Задача по конечным автоматам и регулярным языкам

Прежде, чем решать задачу, необходимо тщательно проработать по учебнику (Белоусов А.И., Ткачев С.Б. Дискретная математика. –5-е изд. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2015; далее просто Учебник) теоретический материал. Это: глава 7, параграфы: 7.1 (кусок про лексикографическую нумерацию можно опустить), 7.4, 7.5 (теорему 7.5 можно не читать), 7.6. Некоторые теоретические положения будут изложены здесь. Также будут рассмотрены некоторые приемы решения систем линейных уравнений, которые существенно упрощают вид решения и которые не отражены в учебнике. Решетка (#) обрамляет части текста, которые можно пропустить без ущерба для понимания решения задачи.

Один пример анализа конечного автомата (КА)

Рассмотрим КА на рис. 1.

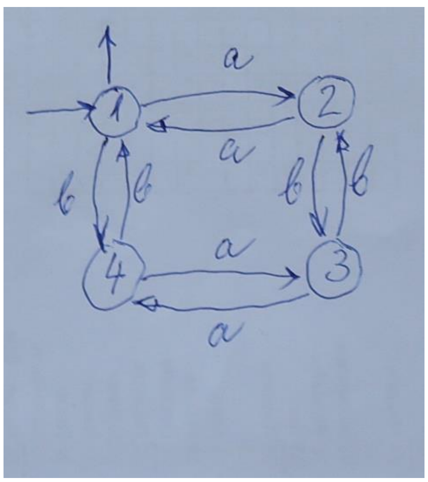

Рис. 1

Используя метод математической индукции, можно доказать, что он допускает те и только те цепочки (слова) в алфавите $\{a,b\}$, которые содержат произвольное четное число вхождений каждого символа (пустое слово считается допустимым: оно содержит нулевое число вхождений каждого символа).

Индукцию проводим по длине слова, прочитав которое автомат останавливается в состоянии 1.

Базис тривиален – пустое слово. Пусть автомат прочитал слово x, обладающее указанным свойством, и достиг состояния 1 (оно и начальное, и заключительное). Чтобы вернуться снова в это состояние, пройдя при этом некий контур, то есть простой замкнутый путь, он должен прочитать либо слова aa, bb, либо $abab, baba$. В любом случае свойство после этого сохранится. #

Выведем регулярное выражение для языка в алфавите $\{a,b\}$, состоящее из всех слов с указанным выше свойством.

Составим систему уравнений:

$$\begin{cases} x_1 = ax_2 + bx_4 + \lambda \\ x_2 = ax_1 + bx_3 \\ x_3 = bx_2 + ax_4 \\ x_4 = bx_1 + ax_3 \end{cases} \qquad (1)$$

В этой системе слагаемое λ стоит в правой части первого уравнения, так как первое состояние заключительное. Поскольку оно также и начальное, то интересующий нас язык L, допускаемый этим автоматом, дается значением неизвестного x_1.

Возникает вопрос: как решать систему?

Заметим, если «ортодоксально» использовать метод последовательного исключения неизвестных, получится очень сложное («нечитаемое») регулярное выражение. Надо как-то упростить процесс решения и сам вид решения.

Глядя внимательно на систему (1), мы замечаем, что второе и четвертое неизвестное уже в системе выражены через первое и третье. Поэтому уже на первой итерации мы можем исключить два неизвестных, получив систему двух уравнений с двумя неизвестными:

$$\begin{cases} x_1 = a(ax_1 + bx_3) + b(bx_1 + ax_3) + \lambda \\ x_3 = b(ax_1 + bx_3) + a(bx_1 + ax_3) \end{cases}$$

Приводя подобные члены в правых частях, получим

$$\begin{cases} x_1 = (aa + bb)x_1 + (ab + ba)x_3 + \lambda \\ x_3 = (ab + ba)x_1 + (aa + bb)x_3 \end{cases} \qquad (2)$$

Важно: слагаемые переставлять можно, но сомножители нельзя ни в коем случае! Сложение в полукольце регулярных языков коммутативно (это объединение множеств, языков как множеств слов), а умножение некоммутативно (это соединение языков).

Нам нужно окончательное выражение для неизвестного x_1. Разумно тогда из второго уравнения системы (2) выразить x_3 через x_1:

$$x_3 = (aa + bb)^*(ab + ba)x_1 \ .$$

Полученное выражение подставим в первое уравнение системы (2):

$$x_1 = (aa + bb)x_1 + (ab + ba)(aa + bb)^*(ab + ba)x_1 + \lambda,$$
$$x_1 = [(aa + bb) + (ab + ba)(aa + bb)^*(ab + ba)]x_1 + \lambda$$

Из последнего уравнения получаем решение в виде:

$$x_1 = [(aa + bb) + (ab + ba)(aa + bb)^*(ab + ba)]^* \cdot \lambda \qquad (3)$$

Используя известное регулярное тождество $(\alpha+\beta)^*=(\alpha^*\beta^*)^*$, где α и β - произвольные регулярные выражения (см. задачу 7.11 на стр. 580), а также учитывая, что сомножитель λ можно не писать (как не пишем мы сомножитель 1 в обычной арифметике), решение (3) можно представить в виде:

$$x_1=[(aa+bb)^*((ab+ba)(aa+bb)^*(ab+ba))^*]^* \qquad (4)$$

Выражение (4) позволяет понять структуру слова, содержащего произвольное четное число каждого символа.

Именно: такое слово может начинаться произвольным чередованием парных вхождений aa или bb, потом, скажем, происходит нарушение четности, то есть появляется цепочка ab или ba. Нарушение четности должно быть уравновешено, четность должна быть восстановлена, а между нарушением четности и ее восстановлением может находиться «прослойка» парных вхождений. Таким образом, мы имеем две «зоны»: зона начальных парных вхождений и зона, построенная по схеме: нарушение четности – парные вхождения – восстановление четности. Зона может как отсутствовать, так и повторяться любое число раз; чередование зон может происходить сколько угодно раз, в том числе и ни разу (пустое слово). Выражение (4) может быть использовано, как своего рода «склад» всех таких цепочек: читая его, можно породить произвольную цепочку со сформулированным свойством. Например:

$$\underbrace{aaaabb}_{1\,зона}\underbrace{abbbaaab}_{2\,зона}\underset{2\,зона}{baab}\ \underset{1\,зона}{bb}\ \underbrace{babbab}_{2\,зона}$$

Перед тем, как рассмотреть решение задачи домашнего задания, сделаем еще одно важное замечание.

Решая в замкнутом полукольце (или полукольце с итерацией) линейное уравнение

$x=\alpha x+\beta$,

решение записываем в виде

$x=\alpha^*\beta$

и ни в коем случае не в виде

$x=\alpha^*+\beta$

Это же относится к ситуации, когда какое-то неизвестное выражается через остальные. Указанная выше грубая ошибка встречается не так редко.

Задача по образцу ДЗ

Исходные данные:

1) конечный автомат (КА), заданный набором $M=(Q_s,Q_f,E)$, где

$Q_s = \{2\}$ - множество начальных состояний (входов);

$Q_f = \{4, 5\}$ - множество заключительных состояний (выходов);

$E = \{(1,2,b),(1,3,a),(1,5,b),(2,5,a),(2,4,b),(2,1,a),(3,2,b),(4,3,a,b),(5,4,a)\}$ - множество дуг с метками; метка каждой дуги – некоторое множество букв из алфавита $\{a, b\}$.

Замечание. В некоторых источниках допускаются КА с несколькими начальными состояниями. Можно показать, что такой автомат преобразуется к эквивалентному КА с одним начальным состоянием и без λ-переходов.

2) Язык $L_0 = \{(ab)^m (ba)^n : m, n \geq 0\}$.

Требуется:

1) найти язык L, допускаемый КА $M = (Q_s, Q_f, E)$;

2) детерминизировать КА $M = (Q_s, Q_f, E)$;

3) построить графы автоматов, допускающих языки $L \cup L_0, L \cdot L_0, L^*$ и из полученных графов удалить λ-переходы.

Таким образом, в задании три части: 1) анализ (определение языка КА), 2) детерминизация и 3) синтез некоторых конструкций из автоматов.

1) Анализ

Граф исходного КА представлен на рис. 2.

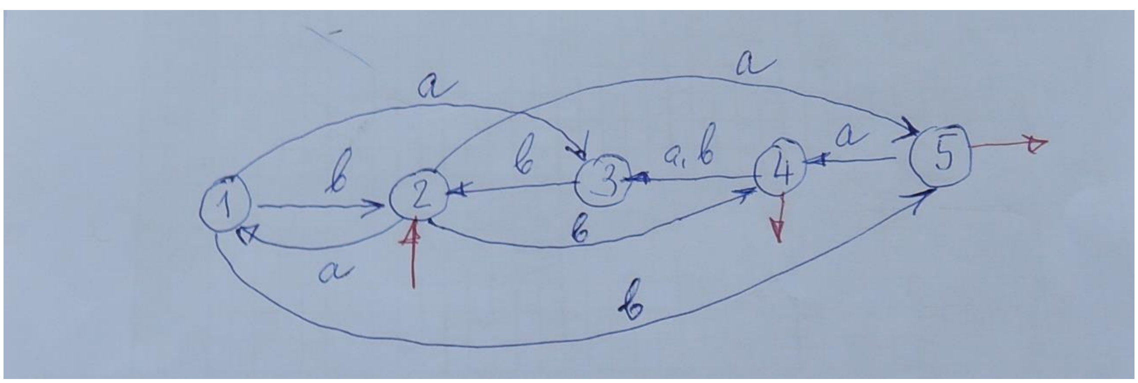

Рис. 2

Чтобы найти язык этого КА, необходимо (и достаточно) решить соответствующую систему линейных уравнений, а именно следующую:

$$
\begin{cases}
x_1 = bx_2 + ax_3 + bx_5 \\
x_2 = ax_1 + bx_4 + ax_5 \\
x_3 = bx_2 \\
x_4 = (a+b)x_3 + \lambda \\
x_5 = ax_4 + \lambda
\end{cases}
\qquad (5)
$$

Принцип, по которому составляется система: коэффициент α_{ij} в уравнении

$x_i = \ldots + \alpha_{ij} x_j + \ldots$

Равен метке дуги, ведущей из i-й вершины в j-ю. Если такой дуги нет, соответствующее слагаемое в правой части отсутствует. Метка дуги – простейший регулярный язык, состоящий из конечного множества однобуквенных цепочек. Записывается в виде регулярного выражения. В уравнениях, номера которых совпадают с номерами заключительных состояний (вершин), записывается в правой части слагаемое λ (регулярное выражение, обозначающее язык, состоящий из одного пустого слова). Чтобы найти язык КА, достаточно найти регулярное выражение для неизвестного, номер которого совпадает с номером начального состояния. В нашем случае это x_2.

Рекомендуется на каждой итерации исключать одно неизвестное, и ни в коем случае не писать преобразований правых частей в строчку. Иначе можно запутаться и получить неудобочитаемое выражение. Кроме того, рекомендуется проверять коэффициенты преобразованной системы способом, который будет описан ниже.

В системе (5) исключаем x_3:

$$
\begin{cases}
x_1 = bx_2 + abx_2 + bx_5 \\
x_2 = ax_1 + bx_4 + ax_5 \\
x_4 = (a+b)bx_2 + \lambda \\
x_5 = ax_4 + \lambda
\end{cases}
$$

Теперь исключаем x_5 и заодно приводим подобные члены в правых частях:

$$
\begin{cases}
x_1 = (b+ab)x_2 + bax_4 + b \\
x_2 = ax_1 + (b+a^2)x_4 + a \\
x_4 = (a+b)bx_2 + \lambda
\end{cases}
\qquad (6)
$$

Вот в системе (6) имеет смысл проверить коэффициенты. Именно, регулярное выражение при x_j в правой части уравнения для x_i должно обозначать язык, состоящих только из тех слов, которые могут быть прочитаны на некотором пути из i-й вершины в j-ю, хотя и не обязательно (на промежуточной итерации) из *всех* таких слов. Свободный член дает множество цепочек, читая которые, автомат переходит из i-го состояния в какое-то заключительное.

Посмотрев на граф, мы видим, что на пути из 1 в 2 читаются цепочка b (на дуге) и цепочка ab на пути 1→ 3→2. На пути из 1 в 4 1→5→4 читается цепочка ba. По букве b из вершины 1 можно перейти в 5 и выйти. Остальные коэффициенты рекомендуем проверить самостоятельно. Такая проверка поможет избежать ошибок в преобразовании системы, а также вовремя эти ошибки замечать и исправлять.

На следующей итерации исключаем x_4:

$$\begin{cases} x_1 = (b+ab)x_2 + ba((a+b)bx_2 + \lambda) + b \\ x_2 = ax_1 + (b+a^2)((a+b)bx_2 + \lambda) + a \end{cases}$$

После раскрытия скобок и приведения подобных членов получим:

$$\begin{cases} x_1 = (b+ab+ba^2b+bab^2)x_2 + ba + b \\ x_2 = ax_1 + (bab+b^3+a^3b+a^2b^2)x_2 + b + a^2 + a \end{cases} \tag{7}$$

Рекомендуется проверить коэффициенты и свободные члены. Покажем на примере коэффициента при x_2 во втором уравнении системы (7). Это регулярное выражение перечисляет метки замкнутых путей, проходящих через вершину 2.

$bab : 2 \underset{\rightarrow}{b} 4 \underset{\rightarrow}{a} 3 \underset{\rightarrow}{b} 2;$

$b^3 : 2 \underset{\rightarrow}{b} 4 \underset{\rightarrow}{b} 3 \underset{\rightarrow}{b} 2;$

$a^3b : 2 \underset{\rightarrow}{a} 5 \underset{\rightarrow}{a} 4 \underset{\rightarrow}{a} 3 \underset{\rightarrow}{b} 2;$

$a^2b^2 : 2 \underset{\rightarrow}{a} 5 \underset{\rightarrow}{a} 4 \underset{\rightarrow}{b} 3 \underset{\rightarrow}{b} 2.$

Обратим внимание на то, что дуга 4→3 имеет метку $a+b$, то есть из 4-го состояния в 3-е КА может перейти, прочитав с ленты любую из указанных двух букв. Поэтому обе цепочки bab и b^3 могут быть прочитаны на одном и том же пути. То же с другой парой цепочек из показанных выше четырех.

Выйти из вершины 2 можно, прочитав букву b на дуге 2→4, прочитав букву a на дуге 2→5, а также прочитав цепочку a^2 на пути 2→5→4.

Вернемся к системе (7). Мы видим, что неизвестное x_1 уже выражено через неизвестное x_2, значение которого и есть искомый язык. Подставляя во второе уравнение системы (7) выражение для x_1 из первого уравнения, получим уравнение для x_2:

$$\{ x_2 = (ab+a^2b+aba^2b+abab^2+bab+b^3+a^3b+a^2b^2)x_2 + aba + ab + b + a^2 + a \tag{8}$$

В коэффициенте при неизвестном здесь представлены метки замкнутых путей, проходящих через вершину 2 (рекомендуется проверить тщательно!), а свободный член содержит метки всех таких путей, ведущих из вершины 2 в какую-либо из заключительных (4 или 5), которые не содержат – как свою часть – ни одного замкнутого пути с вершиной 2. Заметим, что не все эти пути простые. Например, цепочка aba читается на пути 2→1→2→5, который не является простым.

Решая уравнение (8), получим окончательное выражение для языка данного КА:

$$L = x_2 = (ab + a^2b + aba^2b + abab^2 + bab + b^3 + a^3b + a^2b^2)^*(aba + ab + b + a^2 + a) \qquad (9)$$

Это выражение описывает множество всех слов (цепочек), которые могут быть прочитаны на каком-то пути из начального состояния в одно из заключительных. Замкнутый путь, метка которого стоит под знаком итерации, можно проходить сколько угодно раз; может быть, и не разу.

Обратим еще раз внимание на **недопустимость перестановки сомножителей в регулярном выражении!** Например, выражение aba^2b нельзя преобразовать к выражению a^3b^2. Переставлять можно только слагаемые.

Итак, первая часть задачи решена.

2) Детерминизация

Алгоритм детерминизации КА подробно описан в параграфе 7.6 учебника. Поскольку исходный КА не содержит λ-переходов, сразу строим эквивалентный детерминированный КА «методом вытягивания», вычисляя последовательно значения функции переходов δ' нового, детерминированного КА, по формуле:

$$\delta'(S, a) = \bigcup_{q \in S} \delta(q, a) \ , \qquad (10)$$

где δ - функция переходов исходного КА.

Начинаем с состояния-множества (см. пояснение этого условного термина на стр. 524 Учебника), соответствующего начальному состоянию. Если исходный КА имеет несколько начальных состояний, то в начальное состояние-множество войдут все исходные начальные состояния (и только они, разумеется)[1].

В нашем случае имеем одно начальное состояние и будем на первом шаге вычисления новой функции переходов иметь:

$$\delta'(\{2\}, a) = \{1,5\}, \delta'(\{2\}, b) = \{4\}.$$

[1] В детерминированном КА может быть только одно начальное состояние!

Вычисляемые последовательно состояния эквивалентного детерминированного КА (называемые условно состояниями-множествами) можно располагать в виде очереди. Тогда имеем:

$$\delta'(\{1,5\},a)=\delta(1,a)\cup\delta(5,a)=\{3\}\cup\{4\}=\{3,4\};$$
$$\delta'(\{1,5\},b)=\delta(1,b)\cup\delta(5,b)=\{2,5\}\cup\varnothing=\{2,5\}.$$

Пустое множество появилось потому, что в исходном КА из вершины 5 по букве b пойти никуда нельзя, то есть, формально, $\delta(5,b)=\varnothing$.

Дальше в очереди состояния-множества $\{4\}$, $\{3,4\}$, $\{2,5\}$.

Для них получим:

$$\delta'(\{4\},a)=\delta'(\{4\},b)=\{3\};$$
$$\delta'(\{3,4\},a)=\delta(3,a)\cup\delta(4,a)=\varnothing\cup\{3\}=\{3\};$$
$$\delta'(\{3,4\},b)=\delta(3,b)\cup\delta(4,b)=\{2\}\cup\{3\}=\{2,3\};$$
$$\delta'(\{2,5\},a)=\delta(2,a)\cup\delta(5,a)=\{1,5\}\cup\{4\}=\{1,4,5\};$$
$$\delta'(\{2,5\},b)=\delta(2,b)\cup\delta(5,b)=\{4\}\cup\varnothing=\{4\}.$$

Дальше вычисляем последовательно значения новой функции переходов для состояний-множеств $\{3\}$, $\{2,3\}$, $\{1,4,5\}$ (состояние $\{4\}$ уже обработано). Дальше опускаем подробности, но обратим внимание на то, что $\delta'(\{3\},a)=\varnothing$, то есть в графе детерминированного КА будет вершина, соответствующая пустому множеству, и ее следует «зациклить» по всем буквам алфавита, то есть нарисовать в ней петлю, меткой которой будут все буквы алфавита. В нашем случае их две: a и b. Ни в коем случае нельзя заключать символ пустого множества в фигурные скобки[2], а также повторять его столько раз, сколько оно возникает в виде значения новой функции переходов. Это относится ко всем вычисляемым состояниям-множествам: каждое такое состояние фигурирует в графе детерминированного КА ровно один раз.

Покажем еще для примера вычисление новой функции переходов для состояния $\{1,4,5\}$:

$$\delta'(\{1,4,5\},a)=\delta(1,a)\cup\delta(4,a)\cup\delta(5,a)=\{3\}\cup\{3\}\cup\{4\}=\{3,4\};$$
$$\delta'(\{1,4,5\},b)=\delta(1,b)\cup\delta(4,b)\cup\delta(5,b)=\{2,5\}\cup\{3\}\cup=\{2,3,5\}.$$

Продолжая подобным образом до тех пор, пока не перестанут появляться новые состояния-множества, получим граф, изображенный на рис. 3.

[2] Дело в том, что обозначение $\{\varnothing\}$ есть обозначение непустого множества, единственным элементом которого является пустое множество. Таких состояний в новом автомате быть не может.

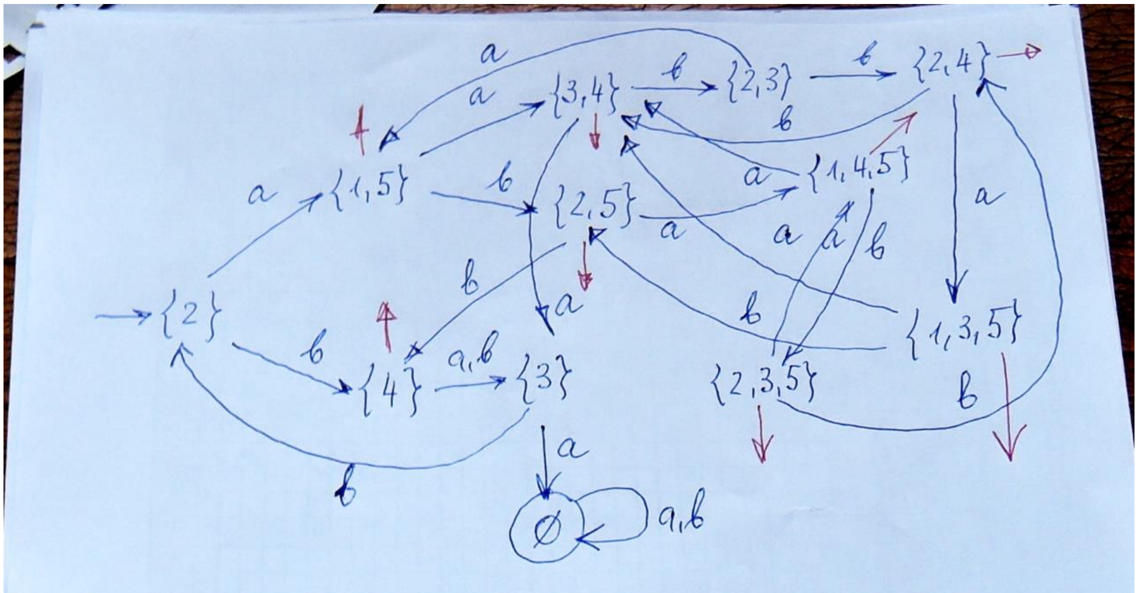

Рис. 3

Заключительные состояния-множества – те, которые содержат заключительное состояние исходного автомата. На рисунке отмечены красными выходными («висячими») стрелками.

Эта процедура требует большого внимания, и по завершении нужно проверить, что построен действительно детерминированный КА. Именно: из каждого состояния по каждой букве определен переход ровно в одно состояние. На графе это видно так: для каждой буквы алфавита из каждой вершины выходит ровно одна дуга, метка которой содержит эту букву.

Заметим еще, что теоретически в детерминированном эквивалентном КА будет 2^5=32 состояния, а мы «вытянули» 12, т.е. более, чем в два раза меньше. Остальные 20 недостижимы из начального состояния и никак не сказываются на языке, допускаемом автоматом. Статистически метод вытягивания дает число состояний, существенно меньшее 2^n, где n – число состояний исходного КА, но можно доказать, что существует такой недетерминированный КА, что в его детерминированном эквиваленте число состояний будет близко к 2^n. К счастью, такие ситуации весьма редки.

3) Конструкции (синтез)

В этой части задания надо построить автоматы для объединения языков $L \cup L_0$, соединения $L \circ L_0$ и итерации L^*.

Используем алгоритмы синтеза, как они изложены в Учебнике на стр. 514-517; доказательство теоремы 7.6) с одним отличием при построении КА для итерации (о чем подробно будет сказано ниже). После этого из построенных графов следует удалить дуги с меткой λ, называемые λ-переходами. Алгоритм удаления таких дуг рассмотрен в параграфе 7.6, но ниже будут сделаны некоторые замечания по этому поводу.

Прежде всего построим КА для языка L_0, который может быть задан регулярным выражением $L_0 = (ab)^*(ba)^*$. Граф показан на рис. 4. Удаляя единственный λ-переход, получим граф на рис. 5.

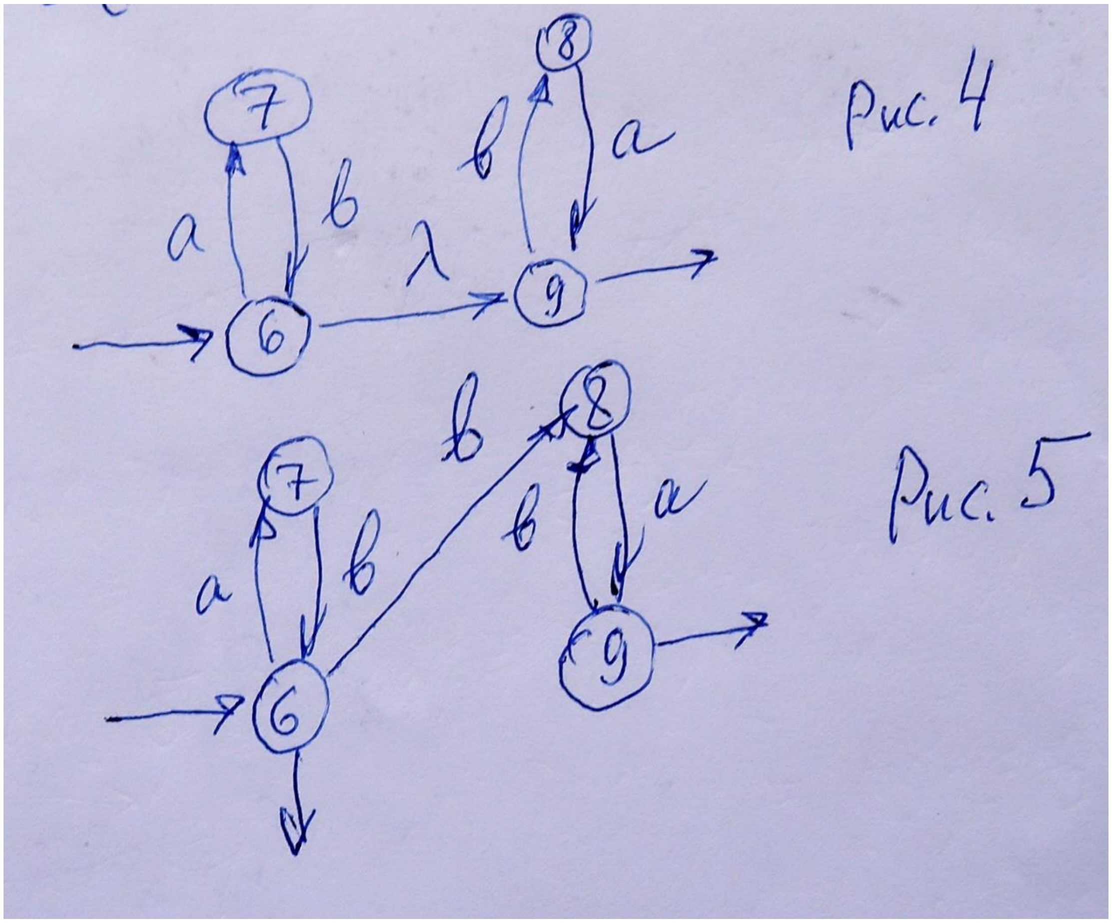

Рис. 4 и 5

Автоматы для упомянутых выше конструкций изображены на рис. 6-9: сначала с λ-переходами, потом – без них. Обратим внимание на то, что КА для итерации строится несколько иначе, чем изложено в Учебнике. Мы проводим λ-переходы из новой начальной вершины (s) в новую заключительную (t) и сразу обратно – из t в s. Рекомендуется использовать именно это построение. Оно дает тот же результат, что и показанное в Учебнике, но проще в реализации.

По поводу удаления λ-переходов необходимо четко усвоить следующее.

1) После удаления в КА остаются только исходное начальное состояние и те состояния исходного КА, в которые заходит хотя бы одна «непустая» дуга, т.е. дуга с меткой, отличной от λ. Состояния, в которые заходят только λ-переходы, исчезают.

2) При перестройке графа надо отследить все такие тройки состояний p, q, r, что существует путь из первого состояния во второе с меткой λ, т.е. проходящий исключительно по λ-переходам, а из второго в третье ведет дуга, метка которой содержит букву входного алфавита ($a \in V$): $p\underset{\to}{\lambda}\ldots\underset{\to}{\lambda}q\underset{\to}{a}r$. Тогда после удаления λ-переходов следует добавить дугу, ведущую из p в r с меткой a: $p\underset{\to}{a}r$. Среднее состояние может исчезнуть, а может остаться в зависимости от того, какие дуги в него заходят (см. стр. 522-23 Учебника). Заметим, что могут совпадать второе и третье состояние (в исходном КА есть петля), а также первое и третье: тогда после удаления λ-переходов возникает новая петля. Первое и второе состояния совпадать не могут, так как по определению невозможны петли с меткой λ.

3) К множеству заключительных состояний нового КА (без λ-переходов) добавляются те состояния исходного КА, из которых ведет путь по λ-переходам (т.е. путь с меткой λ) в заключительное состояние исходного КА (которое опять-таки может исчезнуть, а может и остаться). Из исходного КА останутся, понятно, только те заключительные состояния, в которые заходит хотя бы одна «непустая» дуга.

Поясним подробно работу алгоритма удаления λ-переходов на примере КА для итерации (рис. 8). Именно здесь возникает наибольшее число ошибок у студентов.

Нужно проанализировать все состояния, из которых выходят λ-переходы, чтобы найти все указанные выше в п. (2) тройки состояний.

Начнем с начального состояния *s* (оно остается!).

Из *s* по λ можно попасть в состояние 2, а из него по букве *a* в состояния 1 и 5. Следовательно, должны быть дуги (s, 1) и (s, 5) с меткой *a*.

Из состояния 2 ведет дуга с меткой *b* в состояние 4; значит, надо добавить дугу (s, 4) с меткой *b*.

Это все новые дуги, исходящие из *s*. Можно доказать к тому же, что после удаления λ-переходов *в вершине s не может быть никаких заходящих дуг*. Так что если они вдруг возникли, то это явная ошибка.

Теперь анализируем подобным же образом вершины 4 и 5.

Вершина 4: по λ-переходам идем через t и s в вершину 2, откуда по букве *a* попадаем в 1 и 5; следовательно, в новом КА будут дуги (4, 1) и (4, 5) с меткой *a*. Но поскольку из вершины 2 по букве *b* можно перейти в вершину 4, то возникает петля (4, 4) с меткой *b*.

Вершина 5: по λ-переходам идем через t и s в вершину 2, откуда по букве *a* попадаем в 1 и 5; следовательно, в новом КА будут дуги (5, 1) и (5, 5) – петля! - с меткой *a*. Но поскольку из вершины 2 по букве *b* можно перейти в вершину 4, то возникает дуга (5, 4) с меткой *b*. В последнем случае, поскольку уже есть дуга (5, 4) с меткой *a* и нет «кратных» дуг, то на дуге (5, 4) появляется и буква *b*.

Заключительными состояниями в новом КА будут s, 4 и 5. Вершина t исчезнет (и только она). Заметим, что итерация любого языка содержит пустую цепочку λ. Но КА без λ-переходов может ее допустить лишь в том случае, когда начальное состояние является и заключительным (пустая цепочка читается на пути длины 0).

Граф КА для итерации языка L представлен на рис. 9.

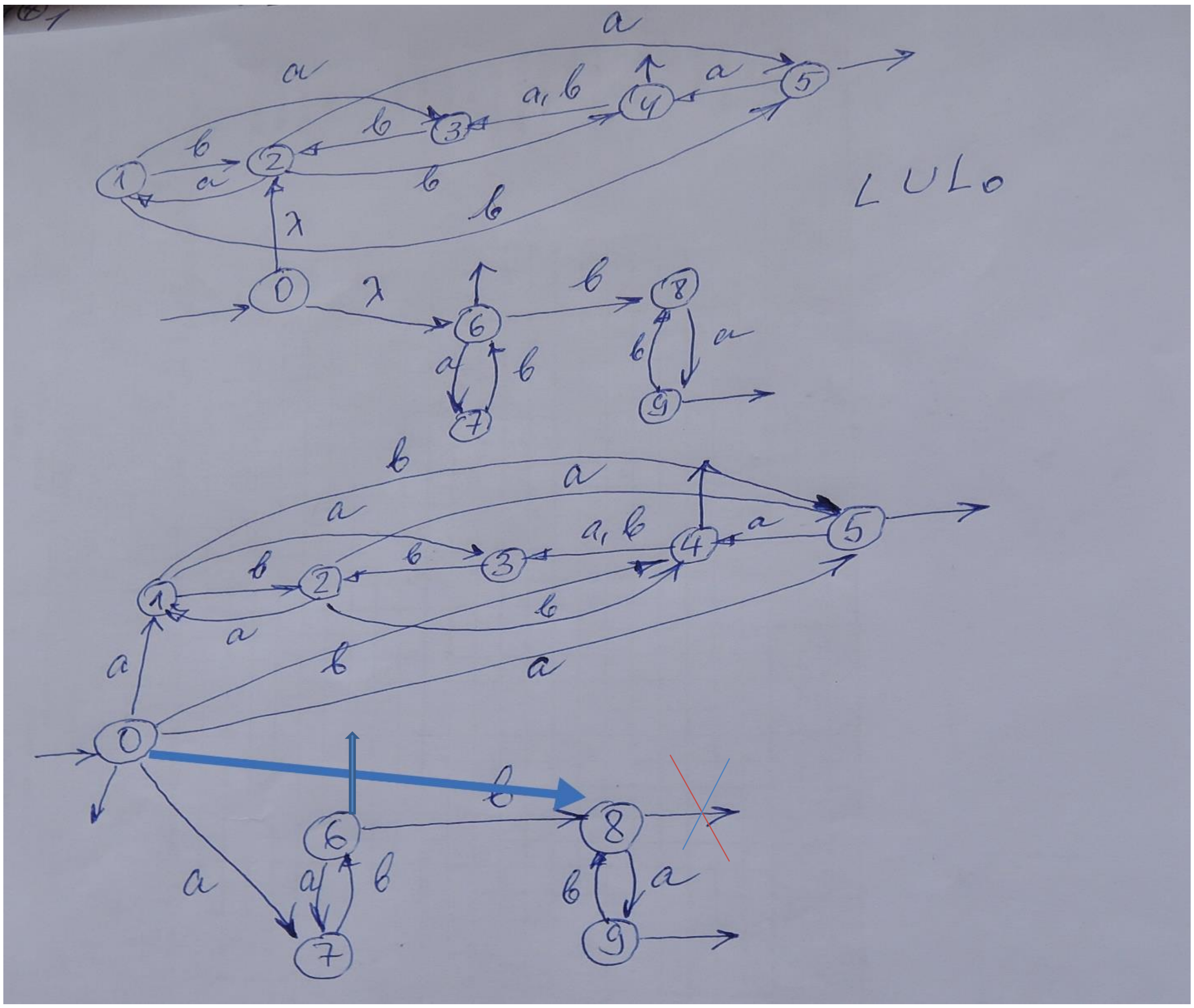

Рис. 6 (метка дуги 0→8 есть *b*; была забыта на самом рисунке, добавлена уже на фото; выходная стрелка в 8 нарисована ошибочно, а выход должен быть в вершине 6.)

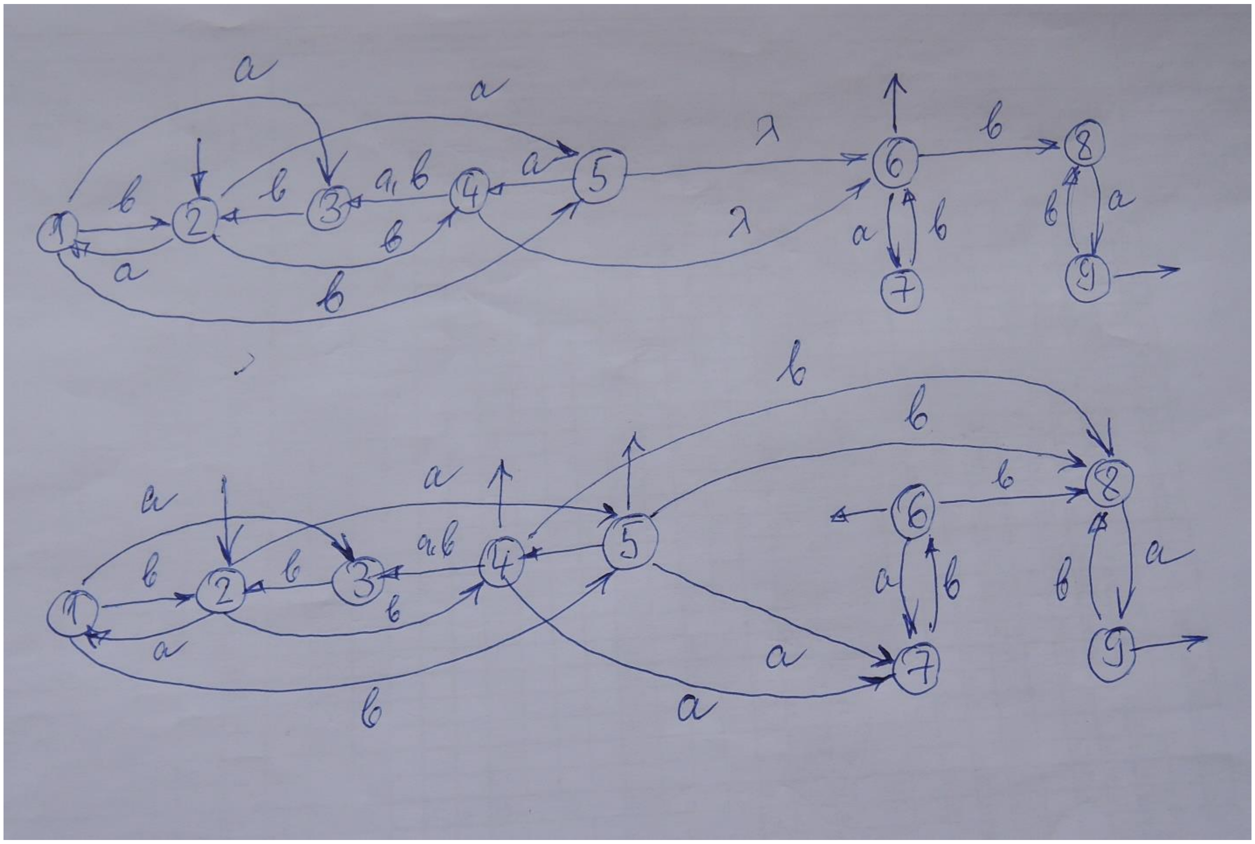

Рис. 7. КА для соединения $L \circ L_0$

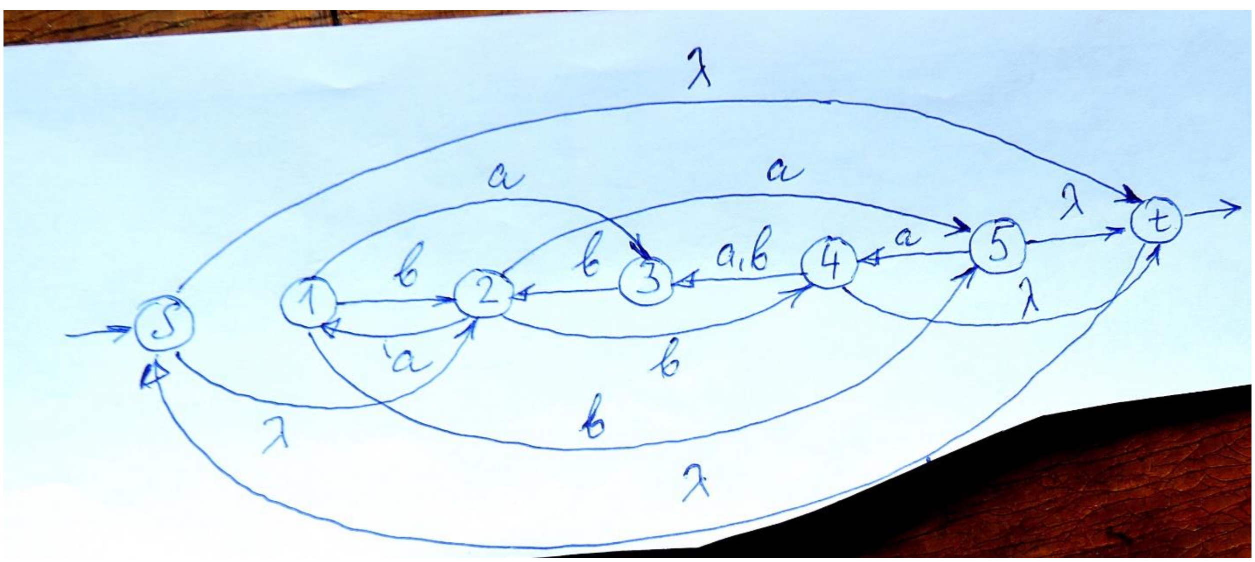

Рис. 8. КА для итерации L* с λ-переходами

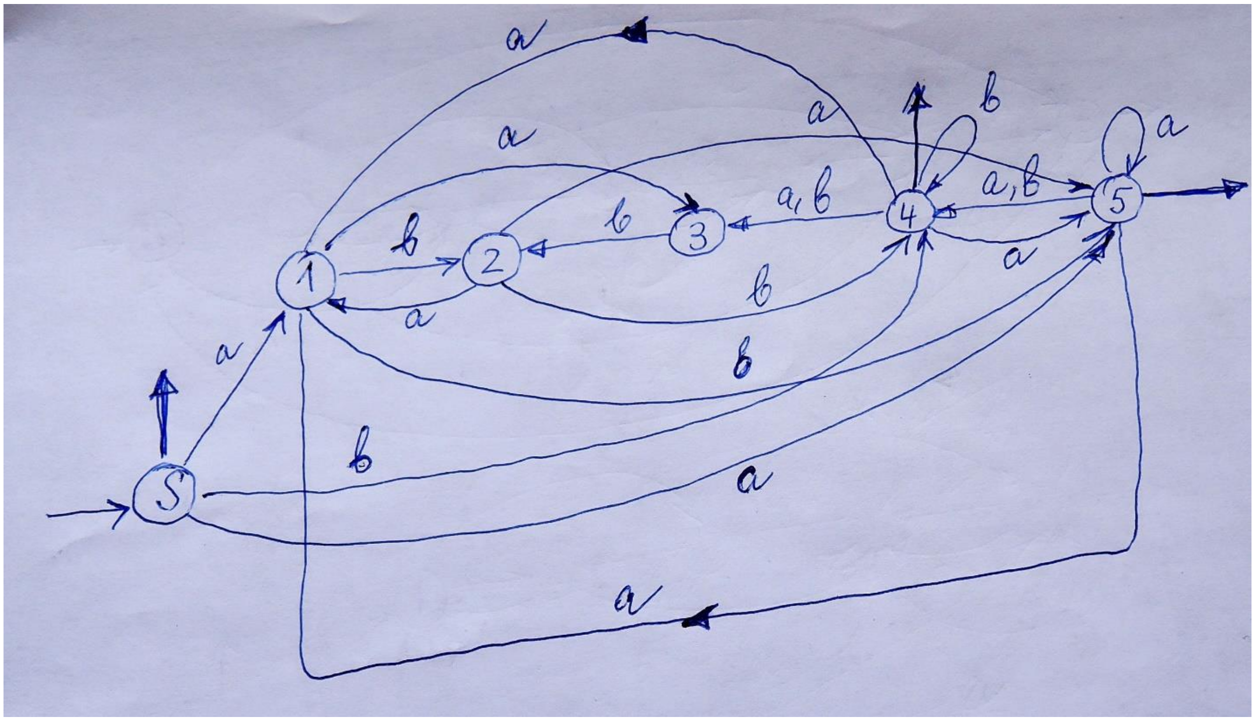

Рис. 9. КА для итерации L* без λ -переходов