УДК 519.178

ОБ ОДНОМ АЛГОРИТМЕ ПОСТРОЕНИЯ УНИВЕРСАЛЬНОГО АВТОМАТА КОНВЕЯ

М. А. Зубова, Б. Ф. Мельников

Тольятинский государственный университет

Поступила в редакцию 8.02.2012 г.

Аннотация. В статье получено доказательство совпадения универсального автомата Конвея и автомата $\mathcal{COM}$, определяющегося через множество всевозможных дуг всех автоматов для заданного регулярного языка. Следствием этого доказательства является возможность замены алгоритма построения универсального автомата на алгоритм построения автомата $\mathcal{COM}$.
Ключевые слова. Универсальный автомат Конвея, множество возможных дуг.

Abstract. In this article, the proof of equality of the universal automaton and automaton $\mathcal{COM}$ was received, where automaton $\mathcal{COM}$ is defined by the set of the possible edges of the regular language. A consequence of this is the possibility of replacing the algorithm for constructing the universal automaton on the algorithm for constructing automaton $\mathcal{COM}$.
Keywords. Conway's universal automaton; the set of possible edges.

1. ВВЕДЕНИЕ

Каждому языку L можно поставить в соответствие минимальный детерминированный автомат. Такой автомат конечен в том и только том случае, когда L является регулярным языком. Кроме того, для любого автомата, распознающего язык L, существует гомоморфизм этого автомата на минимальный детерминированный автомат.

Ещё одним возможным автоматом-эквивалентом (автоматом-инвариантом) языка L является так называемый универсальный автомат, о нём и пойдёт речь в данной статье. Универсальный автомат впервые был определён английским математиком и создателем клеточного автомата «Игра “Жизнь”» Дж. Конвеем в [1], где был использован для поиска наилучшего приближения семейства заданных регулярных языков к языку L.

При исследовании недетерминированных автоматов, задающих язык L, универсальный автомат имеет такую же важность, какую имеет минимальный детерминированный автомат в отношении детерминированных автоматов для этого же языка. Как и в случае с минимальным детерминированным автоматом, универсальный автомат конечен тогда и только тогда, когда L регулярен. К тому же, универсальный автомат обладает другими полезными свойствами — например: всякий задающий язык L минимальный (недетерминированный) автомат является подавтоматом универсального, определяющего этот язык.

2. ПРИМЕНЯЕМЫЕ ОБОЗНАЧЕНИЯ

Пусть $K = \{Q, \Sigma, \delta, S, F\}$ — недетерминированный конечный автомат Рабина—Скотта, определённый согласно [2, 3]. Однако, в отличие от [2, 3], мы будем рассматривать автомат без ε-переходов, т.е. функция переходов δ автомата K будет иметь вид $\delta\colon Q \times \Sigma \to P(Q)$, где $P(Q)$ — множество всех подмножеств множества Q.

Зеркальный автомат для заданного автомата K будем обозначать K^R — т.е. $K^R = \{Q, \Sigma, \delta^R, S, F\}$; здесь $q' \in \delta^R(q, a)$ тогда и только тогда, когда $q \in \delta^R(q', a)$. Очевидно, что K^R определяет язык L^R.

Через $\mathcal{L}_K^{in}(q)$ и $\mathcal{L}_K^{out}(q)$ обозначим входной и выходной языки состояния q — т.е. языки, определяемые автоматами $(Q, \Sigma, \delta, S, \{q\})$ и $(Q, \Sigma, \delta, \{q\}, F)$ соответственно.

Согласно определению, данному в [4, гл. 6, п. 1], автомат $K' = \{Q', \Sigma, \delta', S', F'\}$ называется *подавтоматом* автомата K, если выполняются условия:

- $Q' \subseteq Q$, $S' \subseteq (S \cap Q')$, $F' \subseteq (F \cap Q')$;
- $\delta'(q,a) \subseteq \delta(q,a)$ для любых $q \in Q'$ и $a \in \Sigma$.

Очевидно, что подавтомат определяет язык-подмножество, который в некоторых случаях совпадает с исходным языком.

Определение 1.1. Будем называть автомат $K^* = \{Q^*, \Sigma, \delta^*, S^*, F^*\}$ *псевдоподавтоматом* автомата K, если он для некоторой функции $h\colon Q^* \to Q$ удовлетворяет следующим условиям:

• любому состоянию q^* автомата K^* соответствует некоторое состояние $q = h(q^*)$ автомата K;

• для каждой пары состояний $q_1{}^*$ и $q_2{}^*$ автомата K^* условие $q_2^* \in \delta^*(q_1^*, a)$ выполняется тогда и только тогда, когда для автомата K и его состояний $h(q_1^*)$ и $h(q_2^*)$ выполняется $h(q_2^*) \in \delta(h(q_1^*), a)$.

Определение 1.2. Для любых $a \in \Sigma$, $q_1^*, q_2^* \in Q^*$ дуга $q_1^* \xrightarrow{a} q_2^*$ соответствует дуге $h(q_1^*) \xrightarrow{a} h(q_2^*)$, если выполняются условия $q_2^* \in \delta^*(q_1^*, a)$ и $h(q_2^*) \in \delta(h(q_1^*), a)$.

Очевидно выполнение следующего утверждения.

Утверждение. $\mathcal{L}_{K^*}^{in}(q_1^*) \subseteq \mathcal{L}_{K}^{in}(h(q_1^*))$ и $\mathcal{L}_{K^*}^{out}(q_2^*) \subseteq \mathcal{L}_{K}^{out}(h(q_2^*))$.

Для псевдоподавтомата определим наполненный псевдоподавтомат.

Определение 1.3. *Наполненным* является псевдоподавтомат, в котором можно выделить подмножество дуг, которое соответствует множеству дуг исходного автомата.

3. ОПРЕДЕЛЕНИЕ УНИВЕРСАЛЬНОГО АВТОМАТА

Универсальный автомат имеет несколько определений ([1, 5] и др.). Приведём одно из них [5].

Определение 2.1. *Псевдофакторизацией* заданного регулярного языка L называется такая пара языков (X, Y) над алфавитом Σ, что для языка $XY = \{uv \mid u \in X, v \in Y\}$ выполнено условие $XY \subseteq L$.

Определение 2.2. *Факторизацией* регулярного языка L называется максимальная псевдофакторизация (X, Y), т.е. если $X \subseteq X'$, $Y \subseteq Y'$ и $X'Y' \subseteq L$, то $X = X'$ и $Y = Y'$.

Обозначим множество всех факторизаций языка L как $\mathcal{F}_L$.

Определение 2.3. Универсальным автоматом U_L для заданного регулярного языка L называется пятёрка $U_L = \{\mathcal{F}_L, \Sigma, \delta_L, S_L, F_L\}$, где:

• $\delta_L : \mathcal{F}_L \times \Sigma \to \mathrm{P}(\mathcal{F}_L)$ и $(X', Y') \in \delta_L((X, Y), a)$, если $XaY' \subseteq L$;

• $S_L = \{(X, Y) \in \mathcal{F}_L \mid \varepsilon \in X\}$;

• $F_L = \{(X, Y) \in \mathcal{F}_L \mid \varepsilon \in Y\}$.

Отметим, что данное определение *не* является конструктивным, т.е. оно не даёт алгоритма построения универсального автомата. Для получения такого алгоритма нужны дополнительные построения, которые не нужно выполнять при применении подхода, предлагаемого авторами.

4. ОПРЕДЕЛЕНИЕ АВТОМАТА $\mathcal{COM}$

Следующие определения были приведены в [6], а до того частично — в [7].

Определение 3.1. Пара подмножеств $Q \subseteq \mathcal{Q}$ и $R \subseteq \mathcal{R}$ ($\mathcal{Q}, \mathcal{R}$ — множества состояний канонического автомата к данному автомату и канонического автомата к инверсному соответственно) образует *псевдоблок*, если для любых $A \in Q$ и $X \in R$ выполняется $A \# X$ [3]. Если, кроме того, любая пара множеств Q' и R', таких что $Q \subseteq Q' \subseteq \mathcal{Q}$ и $R \subseteq R' \subseteq \mathcal{R}$, образует псевдоблок тогда и только тогда, когда $Q' = Q$ и $R' = R$, то пара Q и R образует *блок*.

Обозначим псевдоблок записью $B(Q, R)$. Для такого псевдоблока B через $\alpha(B)$ обозначим соответствующие ему состояния $Q \subseteq \mathcal{Q}$, а через $\beta(B)$ — соответствующие состояния $R \subseteq \mathcal{R}$.

Определение 3.2. Положим $\mathcal{B}_2 \in \delta_{\mathcal{B}}(\mathcal{B}_1, a)$ при выполнении следующих условий:

• для любого $q \in \alpha(\mathcal{B}_1)$ выполняется $\delta_{\mathcal{Q}}(q, a) \subseteq \alpha(\mathcal{B}_2)$;

• для любого $q \in \beta(\mathcal{B}_2)$ выполняется $\delta_{\mathcal{R}}(q, a) \subseteq \beta(\mathcal{B}_1)$;

где $\delta_{\mathcal{Q}}$ и $\delta_{\mathcal{R}}$ — функции переходов канонических автоматов, задающих языки L и L^R соответственно.

Определение 3.3. Положим $\mathcal{B}_1 \in S_{\mathcal{B}}$ при выполнении следующих условий:

• найдётся такое $q_1 \in \alpha(\mathcal{B}_1)$, что $q_1 \in S_{\mathcal{Q}}$;

• найдётся такое $q_2 \in \beta(\mathcal{B}_1)$, что $q_2 \in F_{\mathcal{R}}$;

где $S_{\mathcal{Q}}$ и $F_{\mathcal{R}}$ — множество стартовых состояний канонического автомата к данному и множество финальных состояний канонического к инверсному соответственно.

Определение 3.4. Положим $\mathcal{B}_1 \in F_{\mathcal{B}}$ при выполнении следующих условий:

• найдётся такое $q_1 \in \alpha(\mathcal{B}_1)$, что $q_1 \in F_{\mathcal{Q}}$;

• найдётся такое $q_2 \in \beta(\mathcal{B}_1)$, что $q_2 \in S_{\mathcal{R}}$;

где $F_{\mathcal{Q}}$ и $S_{\mathcal{R}}$ — множество финальных состояний канонического автомата к данному и множество стартовых состояний канонического к инверсному соответственно.

Объединив всё вышеперечисленное, определим автомат $\mathcal{COM}$.

Определение 3.5. *Автоматом* $\mathcal{COM}$ для заданного языка L называется пятёрка

$$\mathcal{COM}(L) = \{\mathcal{B}, \Sigma, \delta_{\mathcal{B}}, S_{\mathcal{B}}, F_{\mathcal{B}}\},$$

где

- $\mathcal{B}$ — множество всех блоков для L;
- $S_{\mathcal{B}} \subseteq \mathcal{B}$ множество стартовых состояний;
- $F_{\mathcal{B}} \subseteq \mathcal{B}$ множество финальных состояний.

Отметим, что определение 3.5 даёт *алгоритм* построения автомата $\mathcal{COM}$.

5. СОВПАДЕНИЕ АВТОМАТА КОНВЕЯ И АВТОМАТА $\mathcal{COM}$

Пусть для заданного регулярного языка L построены универсальный автомат и автомат $\mathcal{COM}$.

Теорема. $U_L = \mathcal{COM}(L)$ (с точностью до переобозначения состояний).

Доказательство. Универсальный автомат задаёт язык L [5], и всякий автомат, определяющий этот язык (в том числе и $\mathcal{COM}(L)$), является подавтоматом универсального [8]. Автомат $\mathcal{COM}(L)$, кроме того, является его наполненным псевдоподавтоматом, что следует из определений, приведённых в разделах 2 и 3.

Из [6] следует, что автомат $\mathcal{COM}$ — инвариант языка. Согласно [7], любой автомат для заданного языка является подавтоматом и псевдоподавтоматом $\mathcal{COM}$'а. В частности, универсальный автомат является подавтоматом и наполненным псевдоавтоматом, что следует из определения факторизации.

Итак, автомат $\mathcal{COM}$ является наполненным псевдоподавтоматом универсального, а универсальный автомат, в свою очередь, является наполненным псевдоподавтоматом автомата $\mathcal{COM}$. Следовательно, согласно приведённому выше определению наполненного псевдоподавтомата, они совпадают.

Отметим, что доказывать это утверждение можно и алгоритмически (т.е. согласно приведённым в разделах 3 и 4 определениям), однако при этом доказательство получается очень громоздким. А предлагаемое в данной статье простое доказательство не содержит алгоритмических построений.

Определения раздела 4 дают алгоритм построения универсального автомата. По-видимому, он проще приведённого в [5] (если заранее заданы оба канонических автомата — т.е. для заданного и инверсного языков); однако подробное исследование сложности соответствующих алгоритмов — это возможная тема для дальнейшей работы.

ЛИТЕРАТУРА

1. *Conway J. H.* Regular Algebra and Finite Machines / J. H. Conway — L: Chapman and Hall, 1971. — 147 p.

2. *Melnikov B.* Extended nondeterministic finite automata / B. Melnikov // Fundamenta Informaticae. — Vol. 104 (2010) No.3. — P. 255—265.

3. *Мельников Б.Ф.* Недетерминированные конечные автоматы: монография / Б. Ф. Мельников. — Тольятти: Изд-во ТГУ, 2009. — 160 с.

4. *Лаллеман Ж.* Полугруппы и комбинаторные приложения: пер. с англ. / Ж. Лаллеман. — М.: Мир, 1985. — 440 с.

5. *Lombardy S.* The Universal Automaton / S. Lombardy, J. Sakarovitch // Logic and Automata, Texts in Logic and Games, Vol. 2, Amsterdam Univ. Press. — Amsterdam, 2008. — P. 457—504.

6. *Мельников Б. Ф.* Построение автомата COM на основе базисного автомата / Б. Ф. Мельников, М. А. Зубова // Вектор науки Тольяттинского государственного университета. — Тольятти, 2010. — № 4 (14). — С. 30—32.

7. *Melnikov B.* Possible edges of a finite automaton defining a given regular language / B. Melnikov, N. Sciarini-Guryanova // The Korean J. Comp. and Appl. Math. — Vol. 9 (2002) No.2. — P. 475—485.

8. *Carrez C.* On the minimalization of non-deterministic automaton: technical report / C. Carrez. — Computing Laboratory of the Science Faculty of Lille University, 1970.

Зубова М. А. — студентка кафедры «Прикладная математика и прикладная информатика», Тольяттинский государственный университет

E-mail: nai_ 999801@mail.ru

Мельников Б. Ф. —доктор физико-математических наук, профессор кафедры «Прикладная математика и прикладная информатика», Тольяттинский государственный университет

Тел.: (8482) 53-95-14

E-mail: B.Melnikov@tltsu.ru

Zubova M. A. — student of the Chair "Applied Mathematics and Applied Informatics", Togliatti State University

E-mail: nai_ 999801@mail.ru

Melnikov B. F. Doctor of Physical and Mathematical Sciences, Professor of the Chair "Applied Mathematics and Applied Informatics", Togliatti State University.

Tel.: (8482) 48-70-51

E-mail: B.Melnikov@tltsu.ru