

ПОСТРОЕНИЕ УНИВЕРСАЛЬНОГО КОНЕЧНОГО АВТОМАТА. II. ПРИМЕРЫ РАБОТЫ АЛГОРИТМОВ

В. Н. Долгов, Б. Ф. Мельников

Самарский государственный университет

Поступила в редакцию 25.03.2013 г.

Аннотация: во второй части статьи рассматриваются подробные примеры работы алгоритмов, описанных в первой части. Кроме того, приводится серия примеров, показывающая, как быстро может расти число состояний универсального автомата в зависимости от числа состояний эквивалентного канонического автомата либо канонического автомата для зеркального языка.

Ключевые слова: недетерминированные конечные автоматы, универсальный автомат Конвея, алгоритмы построения.

Abstract: the second part of the paper contains some detailed examples of functioning of the algorithms, described in the first part. It also contains a series of examples which shows, how fast the size of the universal automaton can grow, depending on the number of states of the equivalent canonical automaton or the canonical automaton for the mirror language.

Keywords: nondeterministic finite automata, Conway's universal automaton, constructing algorithms.

6. КРАТКОЕ СОДЕРЖАНИЕ

Эта статья является продолжением части I статьи этих же авторов [1]. Краткое содержание второй части таково.

В разделе 7 мы даём подробный пример работы алгоритма, описанного в предыдущем разделе (разделе 5 первой части). Также приводится пример покрывающего автомата, *эквивалентного* заданному.

В разделе 8 мы рассматриваем серию примеров, которая показывает, насколько быстро может расти число блоков (т.е. размер универсального автомата) – если заданы размеры двух канонических автоматов (т.е. автоматов $\tilde{L}$ и $\tilde{L}^R$). По мнению авторов, эти оценки дополняют приведённые в [2] – поскольку имеют иные «входные данные».

В заключении (раздел 9) мы приводим некоторые возможные направления дальнейших исследований по данной тематике.

7. ПОДРОБНЫЙ ПРИМЕР

Рассмотрим пример языка из [3, Sect. 3]. Для этого приведём автоматы $\tilde{L}$ и $\tilde{L}^R$ для этого языка (рис. 1 и 2), а также таблицу его бинарного отношения $\#$ (Таб. 1):

Таблица 1

#	X	Y	Z	U
A	–	#	#	–
B	#	–	#	–
C	#	#	#	#
D	#	#	#	–

Заметим, что в [3, Sect. 3] мы просто *указали* множество блоков; а в этой работе, мы *используем алгоритм* его построения. Итак, рассмотрим работу алгоритма 3.

Ориентированный граф $\mathcal{DG}$, вершинами которого являются все непустые подмножества множества Q_π , изображён на рис. 3 (при этом подмножества обозначены просто строками, состоящими из их элементов). Приведём некоторые комментарии к этому рисунку.

Множества, помеченные 3 овалами – т.е. $\{A, B, C, D\}$, $\{A, C, D\}$, $\{B, C, D\}$ и $\{C\}$ – были построены шагом 1 алгоритма 3.

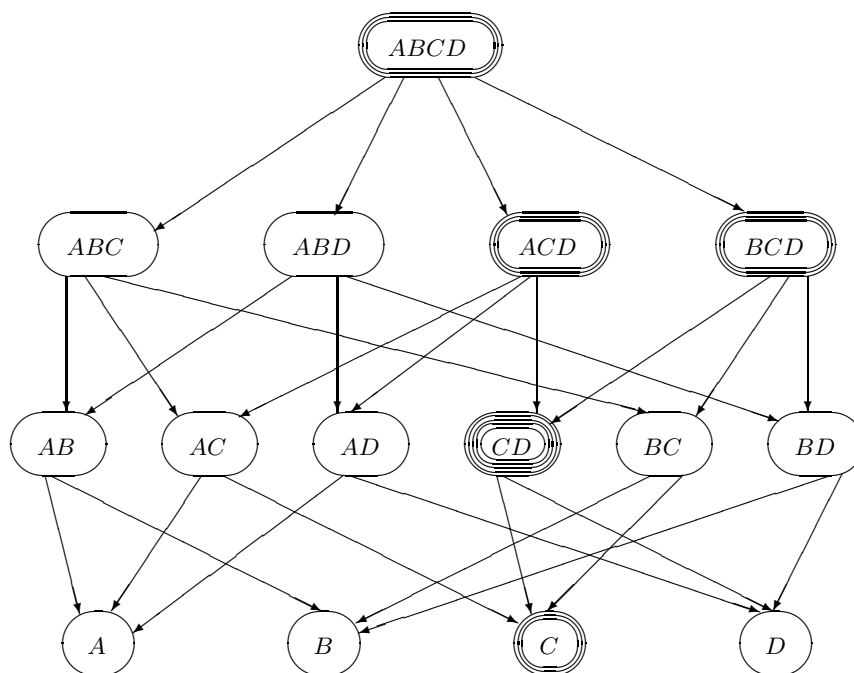


Рисунок 3

На шагах 2 и 3 этого алгоритма мы рассматриваем и другие подмножества (другие вершины графа $\mathcal{DG}$). При этом у нас есть единственная вершина (а именно, $\{C, D\}$), для которой существуют по крайней мере 2 такие вершины, из которых имеются дуги в $\{C, D\}$; мы поместили эту «новую подходящую» вершину одновременно несколькими овалами. Итак, все 5 упомянутых вершин (и только они) являются элементами множества $\alpha(\mathcal{B})$ для некоторых блоков $\mathcal{B}$.

С помощью шага 4 алгоритма 3 мы выбираем для них следующее множество блоков¹:

$$\begin{aligned}\zeta &= \{A, C, D\} \times \{Y, Z\}, \quad \eta = \{A, B, C, D\} \times \{Z\}, \\ \vartheta &= \{B, C, D\} \times \{X, Z\}, \\ \nu &= \{C\} \times \{X, Y, Z, U\}, \quad \xi = \{C, D\} \times \{X, Y, Z\}.\end{aligned}$$

Далее эти же буквы (т.е., ζ , η , ϑ , ν и ξ) будут обозначать и состояния автомата $\mathcal{U}_L$.

С помощью шага 5 алгоритма 3 мы получаем следующие множества входов и выходов $\mathcal{U}_L$:

$$S_{\text{вх}} = \{\zeta, \eta\}, \quad F_{\text{вых}} = \{\vartheta, \nu, \xi\}.$$

Также с помощью шага 5 (т.е. по определению автомата $\mathcal{COM}(L)$, приведённому в разделе 3), мы получаем функцию переходов δ_{COM} – следующим способом.

Во-первых, рассмотрим таб. 2 – таблицу функции переходов для буквы a , получаемую на основе канонического автомата:

Таблица 2

блоки ($\mathcal{B}$)	$\alpha(\mathcal{B})$	$\xrightarrow[\delta_\pi]{a}$	ζ	η	ϑ	ν	ξ
ζ	ACD	BC	–	+	+	–	–
η	$ABCD$	–	–	–	–	–	–
ϑ	BCD	–	–	–	–	–	–
ν	C	C	+	+	+	+	+
ξ	CD	BC	–	+	+	–	–

Здесь мы используем упрощённые обозначения – аналогично тому, как мы поступали ранее. Т.е. элементы 2-го и 3-го столбцов этой таблицы являются подмножествами множеств элементов Q_π . Однако нужно подробно пояснить эти обозначения – поскольку символы «–» в клетках 3-го столбца таблицы *не* обозначают пустого множества.

Итак, элементы 2-го столбца суть элементы $\alpha(\mathcal{B})$ рассматриваемого блока $\mathcal{B}$ (который находится в 1-м столбце). Далее, если для каждого элемента этого подмножества (пусть сам этот элемент, т.е. само это состояние – A) существует переход $A \xrightarrow[\delta_\pi]{a} B$ для некоторого состояния $B \in Q_\pi$, то в этой строке соответствующее множество в 3-м столбце является объединением всех таких B .² Иначе – т.е. если для некоторого элемента A (из 2-го столбца) *не существует* перехода $A \xrightarrow[\delta_\pi]{a} B$ для какого-либо $B \in Q_\pi$ – мы помещаем в 3-м столбце символ «–».

Правая часть этой таблицы (т.е. столбцы начиная с 4-го) – квадратная матрица. Например, для её элемента, соответствующего строке ζ и столбцу η , мы ставим + в клетке таблицы тогда и только тогда, когда элемент в 3-м столбце строки ζ :

- не есть «–»;
- кроме того, он является подмножеством множества, которое мы уже поставили в 3-м столбце строки η ;

¹ Ещё раз отметим, что в [3] мы просто *указали* множество блоков. А в этом разделе мы использовали алгоритм 3 для их *построения*.

² Напомним, что мы рассматриваем букву a . Также напомним, что мы рассматриваем канонические автоматы *без* «мёртвых» состояний.

иначе же мы ставим в эту клетку «-». ³ Итак, правая часть этой таблицы образует квадратную матрицу (т.е. матрицу

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix} \quad (5)$$

для рассматриваемого нами примера), которая действительно определяет условие (1) для буквы a .

Таблица 3

блоки ($\mathcal{B}$)	$\beta(\mathcal{B})$	$\xrightarrow[\delta_\rho]{a}$	ζ	η	ϑ	ν	ξ
ζ	YZ	YU	-	-	-	+	-
η	Z	Y	+	-	-	+	+
ϑ	XZ	Y	+	-	-	+	+
ν	$XYZU$	YU	-	-	-	+	-
ξ	XYZ	YU	-	-	-	+	-

Для той же самой буквы a и другого канонического автомата, а именно $\widetilde{L}^R$, мы получаем таб. 3. Аналогично предыдущему, правая часть этой таблицы определяет для буквы a условие (2).

Далее. Поэлементная конъюнкция:

- первой матрицы (в рассматриваемом нами примере – матрицы (5))
- и транспонированной матрицы из таб. 3

– даёт матрицу наличия a -переходов. Заметим, что в рассмотренном примере матрицы для поэлементной конъюнкции совпадают⁴ – однако существуют и примеры, где эти матрицы различаются.

Рассмотрим аналогичные таблицы для буквы b . Для автомата $\widetilde{L}$ мы получаем следующую таб. 4:

Таблица 4

блоки ($\mathcal{B}$)	$\alpha(\mathcal{B})$	$\xrightarrow[\delta_\pi]{b}$	ζ	η	ϑ	ν	ξ
ζ	ACD	C	+	+	+	+	+
η	$ABCD$	CD	+	+	+	-	+
ϑ	BCD	CD	+	+	+	-	+
ν	C	C	+	+	+	+	+
ξ	CD	C	+	+	+	+	+

А для автомата $\widetilde{L}^R$ – следующую таб. 5:

³ Особо отметим, что символ «-» в 3-м столбце подразумевает «-» в каждой клетке этой строки в правой части таблицы.

⁴ Для буквы b обе соответствующие матрицы также совпадают – см. ниже.

Таблица 5

блоки ($\mathcal{B}$)	$\beta(\mathcal{B})$	$\xrightarrow[\delta_p]{a}$	ζ	η	ϑ	ν	ξ
ζ	YZ	Z	+	+	+	+	+
η	Z	Z	+	+	+	+	+
ϑ	XZ	Z	+	+	+	+	+
ν	$XYZU$	YZ	+	–	–	+	+
ξ	XYZ	Z	+	+	+	+	+

На основе последних двух таблиц мы для автомата $\mathcal{CM}(L)$ получаем следующую матрицу b -переходов:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (6)$$

(в ней порядок блоков совпадает с предыдущим – т.е. $\zeta, \eta, \vartheta, \nu, \xi$).

На основе матриц (5) и (6) мы для рассматриваемого языка получаем следующий автомат $\mathcal{CM}(L)$:

Таблица 6

$\mathcal{CM}(L)$	a	b
$\rightarrow \zeta$	η, ϑ	$\zeta, \eta, \vartheta, \nu, \xi$
$\rightarrow \eta$	–	$\zeta, \eta, \vartheta, \xi$
$\leftarrow \vartheta$	–	$\zeta, \eta, \vartheta, \xi$
$\leftarrow \nu$	$\zeta, \eta, \vartheta, \nu, \xi$	$\zeta, \eta, \vartheta, \nu, \xi$
$\leftarrow \xi$	η, ϑ	$\zeta, \eta, \vartheta, \nu, \xi$

Рассмотрим покрывающие подмножества множества блоков этого автомата. Одно из них⁵ – следующее: $\{\zeta, \vartheta, \nu\}$. Далее, используя определение покрывающего автомата (раздел 4), мы получаем следующий *соответствующий* покрывающий автомат $\mathcal{CM}_{\{\zeta, \vartheta, \nu\}}(L)$ (т.е., покрывающий автомат для множества $\{\zeta, \vartheta, \nu\}$, см. таб. 7 и рис. 4):

Таблица 7

$\mathcal{CM}_{\{\zeta, \vartheta, \nu\}}(L)$	a	b
$\rightarrow \zeta$	ϑ	ζ, ϑ, ν
$\leftarrow \vartheta$	–	ζ, ϑ
$\leftarrow \nu$	ζ, ϑ, ν	ζ, ϑ, ν

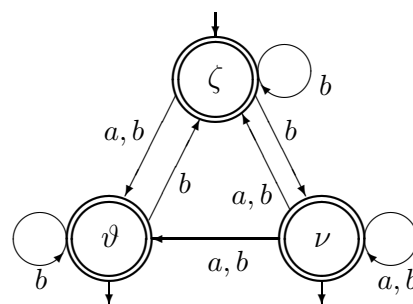


Рисунок 4

Легко показать, что последний НКА действительно определяет заданный язык.

Таким образом, приведённый пример даёт *эквивалентный* покрывающий автомат. Однако существуют примеры языков, для которых существуют покрывающие автоматы, *не*

⁵ Как уже было отмечено, мы не будем рассматривать *алгоритмы* для построения таких подмножеств для произвольного автомата. В нашем примере очевидно, что существует *единственное* покрывающее подмножество, содержащее не более 3 блоков.

определяющие заданный язык. Мы предполагаем продолжить изучение таких примеров в следующих работах.

8. КРАТКОЕ ОПИСАНИЕ СЕРИИ ПРИМЕРОВ

В [2, Th. 5.1], была получена точная верхняя граница размера универсального автомата – для регулярного языка, который может быть задан с помощью некоторого НКА, имеющего k состояний. Это значение оказалась k -м числом Дедекинда ($D(k)$, см. [5] и др.).

В этом разделе мы рассмотрим серию примеров, которая показывает, насколько быстро может расти число блоков (т.е. размер универсального автомата) – если нам даны размеры обоих *канонических* автоматов (т.е. $\tilde{L}$ и $\tilde{L}^R$). По мнению авторов, эти примеры дополняют рассмотренные в [2].

Итак, сначала рассмотрим следующую матрицу размера 12×12 :

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (7)$$

Очевидно, что мы можем рассматривать подобные матрицы для любого $x \geq 2$ (где $x \times x$ – размеры «блока нулей») и любого $n \geq 2$, такого что $n \bmod x = 0$ (где $n \times n$ – размеры матрицы). В приведённом выше примере (7) – $x = 4$ и $n = 12$. (Более того, можно взять блоки нулей разного размера, см. ниже.) Однако формулы для элементов таких матриц мы записывать не будем.

Для каждой подобной матрицы мы можем построить соответствующий НКА, имеющий в точности такую таблицу отношения $\#$: на основе доказанного в [4] (см. также [6]) можно показать, что такой автомат всегда существует.

Строя *некоторые* блоки для примера (7), можно выбрать ровно одну строку из множества $\{1, 2, 3, 4\}$, далее ровно одну строку из множества $\{5, 6, 7, 8\}$ и, наконец, ровно одну строку из множества $\{9, 10, 11, 12\}$; при этом очевидно, что все номера выбираемых столбцов должны совпадать с номерами ранее выбранных строк. Таким образом, в примере (7) имеется, по крайней мере, $4^3 = 64$ блоков.

Далее, пусть

$$n = a_1 + a_2 + \dots + a_k,$$

и имеются блоки нулей размерами $a_1, a_2, \dots, a_k$; тогда соответствующее число блоков (состояний канонического автомата) будет равно $a_1 \cdot a_2 \cdot \dots \cdot a_k$. Наша задача – найти максимум этого значения при фиксированном значении n . Для этого нужно учесть, что

$$\begin{aligned} 4 &= 2 + 2, & 2 \cdot 2 &= 4 \geq 4; \\ 5 &= 2 + 3, & 2 \cdot 3 &= 6 > 5; \\ 6 &= 2 + 2 + 2 = 3 + 3, & 3 \cdot 3 &> 2 \cdot 2 \cdot 2 \quad \text{и т.д.} \end{aligned}$$

– т.е. каждое слагаемое, большее или равное 4, можно разбить на двойки и тройки, и при этом произведение не уменьшится. Также очевидно, что при разбиении слагаемых необходимо отдавать предпочтение тройкам (см. последнее неравенство).

Итак, искомый максимум имеет один из следующих видов:

$$\begin{aligned} & 3 \cdot 3 \cdot \dots \cdot 3 \\ & 3 \cdot 3 \cdot \dots \cdot 3 \cdot 2 \cdot 2 \\ & 3 \cdot 3 \cdot \dots \cdot 3 \cdot 2, \end{aligned}$$

– в зависимости от остатка при делении n на 3. Для простоты рассуждений ограничимся случаем $n \bmod 3 = 0$. В этом случае число блоков будет не менее $3^{n/3}$.

Таким образом, наша серия примеров показывает, что размер универсального автомата может расти экспоненциально (с основанием $3^{1/3}$) по отношению к размеру канонических автоматов – т.е. по отношению к $n = \min(|Q_\pi|, |Q_\rho|)$. Повторим, что наши примеры *не дают* точного числа блоков, которое может образоваться в случае если нам про автоматы $\tilde{L}$ и $\tilde{L}^R$ известно лишь значение $n = \min(|Q_\pi|, |Q_\rho|)$.

Отметим, что нет смысла комбинировать описанную здесь серию примеров с очевидной оценкой $n = 2^k$ для числа состояний канонического автомата: получающаяся функция $3^{n/3} = 3^{2^k/3}$ растёт значительно быстрее, чем $D(k)$. Это означает лишь то, что языки с отношением $\#$, аналогичным (7), не могут иметь эквивалентных автоматов с числом состояний, намного меньшим, чем n .

9. ЗАКЛЮЧЕНИЕ

В последующих работах мы предполагаем рассмотреть:

- циклы базисного автомата $\mathcal{BA}(L)$ и их использование для выражения пометок циклов автомата $\mathcal{COM}(L)$;
- детальное исследование примера, в котором покрывающий автомат, в отличие от рассмотренного в разделе 7, *не* определяет заданный язык; иными словами – привести подробное исследование т.н. автомата Ватерлоо ([7]) «с точки зрения» эквивалентного базисного автомата;
- исследование серии примеров, которая доставляет *конструктивное* доказательство следующего факта: примеры, подобные Ватерлоо, могут быть построены *для каждой* таблицы отношения $\#$, имеющей следующее дополнительное свойство: существует *собственное* подмножество множества блоков, являющееся покрывающим;
- специальную математическую модель, описывающую понятие «средний конечный автомат» в терминологии так называемых No-Free-Lunch-теорем ([8]), – и показать, что рассмотренные нами алгоритмы более эффективны (с точки зрения этой модели), чем алгоритмы, строящие $\mathcal{U}_L$ по определению, т.е. использующие (суб)факторизации и т.п. понятия⁶.

БЛАГОДАРНОСТИ

Авторы выражают благодарность профессору Антону Клячко (Московский государственный университет) за его ценные замечания.

⁶ Отметим, что работы, соответствующие этому пункту, начались публикацией [9].

СПИСОК ЛИТЕРАТУРЫ

- [1] Долгов В.Н. Построение универсального конечного автомата. I. От теории к практическим алгоритмам / В.Н. Долгов, Б.Ф. Мельников // Вестн. Воронеж. гос. ун-та. Сер. Физика, математика. — 2013. — № 2. — С. 131–139.
- [2] S.Lombardy, J.Sakarovitch: The Universal Automaton – *Logic and Automata, Texts in Logic and Games, Amsterdam Univ. Press*, Vol. 2 (2008) 457–504.
- [3] В. Melnikov: Once more on the edge-minimization of nondeterministic finite automata and the connected problems – *Fundamenta Informaticae*, Vol. 104, No. 3 (2010) 267–283.
- [4] А. Мельникова: Некоторые свойства базисного автомата – *Вестник Воронежского государственного университета. Серия «Физика. Математика»*, № 2 (2012) 184–189.
- [5] R. Dedekind: *Über Zerlegungen von Zahlen durch ihre größten gemeinsamen Teiler*. Gesammelte Werke, Vol. 2, 103–148.
- [6] М. Зузанова: Некоторые алгоритмы эквивалентного преобразования недетерминированных конечных автоматов. Дисс. . . канд. физ.-мат. наук: 05.13.18, Тольяттинский государственный университет, Д 212.264.03, защищена 01.07.2010.
- [7] T. Kameda, P. Weiner: On the state minimization of nondeterministic finite automata – *IEEE Transactions on Computation*, Vol. C-19, No. 7 (1970) 617–627.
- [8] D. Wolpert, W. Macready: No Free Lunch Theorems for Optimization – *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 1 (1997) 67–82.
- [9] Б. Мельников, С. Пивнева, О. Рогова: Репрезентативность случайно сгенерированных недетерминированных конечных автоматов с точки зрения соответствующих базисных автоматов – *Стохастическая оптимизация в информатике*, № 6 (2010) 74–82.

Долгов В.Н., аспирант кафедры «Прикладная математика и информатика» Тольяттинского государственного университета
E-mail: terenga74@mail.ru
Тел.: (927) 8132544

Dolgov V.N., post-graduate student of department “Applied Mathematics and Informatics”, Togliatti State University
E-mail: terenga74@mail.ru
Tel.: (927) 8132544

Мельников Б.Ф., доктор физико-математических наук, профессор кафедры «Прикладная математика и информатика» Тольяттинского филиала Самарского государственного университета
E-mail: bormel@rambler.ru
Тел.: (927) 6109043

Melnikov B.F., doctor of physical and mathematical sciences, professor of department “Applied Mathematics and Informatics”, Togliatti Branch of Samara State University
E-mail: bormel@rambler.ru
Tel.: (927) 6109043