметим, что если красных букв в слове вообще нет, то оно очевидно лежит в языке $(a \mid b)^+(a \mid b)^+$. Если же есть красная часть, она однозначно разбивается на блоки $aa \mid bb \mid abab \mid baba$. Разбили на такие блоки, пронумеровали каждый блок. Теперь по индукции можем доказать, что для каждого подкрашенного слова можно получить такое же черное на этом автомате. Базу рассмотрели выше, переход: для каждого слова $(black^*)(red_1)(red_2)...(red_n)(black^*)$ можно этим же автоматом получить слово $(black^*)(red_1)(red_2)...(red_n)(black^*)$. Доказывается это перебором всех 4 случаев для блоков. (техническая, очевидная часть).

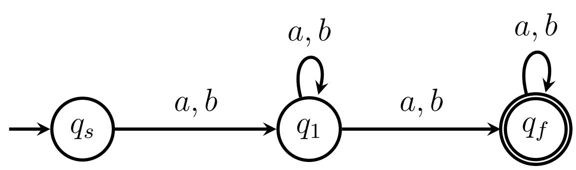
Итого доказали, что красную часть автомата можно выкинуть, заменив на пустые переходы. Таким образом д-ли эквивалентность языков. (Вообще, вроде сразу видно, что они эквивалентны, но если руками махать не хочется, такое доказательство должно работать).

Решение для $(a \mid b)^+(a \mid b)^+$

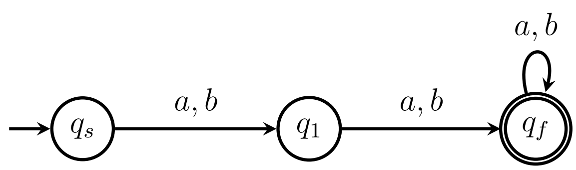
(a)



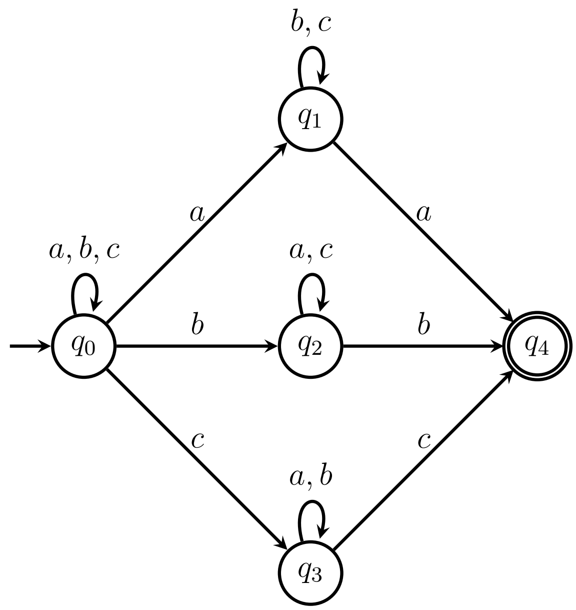
(b)



(c)



3. Построить регулярное выражение, распознающее тот же язык, что и автомат:



Решение:

$(a \mid b \mid c)^*((a(b \mid c)^*a) \mid (b(a \mid c)^*b) \mid (c(a \mid b)^*c))$

4. Определить, является ли автоматным язык $\{\omega\omega^r \mid \omega \in \{0,1\}^*\}$. Если является — построить автомат, иначе — доказать.

   Решение:

   Воспользуемся Леммой о накачке, чтобы доказать нерегулярность языка. Пусть язык регулярный, тогда по лемме о накачке найдетсу n, т.ч слово $\omega = 0^n1^n1^n0^n$ можно разбить и накачать. Пусть $xyz = \omega, |xy| < n, y \neq \varepsilon$, тогда $y = 0^b$. Возьмем $k = 3$, тогда максимальный префикс из нулей для слова xy^kz равен $n - b + 3 * b = n + 2 * b > n$, а максимальный суффикс из нулей имеет размер n, что неверно ни для одного слова из языка. Противоречие.

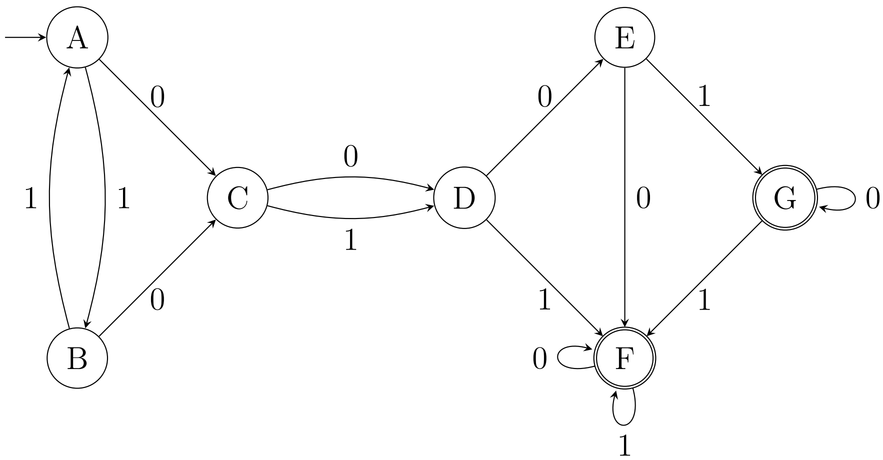
5. Определить, является ли автоматным язык $\{uaav \mid u, v \in \{a,b\}^*, |u|_b \geq |v|_a\}$. Если является — построить автомат, иначе — доказать.

Решение:

Воспользуемся Леммой о накачке, чтобы доказать нерегулярность языка. Возьмем слово $\omega = b^naa(ba)^n$, тогда $y = b^m, m > 0$. Рассмотрим $k = 0$ и слово xz. В нем только одно место, где есть две ашки подряд, значит разделить на u и v можно только так, что слева будет $n-m$ b-шек, а справа n. Противоречие.

Пример применения алгоритма минимизации

Минимизируем данный автомат:



Автомат полный, в нем нет недостижимых вершин — продолжаем.

Строим обратное δ отображение.

δ^{-1}	0	1
A	—	B
B	—	A
C	A B	—
D	C	C
E	D	—
F	E F	D F G
G	G	E

Отмечаем в таблице и добавляем в очередь пары состояний, различаемых словом ε: все пары, один элемент которых — терминальное состояние, а второй — не терминальное состояние. Для данного автомата это пары

$(A, F), (B, F), (C, F), (D, F), (E, F), (A, G), (B, G), (C, G), (D, G), (E, G)$

Дальше итерируем процесс определения неэквивалентных состояний, пока очередь не оказывается пуста.

(A, F) не дает нам новых неэквивалентных пар. Для (B, F) находится 2 пары: $(A, D), (A, G)$. Первая пара не отмечена в таблице — отмечаем и добавляем в очередь. Вторая пара уже отмечена в таблице, значит, ничего делать не надо. Переходим к следующей паре из очереди. Итерируем дальше, пока очередь не опустошится.

Результирующая таблица (заполнен только треугольник, потому что остальное симметрично) и порядок добавления пар в очередь.

	A	B	C	D	E	F	G
A							
B							
C	✓	✓					
D	✓	✓	✓				
E	✓	✓	✓	✓			
F	✓	✓	✓	✓	✓		
G	✓	✓	✓	✓	✓		

Очередь:
$(A, F), (B, F), (C, F), (D, F), (E, F), (A, G), (B, G), (C, G), (D, G), (E, G),$
$(B, D), (A, D), (A, E), (B, E), (C, E), (C, D), (D, E), (A, C), (B, C))$

В таблице выделились классы эквивалентных вершин: $\{A, B\}, \{C\}, \{D\}, \{E\}, \{F, G\}$. Остается только нарисовать результирующий автомат с вершинами-классами. Переходы добавляются тогда, когда из какого-нибудь состояния первого класса есть переход в какое-нибудь состояние второго класса. Минимизированный автомат: