Известия вузов. Математика
2009, № 4, с. 67–71

http://www.ksu.ru/journals/izv_vuz/
e-mail: izvuz.matem@ksu.ru

Краткое сообщение

Б.Ф. МЕЛЬНИКОВ, М.Р. САЙФУЛЛИНА

О НЕКОТОРЫХ АЛГОРИТМАХ ЭКВИВАЛЕНТНОГО ПРЕОБРАЗОВАНИЯ НЕДЕТЕРМИНИРОВАННЫХ КОНЕЧНЫХ АВТОМАТОВ

Аннотация. В данной статье рассматриваются алгоритмы, позволяющие объединять несколько состояний недетерминированного конечного автомата в одно. Кроме алгоритмов объединения состояний, в настоящей статье рассмотрен еще один алгоритм эквивалентного преобразования недетерминированного конечного автомата, а именно, алгоритм добавления циклов. Вопросы, рассмотренные авторами, ориентированы на создание эффективных компьютерных программ.

Ключевые слова: недетерминированные конечные автоматы, эквивалентное преобразование, расширенный автомат, объединение состояний.

УДК: 519.6

Abstract. In this paper we consider algorithms which allow one to combine several states of a nondeterministic finite automaton into one state. Along with the algorithms for combining states, we adduce one more algorithm for the equivalent transformation of a non-deterministic finite automaton, namely, an algorithm for adding cycles. Problems under consideration imply the development of robust computer programs.

Keywords: nondeterministic finite automata, equivalent transformation, extended automaton, combining states.

1. Введение

Данная статья является продолжением предыдущих работ авторов, где регулярные языки и недетерминированные конечные автоматы Рабина–Скотта (ниже НКА) рассматривались с точки зрения так называемой функции разметки состояний.

В статье рассматриваются алгоритмы, позволяющие объединять несколько состояний некоторого НКА в одно. При этом не ставится цель получения какого-либо нового алгоритма вершинной минимизации НКА, т. е. получения автомата, определяющего заданный регулярный язык и имеющего минимально возможное число состояний. Такие алгоритмы применяются в различных задачах, в частности, для решения задач теории регулярных языков ([1], [2]).

Поступили: полный текст 11.01.2005, краткое сообщение 16.09.2008.

Вопросы, рассмотренные в данной статье, ориентированы на создание эффективных компьютерных программ. Поэтому авторы сознательно опускают оценку сложности приведенных здесь алгоритмов. При доведении этих алгоритмов до завершенных компьютерных программ применяются специальные эвристики для уменьшения среднего времени работы алгоритмов ([3]).

Кроме алгоритмов объединения состояний, в данной статье рассмотрен еще один алгоритм эквивалентного преобразования НКА, а именно, алгоритм добавления циклов. Авторы считают это преобразование не усложнением, а даже упрощением. Например, оно позволяет добавлять циклы, отсутствовавшие в исходном конечном автомате, но присутствовавшие в эквивалентном ему базисном.

Подобные алгоритмы применяются и при других преобразованиях автоматов, среди которых имеется описанная в данной статье возможность последовательных преобразований алгоритма, позволяющая по любому автомату для заданного регулярного языка получить любой ему эквивалентный.

2. Алгоритмы объединения состояний автомата

Будем использовать определения и обозначения из [4]. Пусть

$$K = (Q, \Sigma, \delta, S, F) \tag{1}$$

— некоторый НКА, определяющий регулярный язык $L = \mathcal{L}(K)$.

Теорема 1 ([4])**.** *Пусть для автомата* (1) *выполнено условие* $\varphi_K^{\text{in}}(q) = \varphi_K^{\text{in}}(q')$ *для некоторых его состояний* $q, q' \in Q$. *Тогда автомат* $K' = \mathcal{J}^{qq'}(K)$ *эквивалентен* K *и*

$$\varphi_K^{\text{in}}(q) = \varphi_{K'}^{\text{in}}(q). \tag{2}$$

Схема доказательства. Рассмотрим такое слово $v \in \mathcal{L}(K')$, что автомат K' при его принятии проходит через состояние q. Поэтому запишем v в виде

$$v = uv_1v_2\dots v_nw, \tag{3}$$

где

- $u \in \mathcal{L}_{K'}^{\text{in}}(q)$,
- $(\forall i \in \{1, 2, \dots, n\})(v_i \in \mathcal{L}_{K'}^{\text{io}}(q))$,
- $w \in \mathcal{L}_{K'}^{\text{out}}(q)$,

и при этом

- u не может быть записано как $u = u'u''$, где $u' \in \mathcal{L}_{K'}^{\text{in}}(q)$, $u'' \in \mathcal{L}_{K'}^{\text{io}}(q)$,
- никакое v_i не может быть записано в виде $v_i = v'v''$, где $v', v'' \in \mathcal{L}_{K'}^{\text{io}}(q)$,
- w не может быть записано как $w = w'w''$, $w' \in \mathcal{L}_{K'}^{\text{io}}(q)$, $w'' \in \mathcal{L}_{K'}^{\text{out}}(q)$.

(Значит, автомат K' в процессе чтения слова v проходит состояние q ровно $n+1$ раз. Заметим, что, вообще говоря, K' не является однозначным автоматом, т. е. может существовать и другой способ принятия автоматом K' слова v.)

Докажем индукцией по n, что для любого v, которое может быть записано в виде (3), выполнено условие $v \in \mathcal{L}(K)$. Базис индукции очевиден, поскольку если $n = 0$, то автомат K' может принять слово v без прохождения через состояние q; следовательно, автомат K может принять v без прохождения как q, так и q'. Докажем шаг индукции.

Обозначим $v' = v_1v_2\dots v_{n-1}$, $v'' = v_n$. Без ограничения общности предположим, что $u \in \mathcal{L}_K^{\text{in}}(q)$ (случай $u \in \mathcal{L}_K^{\text{in}}(q')$ рассматривается аналогично). Поэтому рассмотрим следующие восемь случаев:

A) $v' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $w \in \mathcal{L}_K^{\mathrm{out}}(q')$;
B) $v' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $w \in \mathcal{L}_K^{\mathrm{out}}(q')$;
C) $v' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $w \in \mathcal{L}_K^{\mathrm{out}}(q')$;
D) $v' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $w \in \mathcal{L}_K^{\mathrm{out}}(q')$;
E) $v' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $w \in \mathcal{L}_K^{\mathrm{out}}(q)$;
F) $v' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $w \in \mathcal{L}_K^{\mathrm{out}}(q)$;
G) $v' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $w \in \mathcal{L}_K^{\mathrm{out}}(q)$;
H) $v' \in \mathcal{L}_K^{\mathrm{io}}(q',q)$, $v'' \in \mathcal{L}_K^{\mathrm{io}}(q,q')$, $w \in \mathcal{L}_K^{\mathrm{out}}(q)$.

Чтобы доказать эквивалентность автоматов, необходимо рассмотреть случаи A)–H) для условия $\mathcal{L}(K') \subseteq \mathcal{L}(K)$, поскольку условие $\mathcal{L}(K) \subseteq \mathcal{L}(K')$ очевидно. Согласно [4] равенство (2) выполняется по определению детерминированного автомата и функции φ^{in} (поскольку существует канонический конечный автомат, который эквивалентен обоим рассматриваемым автоматам K и K').

Аналогично предыдущей доказываются две следующие теоремы.

Теорема 2. *Пусть для автомата* (1) $\varphi_K^{\mathrm{out}}(q) = \varphi_K^{\mathrm{out}}(q')$ *для двух состояний* $q, q' \in Q$. *Тогда автомат* $K' = \mathcal{J}^{qq'}(K)$ *эквивалентен* K *и* $\varphi_K^{\mathrm{out}}(q) = \varphi_{K'}^{\mathrm{out}}(q)$.

Теорема 3. *Если для автомата* (1) *и двух его состояний* $q, q' \in Q$ *выполнены следующие условия*: $\varphi_K^{\mathrm{in}}(q') \subseteq \varphi_K^{\mathrm{in}}(q)$, $\varphi_K^{\mathrm{out}}(q') \subseteq \varphi_K^{\mathrm{out}}(q)$ *и* $\varphi_K^{\mathrm{in}}(q') \neq \emptyset$, $\varphi_K^{\mathrm{out}}(q') \neq \emptyset$, *то автомат* $K' = \mathcal{J}^{qq'}(K)$ *эквивалентен* K *и* $\varphi_K^{\mathrm{in}}(q) = \varphi_{K'}^{\mathrm{in}}(q)$, $\varphi_K^{\mathrm{out}}(q) = \varphi_{K'}^{\mathrm{out}}(q)$.

Заметим, что объединение состояний может привести к изменению функций разметки [4].

3. Расширенный базисный автомат

Определение 1. Пару $\mathcal{K} = (K, \zeta)$, где $K = (Q, \Sigma, \delta, S, F)$ — автомат (1), а ζ — функция вида $\zeta : Q \to \mathcal{P}(\Sigma^*)$, будем называть расширенным конечным автоматом.

Определение 2. Пусть $q_1, q_2 \in Q$ — два состояния расширенного автомата. Определим язык $\mathcal{L}_{\mathcal{K}}^{\mathrm{io}}(q_1, q_2)$ следующим образом:

- если $u \in \mathcal{L}_K^{\mathrm{io}}(q_1, q_2)$, то $u \in \mathcal{L}_{\mathcal{K}}^{\mathrm{io}}(q_1, q_2)$;
- если для некоторого $q_3 \in Q$ выполняются условия $u_1 \in \mathcal{L}_{\mathcal{K}}^{\mathrm{io}}(q_1, q_3)$, $u_2 \in \zeta(q_2)$ и $u_3 \in \mathcal{L}_{\mathcal{K}}^{\mathrm{io}}(q_3, q_2)$, то $u_1 u_2 u_3 \in \mathcal{L}_{\mathcal{K}}^{\mathrm{io}}(q_1, q_2)$;
- ничто другое не является словом языка $\mathcal{L}_{\mathcal{K}}^{\mathrm{io}}(q_1, q_2)$.

Язык расширенного автомата определяется как $\mathcal{L}(\mathcal{K}) = \bigcup\limits_{s \in S,\, f \in F} \mathcal{L}_{\mathcal{K}}^{\mathrm{io}}(s, f)$.

Не каждый язык, заданный согласно определению 2, является регулярным. Однако для определенного ниже автомата $\mathcal{BE}(L)$ данный факт является следствием теоремы 4.

Утверждение 1. *Пусть для некоторого регулярного языка* L *и определяющего его автомата* (1) *выполнено следующее*: $q \in Q$, $q \in \left\langle \begin{smallmatrix} A \\ X \end{smallmatrix} \right\rangle_K$, $\mathcal{L}_K^{\mathrm{io}}(q) \ni v$, $\mathcal{L}_{\mathcal{BA}(L)}^{\mathrm{in}} \left(\begin{smallmatrix} A \\ X \end{smallmatrix} \right) \ni u$ *и* $\mathcal{L}_{\mathcal{BA}(L)}^{\mathrm{out}} \left(\begin{smallmatrix} A \\ X \end{smallmatrix} \right) \ni \omega$. *Тогда для каждого* $i \geq 0$ $uv^i w \in L$.

Итак, изучен специальный случай расширенных конечных автоматов. Схожие объекты рассматривались в [5]. Пусть регулярный язык L задан. Построим автомат $\mathcal{BE}(L)$ по следующему алгоритму. Определим языки $\zeta_{\mathcal{BE}(L)}$, соответствующие состояниям (не расширенного) автомата $\mathcal{BA}(L)$ [6].

Определение 3. Пусть задан регулярный язык L, и $\frac{A}{X}$ — некоторое состояние автомата $\mathcal{BA}(L)$. Пусть также (1) — некоторый определяющий L конечный автомат, а $q \in Q$ такое его состояние, что $q \in \left\langle \frac{A}{X} \right\rangle_K$. Пусть $u \in \mathcal{L}_K^{\rm io}(q)$; тогда считаем, что $u \in \zeta_{\mathcal{BE}(L)}\left(\frac{A}{X}\right)$. Никакое иное слово не принадлежит языку $\zeta_{\mathcal{BE}(L)}\left(\frac{A}{X}\right)$. Расширенный автомат $\mathcal{BE}(L)$ определяется как $\mathcal{BE}(L) = (\mathcal{BA}(L), \zeta_{\mathcal{BE}(L)})$.

Теорема 4. $\mathcal{L}(\mathcal{BE}(L)) = L$.

Утверждение 2. *Если q — состояние автомата $\mathcal{BE}(L)$ и $q \xrightarrow[\mathcal{BE}(L)]{u} q$, то $u \in \zeta_{\mathcal{BE}(L)}(q)$.*

Определения 4 и 5 естественным образом определяют автомат с объединенными и продублированными состояниями соответственно.

Определение 4. Для автомата (1) и двух его состояний $q_1, q_2 \in Q$ записью $\mathcal{J}^{q_1 q_2}(K)$ обозначим автомат, чей граф переходов получается из графа переходов автомата K следующим образом:

- для каждого $r \in Q$ множества дуг $\gamma(q_1, r)$ заменяются на множества $\gamma(q_1, r) \cup \gamma(q_2, r)$;
- аналогично, для каждого $r \in Q$ множества дуг $\gamma(r, q_1)$ заменяются на множества $\gamma(r, q_1) \cup \gamma(r, q_2)$;
- состояние q_2 удаляется из автомата вместе со входящими и выходящими дугами.

Кроме того, состояние q_1 автомата $\mathcal{J}^{q_1 q_2}(K)$ является стартовым (финальным) тогда и только тогда, когда по крайней мере одно из состояний q_1 и q_2 является стартовым (соответственно финальным) для K.

Определение 5. Для автомата (1) и его состояния $q \in Q$ записью $\mathcal{R}^q(K)$ обозначим автомат, чей граф переходов получается из графа переходов автомата K следующим образом:

- состояние q переименовывается в q_1 и добавляется состояние q_2;
- для всех $r \in Q$ создаются множества $\gamma(q_2, r)$ такие, что $\gamma(q_1, r) = \gamma(q_2, r)$;
- аналогично, для каждого $r \in Q$ создаются множества дуг $\gamma(r, q_2)$ такие, что $\gamma(r, q_1) = \gamma(r, q_2)$.

Кроме того, состояние q_2 автомата $\mathcal{R}^q(K)$ является стартовым (финальным) тогда и только тогда, когда состояние q_1 является стартовым (соответственно финальным) для K.

Теорема 5. *Пусть даны автоматы K_1 и K_2, определяющие один и тот же регулярный язык. С помощью конечного количества применений двух примитивов $\mathcal{J}^{q_1 q_2}$ и $\mathcal{R}^q$ из K_1 можно получить K_2.*

4. Заключение

Теорема 2 является результатом совместной работы авторов. Теорема 5 сформулирована и доказана М.Р. Сайфуллиной. Остальные результаты статьи принадлежат Б.Ф. Мельникову.

Вопросы, рассмотренные в данной статье, могут быть продолжены в следующих направлениях.

Сформулировать различные варианты достаточных условий для функции разметки состояний исходного НКА, при которых объединение состояний не изменяет значений функции разметки.

Исследовать автоматы без выходных состояний (а возможно также и без стартовых состояний [7]), изучить возможности объединения двух состояний таких автоматов, при которых остаются без изменений определяемые этими автоматами языки (точнее, ω-языки и 2ω-языки).

Литература

[1] Брауэр В. *Введение в теорию конечных автоматов.* – М.: Радио и связь, 1987. – 392 с.
[2] Саломаа А. *Жемчужины теории формальных языков.* – М.: Мир, 1986. – 159 с.
[3] Мельников Б.Ф. *Мультиэвристический подход к задачам дискретной оптимизации* // Кибернетика и системный анализ (НАН Украины). – 2006. – № 3. – С. 32–42.
[4] Melnikov B., Sciarini-Guryanova N. *Possible edges of a finite automaton defining the given regular language* // The Korean J. Comput. and Appl. Math. – 2002. – V. 9. – № 2. – P. 475–485.
[5] Han Y.-S., Wood D. *The generalization of generalized automata: expression automata* // The 9th Internat. Confer. on Implement. and Appl. of Automata. – 2004. – P.114–122.
[6] Melnikov B. *On an expansion of nondeterministic finite automata* // J. Appl. Math. Comput. – 2007. – V. 24. – № 1. – P. 155–165.
[7] Мельников Б.Ф. *Об ω-языках специальных биллиардов* // Дискретная матем. (РАН). – 2002. – № 3. – С. 95–108.

Б.Ф. Мельников

профессор, кафедра прикладной математики и информатики,
Тольяттинский государственный университет,
445667, г. Тольятти, ул. Белорусская, д. 14,

e-mail: B.Melnikov@tltsu.ru

М.Р. Сайфуллина

аспирант, кафедра прикладной математики и информатики,
Тольяттинский государственный университет,
445667, г. Тольятти, ул. Белорусская, д. 14,

e-mail: M.Sayfullina@tltsu.ru, mariasayfullina@gmail.com

B.F. Melnikov

Professor, Chair of Applied Mathematics and Informtion Science,
Togliatti State University, 14 Belorusskaya str., Togliatti, 445667 Russia,

e-mail: B.Melnikov@tltsu.ru

M.R. Saifullina

Postgraduate, Chair of Applied Mathematics and Informtion Science,
Togliatti State University,
14 Belorusskaya str., Togliatti, 445667 Russia,

e-mail: M.Sayfullina@tltsu.ru, mariasayfullina@gmail.com