

На правах рукописи

МЕЛЬНИКОВА
Александра Александровна

**БАЗИСНЫЙ
КОНЕЧНЫЙ АВТОМАТ
КАК ИНВАРИАНТ
РЕГУЛЯРНОГО ЯЗЫКА**

01.01.09 – дискретная математика и математическая кибернетика

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата физико-математических наук

Казань – 2013

Работа выполнена в Димитровградском инженерно-технологическом институте – филиале ФГАОУ ВПО «Национальный исследовательский ядерный университет “Московский инженерно-физический институт”» на кафедре «Высшая математика».

Научный руководитель: доктор физико-математических наук
профессор Мельников Борис Феликсович
(заведующий кафедрой «Прикладная математика и информатика» Тольяттинского филиала ФГБОУ ВПО «Самарский государственный университет»)

Официальные оппоненты: доктор физико-математических наук
доцент Захаров Владимир Анатольевич
(заведующий лабораторией математических проблем компьютерной безопасности ФГБОУ ВПО «Московский государственный университет им. М. В. Ломоносова»)

кандидат физико-математических наук
доцент Богомолов Алексей Сергеевич
(доцент кафедры математической кибернетики и компьютерных наук ФГБОУ ВПО «Саратовский государственный университет имени Н. Г. Чернышевского»)

Ведущая организация: ФГБОУ ВПО «Самарский государственный университет путей сообщения»

Защита диссертации состоится 27 декабря 2013 г. в 14:30 часов на заседании диссертационного совета Д 212.081.24 при Казанском (Приволжском) федеральном университете по адресу: г. Казань, ул. Кремлёвская, д. 35, корпус 2, аудитория 1011.

С диссертацией можно ознакомиться в библиотеке Казанского (Приволжского) федерального университета.

Отзывы по данной работе в двух экземплярах, заверенные печатью организации, просим направлять по адресу: 420008, г. Казань, ул. Кремлёвская, д. 35, диссертационный совет Д 212.081.24.

Автореферат разослан « » ноября 2013 г.

Учёный секретарь
диссертационного совета Д 212.081.24,
кандидат физико-математических наук, доцент
Еникеев А.И.

1 Общая характеристика работы

Актуальность темы. В 1960-х годах активно развивалась теория конечных автоматов Рабина-Скотта, в литературе было очень много соответствующих научных публикаций¹. Но примерно с начала 1970-х годов, с появлением развитой теории детерминированных конечных автоматов, в научных статьях часто стало прослеживаться мнение, что в теории регулярных языков и конечных автоматов уже решены практически все возможные задачи. Когда же данная теория стала применяться к различным вопросам смежных наук, оказалось, что остаётся важным способ, которым задаётся регулярный язык, а не только сам исследуемый язык². Именно подобные проблемы приводят к необходимости рассматривать представление регулярных языков с помощью базисных конечных автоматов, которым и посвящена настоящая диссертация. Базисный автомат является инвариантом заданного регулярного языка, альтернативным каноническому автомату. Развиваемая в диссертации теория применяется для решения различных задач теории регулярных языков, а также некоторых задач дискретной оптимизации, причём как точными, так и эвристическими методами.

Недетерминированные конечные автоматы (ниже – НКА), изучаемые в диссертационной работе, впервые рассматривались в середине прошлого века Ю.Медведевым³, а также М.Рабином и Д.Скоттом⁴. Позднее, прежде всего – при применении НКА к решению различных прикладных задач, указанные автоматы модифицировались и обобщались многими авторами⁵. Важным обобщением стали недетерминированные автоматы-распознаватели с магазинной памятью (push-down automata, МП-автоматы), появившиеся как одно из средств автоматического перевода, а также для различных целей программирования, широко используемые в теории трансляции⁶. Как МП-автоматы, так и НКА служат примером так называемых автоматов-распознавателей – в отличие от конечных автоматов-преобразователей (например, конечных автоматов Мили, Мура), которые в представляемой диссертационной работе не рассматриваются. Конечные автоматы и такие тесно связанные с ними конструкции, как, например, линейные грамматики и регулярные выражения, относятся к основным понятиям информатики.

Различные варианты конечных автоматов и близкие им математические объекты служат для описания и анализа технических устройств, различных систем и процессов, программ и алгоритмов⁷. Эффективность применения именно конечных автоматов

¹ Подробные библиографии работ 1960–70-х годов можно найти, например, в следующих монографиях:

- Кудрявцев В., Алёшин С., Подколзин А. Введение в теорию автоматов // М. – Наука. – 1985. – 319 с.
- Иванов Н., Михайлов Г., Руднев В., Таль А. Конечные автоматы: эквивалентность и поведение // М. – Наука. – 1984. – 192 с.

² Sheng Yu. *A renaissance of automata theory?* // Bulletin of the Eur. Assoc. Theor. Comput. Sci., No. 72. – 2000. – 270–272.

³ Медведев Ю. *О классе событий, допускающих представление в конечном автомате* // Автоматы. – М. – Иностранная Литература. – 1956. – 385–401.

⁴ Rabin M., Scott D. *Finite automata and their decision problems* // IBM J. Research and Development. – 1959. – Vol. 3. – 115–125.

⁵ См., например:

- Гилл А. Линейные последовательностные машины // М. – Наука. – 1974. – 288 с.
- Чирков М. Основы общей теории конечных автоматов // Л. – Изд-во ЛГУ. – 1975. – 280 с.
- Melnikov V. *On an expansion of nondeterministic finite automata* // Journal of Applied Math. and Computing. – Springer Berlin/Heidelberg. – 2007. – V.24. – No.1-2. – 155–165.
- Твердохлебов В. Геометрические образы законов функционирования автоматов // Саратов: ООО Издательский Центр «Научная книга». – 2008. – 183 с.

⁶ См., например: Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции // М. – Мир. – 1978. – Т. 1. – 308 с.

⁷ См., например: Брауэр В. Введение в теорию конечных автоматов // М. – Радио и связь. – 1987.

в различных задачах теории регулярных языков связана с тем, что, в отличие от некоторых других задающих регулярный язык формализмов, конечный автомат является вычислительной моделью⁸.

Итак, в различных прикладных задачах теории формальных языков желательно получать экономное (с различных точек зрения) представление недетерминированных конечных автоматов. Это обосновано, например, тем, что экономное представление автомата в памяти часто связано с ускорением работы алгоритмов, моделирующих функционирование автоматов. При разных способах представления автомата в памяти компьютера более важными могут быть либо минимально возможное число вершин, либо минимально возможное число дуг или минимальное число вершин, и минимальное суммарное число вершин и дуг⁹. Среди других возможных областей применения можно назвать такие основные, как теория трансляции, теория «естественных» языков, теория искусственного интеллекта, математическое моделирование. Особо отметим возможное применение исследованных автором диссертации функций разметки для создания алгоритмов для ДНК-компьютеров¹⁰.

Основные задачи исследования заключались в разработке и применении подхода к анализу регулярных языков и соответствующих им недетерминированных конечных автоматов – с использованием определённых и исследованных автором базисных конечных автоматов.

Объектом исследования являлись недетерминированные конечные автоматы Рабина-Скотта (Медведева), в первую очередь – определённые автором базисные НКА.

Предметом исследования являлись алгоритмы работы с недетерминированными конечными автоматами.

Методы исследования. В качестве аппарата исследований применялись методы доказательства теорем дискретной математики и методы разработки алгоритмов.

Результаты исследования. Результатами диссертационного исследования являются новые утверждения и теоремы теории формальных языков, а также новые вычислительные алгоритмы.

Научная новизна. Новые научные результаты, полученные в диссертации, заключаются в следующем: на основе специального бинарного отношения $\#$, связанного с исследованными автором т.н. функциями разметки состояний, описан автомат для заданного регулярного языка, названный базисным автоматом. Этот новый автомат является ещё одним, наряду с каноническим автоматом, инвариантом регулярного языка, однако часто в конкретных случаях лучше отражающим структуру языка. С помощью этого базисного автомата получаются различные свойства рассматриваемого регулярного языка, в том числе – свойства функций разметки состояний. Указанные свойства применены в различных задачах теории регулярных языков и конечных автоматов,

– 392 с.

⁸ См. в вышеуказанных монографиях, а также, например, в следующей: Hromkovič J. *Algorithmics for Hard Problems. Introduction to Combinatorial Optimization, Randomization, Approximation, and Heuristics* // Springer. – 2003. – 544 p.

⁹ См., например:

- Hashiguchi K. *Algorithms for determining the smallest number of nonterminals (states) sufficient for generating (accepting) a regular language* // ICALP. – 1991. – 641–648.
- Kameda T., Weiner P. *On the state minimization of nondeterministic finite automata* // IEEE Trans. on Computers. – 1970. – C-19. – 617–627.
- Jiang T., Ravikumar B. *Minimal NFA Problems are Hard* // SIAM J. Comput. – 1993. – V. 22, No. 1. – 1117–1141.

¹⁰ Работы в этом направлении только начались. См.: Крайнюков Н., Мельников Б. *Функция разметки состояний и базисные слова для конечных автоматов* // Межд. конф. «Совр. проблемы математики, информатики и биоинформатики», посвящённая 100-летию А.А.Ляпунова. – 2010. – 65–66.

прежде всего – при решении проблемы дуговой минимизации.

Доказанные утверждения и теоремы, а также описанные алгоритмы эквивалентного преобразования недетерминированных конечных автоматов являются оригинальными разработками автора диссертации. Некоторые утверждения доказаны совместно с научным руководителем.

Практическая значимость исследования. Полученные результаты могут быть применены при доказательстве теорем, связанных с регулярными языками и недетерминированными конечными автоматами, а также при описании алгоритмов эквивалентного преобразования конечных автоматов¹¹, в том числе – при построении автоматов, эквивалентных заданному¹². Кроме того, результаты диссертации могут быть применены в различных задачах минимизации недетерминированных конечных автоматов – в вершинной¹³, дуговой ([18]) и т.н. звёздно-высотной минимизации¹⁴, а также минимизации по некоторым комбинированным критериям. Для каждой из этих задач может помочь построение т.н. универсального автомата Конвэя – которое также удаётся выполнить на основе базисного автомата¹⁵. При этом результаты диссертации (прежде всего – результаты главы 5) могут быть применены и для создания *эвристических* алгоритмов решения перечисленных задач дискретной оптимизации¹⁶. Результаты представляемой диссертации применяются также в смежных областях – а именно, для проверки репрезентативности случайно сгенерированных дискретных структур и описания специальных подходов к кластеризации¹⁷ и для описания специальных математических моделей¹⁸.

Достоверность результатов подтверждается приведёнными в диссертации доказательствами утверждений и теорем.

¹¹ Мельников Б., Сайфуллина М. *О некоторых алгоритмах эквивалентного преобразования недетерминированных конечных автоматов* // Известия ВУЗов. Математика. – 2009. – №3. – 67–71.

¹² Например, схожие конструкции были применены в следующей статье, опубликованной через 9 лет после статьи [17] соискателя данной диссертации: van Glabbeek R., Ploeger B. *Five determinisation algorithms* // In: O.Ibarra and B.Ravikumar, editors. Implementation and Applications of Automata. – Springer, LNCS. – 2008. – V.5148. – 161–170.

¹³ См., например:

- статьи, указанные в двух предыдущих сносках;
- следующую диссертацию: Зузанова М. *Некоторые алгоритмы эквивалентного преобразования недетерминированных конечных автоматов* // Дисс. . . . канд. физ.-мат. наук: 05.13.18, Тольяттинский государственный университет, Д 212.264.03, защищена 01.07.10;
- а также цитируемые ниже статьи, выполненные под рук. Цыганова.

¹⁴ Баумгертнер С., Мельников Б. *Мультиэвристический подход к проблеме звёздно-высотной минимизации недетерминированных конечных автоматов* // Вестник Воронежского государственного университета, «Системы управления и информационные технологии». – 2010. – №1. – 5–7.

¹⁵ Мельников Б., Зубова М. *Построение автомата СМ на основе базисного автомата* // Вектор науки Тольяттинского государственного университета. – 2010. – №4. – 30–32.

¹⁶ Про применение базисных конечных автоматов для эвристических алгоритмов звёздно-высотной минимизации см. вышеуказанную статью Баумгертнер и Мельникова. Про применение для эвристических алгоритмов вершинной минимизации см. вышеуказанную диссертацию Зузановой, а также следующую серию статей:

- Цыганов А., Булычев О. *HeO: библиотека метаэвристик для задач дискретной оптимизации* // Программные продукты и системы. – 2009. – №4. – 148–151.
- Цыганов А., Булычев О. *Параллельные эвристические алгоритмы для задачи вершинной минимизации недетерминированных конечных автоматов* // Вестник Воронежского государственного университета, «Системы управления и информационные технологии». – 2010. – №1. – 60–63.

¹⁷ Мельников Б., Пивнева С., Рогова О. *Репрезентативность случайно сгенерированных конечных автоматов с точки зрения базисных автоматов* // Стохастическая оптимизация в информатике (изд-во СПбГУ). – 2010. – Том 6. – 74–82.

¹⁸ Также см. вышеуказанную диссертацию Зузановой.

Основные положения, выносимые на защиту.

- Определение функций разметки состояний заданного конечного автомата.
- Определение базисного конечного автомата как инварианта регулярного языка, альтернативного каноническому автомату.
- Описание двух алгоритмов построения базисного конечного автомата.
- Формулировка свойств базисного конечного автомата:
 - эквивалентность базисного автомата и канонического;
 - однозначность базисного автомата – т.е. единственность последовательности состояний базисного автомата при принятии им некоторого слова рассматриваемого языка;
 - утверждение о связи любого допускающего пути любого автомата, определяющего заданный регулярный язык, с соответствующим путём базисного автомата – а именно, непустое пересечение значений функций разметки соответствующих вершин;
 - утверждения о свойствах таблицы состояний конечного автомата;
 - утверждения о свойствах функций разметки.

Доказательство соответствующих утверждений о свойствах базисного автомата.

- Применение основных свойств базисного автомата для описания вспомогательных алгоритмов, применяемых для алгоритмов вершинной минимизации недетерминированных конечных автоматов.
- Описание двух алгоритмов дуговой минимизации недетерминированных конечных автоматов на основе определения базисного конечного автомата. Доказательство корректности описанных алгоритмов.

Апробация работы. Основные научные и практические результаты диссертации были представлены на следующих российских и международных научных конференциях и семинарах.

- Научная конференция «Математическое моделирование физических, экономических, социальных систем и процессов», Ульяновск, УлГУ, 1998 и 1999 г.
- Международная алгебраическая конференция памяти А.Г.Куроша, Москва, МГУ, 1998 г.
- VII Международный семинар «Дискретная математика и её приложения», Москва, МГУ, 2001 г.
- Международная научно-практическая конференция «Методы и алгоритмы прикладной математики в технике, медицине и экономике», Новочеркасск 2001 г.
- Научный семинар для аспирантов кафедры МАТИС (под руководством д.ф.-м.н. В.Бувевича и С.Алёшина), Москва, МГУ, 2001 и 2002 г.г.
- Научный семинар (под руководством д.ф.-м.н. Г.Осипова), Переславль-Залесский, ИПСРАН, 2001 и 2003 г.г.
- XIII Международная конференция «Проблемы теоретической кибернетики», Казань, 2002 г.

- Международная научно-практическая конференция «Проблемы математического образования и культуры», Тольятти, ТГУ, 2003 г.
- Международная конференция «Общие проблемы управления и их приложения. Проблемы преподавания математики» Тамбов, ТГУ, 2007 и 2009 г.г.
- Научный семинар по математической кибернетике и дискретной математике при Диссертационном совете Д 212.081.24, Казань, КГУ (КПФУ), 2009, 2011 и 2013 г.г.
- Научный семинар кафедры математического анализа Ульяновского государственного педагогического университета, 2010 и 2012 г.г.
- Научный семинар по дискретной математике, Димитровград, Филиал УлГУ и ДИТИ – филиал НИЯУ «Московский инженерно-физический институт», 1999–2013 г.г.

Публикации. Основные результаты диссертации опубликованы в 20 работах, 7 из которых входят в список изданий, рекомендованных ВАК. Список работ приведён в конце автореферата.

Личный вклад. Изложенные утверждения и теоремы доказаны лично автором, либо совместно с научным руководителем. Результаты, полученные в совместных с научным руководителем работах [1, 5, 6, 8, 14, 16, 18, 19], принадлежат в части, касающейся базисных автоматов, автору диссертации.

Структура и объём диссертации. Общий объём диссертационной работы – 102 страницы 14-го кегля. Диссертация состоит из 5 глав (включая введение и заключение) и списка используемой литературы из 106 наименований. В тексте диссертации содержатся 31 рисунок и 26 таблиц.

2 Содержание работы

Глава 1 («Введение»)

В главе 1 формулируются цели и задачи работы, даётся краткий исторический обзор, освещается актуальность и новизна исследуемой темы, конкретизируется предмет исследования – недетерминированный конечный автомат Рабина-Скотта (Медведева) для некоторого заданного регулярного языка L , говорится о различных способах задания автомата. В работе в первую очередь применяется задание НКА для некоторого регулярного языка L пятёркой вида $K = (Q, \Sigma, \delta, S, F)$, где:

- Q – некоторое конечное множество состояний (вершин);
- Σ – рассматриваемый алфавит (автомат K вводится для задания языка L над этим алфавитом);
- δ – функция переходов вида $\delta : Q \times (\Sigma \cup \{\varepsilon\}) \rightarrow \mathcal{P}(Q)$, (где ε – пустое слово, \mathcal{P} – множество всех подмножеств данного множества);
- $S \subseteq Q$ – множество стартовых состояний (входов);
- $F \subseteq Q$ – множество финальных состояний (выходов).

В работе используется также функция переходов вида $\gamma : Q \times Q \rightarrow \mathcal{P}(\Sigma \cup \{\varepsilon\})$, такая что для некоторой пары состояний $q_1, q_2 \in Q$ и некоторой буквы $a \in \Sigma$ условие $\gamma(q_1, q_2) \ni a$ считается выполненным при $\delta(q_1, a) \ni q_2$. Вводится определение зеркального к K конечного автомата $K^R = (Q, \Sigma, \delta^R, F, S)$, для которого $\delta^R(q_1, a) \ni q_2$ выполнено тогда

и только тогда, когда $\delta(q_2, a) \ni q_1$. Автомат K^R задаёт зеркальный (инверсный) к L язык L^R .

Через $\mathcal{L}_K^{in}(q)$ и $\mathcal{L}_K^{out}(q)$ обозначаются входной и выходной языки состояния q , т.е. языки, определяемые автоматами

$$(Q, \Sigma, \delta, S, \{q\}) \quad \text{и} \quad (Q, \Sigma, \delta, \{q\}, F)$$

соответственно. После получения детерминированного всюду определённого автомата¹⁹

$$\dot{K} = (\dot{Q}, \Sigma, \dot{\delta}, \{\dot{s}\}, \dot{F})$$

на основе некоторого недетерминированного автомата K , рассмотрим на множестве, содержащем все неупорядоченные пары различных элементов \dot{Q} , а также специальный элемент \mathcal{F} , следующее бинарное отношение \succ :

- $\{q_1, q_2\} \succ \{q_3, q_4\}$, если для некоторого $a \in \Sigma$, имеем $\dot{\delta}(q_1, a) = \{q_3\}$ и $\dot{\delta}(q_2, a) = \{q_4\}$.
- $\{q_1, q_2\} \succ \mathcal{F}$, если из двух состояний q_1, q_2 ровно одно финальное.

Рассмотрим \succ^+ – транзитивное замыкание отношения \succ . При этом $\mathcal{L}_K^{out}(q_1) = \mathcal{L}_K^{out}(q_2)$ тогда и только тогда, когда не выполнено условие $\{q_1, q_2\} \succ^+ \mathcal{F}$. Если выходные языки состояний q_1 и q_2 равны (т.е. эти состояния эквивалентны), то можно объединить эти состояния (при этом мы без ограничения общности считаем, что q_2 – не вход):

- для всех вершин $r \in \dot{Q}$ множество дуг $\dot{\gamma}(r, q_1)$ заменяем на множество $\dot{\gamma}(r, q_1) \cup \dot{\gamma}(r, q_2)$
- удаляем q_2 и все дуги множеств $\dot{\gamma}(r, q_2)$ и $\dot{\gamma}(q_2, r)$.

Задающий язык L канонический автомат \tilde{L} , единственный с точностью до переобозначения состояний, получается из \dot{K} объединением всех пар эквивалентных состояний. \dot{K} будет всюду определённым, если добавить (при необходимости) одно бесполезное состояние N .

Кроме того, во введении излагаются основные результаты и описывается структура диссертации. В частности отмечается, что исследование соответствующих языков с помощью новых, определённых автором данной работы базисных конечных автоматов важно для различных прикладных задач – например, для задач минимизации НКА по различным критериям.

Глава 2 («Базисные конечные автоматы – определения, примеры и алгоритмы построения»)

Во 2-й главе определяется новый объект в теории конечных автоматов Рабина-Скотта – так называемый базисный автомат для заданного регулярного языка L . Этот автомат является единственным для данного языка L , причём является его инвариантом. Исследование базисных автоматов и является предметом представляемой диссертации.

В разделе 2.1 рассматриваются дополнительные определения, необходимые для введения понятия базисного автомата. А именно, определяется бинарное отношение $\#$ на множестве пар состояний двух канонических автоматов – для заданного языка L и для инверсного языка L^R .

Для произвольного НКА

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

¹⁹ Всюду определённый (total) автомат определяется согласно следующей монографии: Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции // М. – Мир. – 1978. – Т. 1.

обозначим через $\mathcal{L}(K)$ язык, определяемый K . Регулярный язык L первоначально задаётся некоторым автоматом K , т.е. $L = \mathcal{L}(K)$.

Автомат

$$\tilde{L} = (\mathcal{Q}, \Sigma, \delta_{\mathcal{Q}}, \{s_{\mathcal{Q}}\}, F_{\mathcal{Q}}) \quad (2)$$

– эквивалентный автомат в канонической форме, единственный с точностью до переобозначения состояний, а

$$\tilde{L}^R = (\mathcal{R}, \Sigma, \delta_{\mathcal{R}}, \{s_{\mathcal{R}}\}, F_{\mathcal{R}}) \quad (3)$$

автомат в канонической форме, определяющий зеркальный к L язык L^R .

Бинарное отношение $\#$ определяется следующим образом:

$$A\#X \quad \text{тогда и только тогда, когда} \quad (\exists uv \in L) \left(u \in \mathcal{L}_{\tilde{L}}^{in}(A), v^R \in \mathcal{L}_{\tilde{L}^R}^{in}(X) \right),$$

где $\# \subseteq \mathcal{Q} \times \mathcal{R}$, а \mathcal{Q} и \mathcal{R} – множества состояний автоматов в канонической форме \tilde{L} и \tilde{L}^R соответственно. Очевидно, что существуют несложные алгоритмы построения этого бинарного отношения, т.е. приведённое определение конструктивно.

В разделе 2.2 приводится основное введённое автором определение – определение базисного автомата.

Определение 1 Автомат

$$(\mathcal{T}, \Sigma, \delta_{\mathcal{T}}, S_{\mathcal{T}}, F_{\mathcal{T}})$$

над некоторым алфавитом Σ для следующих множеств \mathcal{T} , $\delta_{\mathcal{T}}$, $S_{\mathcal{T}}$ и $F_{\mathcal{T}}$:

- $\mathcal{T} = \{(A, X) \mid \text{для любого } A \in \mathcal{Q}, X \in \mathcal{R}, A\#X\}$ (обозначим A как $\alpha(T)$, X как $\beta(T)$ для $T \in \mathcal{T}$);
- для любых $T_1, T_2 \in \mathcal{T}$, $a \in \Sigma$ мы имеем $\delta_{\mathcal{T}}(T_1, a) \ni T_2$ если и только если $\delta_{\mathcal{Q}}(\alpha(T_1), a) \ni \alpha(T_2)$ и $\delta_{\mathcal{R}}(\beta(T_2), a) \ni \beta(T_1)$;
- $T \in S_{\mathcal{T}}$ если и только если $\alpha(T) = s_{\mathcal{Q}}$ (это означает, что $\alpha(T)$ – единственное начальное состояние \tilde{L} , т.е. $s_{\mathcal{Q}}$);
- $T \in F_{\mathcal{T}}$ если и только если $\beta(T) = s_{\mathcal{R}}$ (это означает, что $\beta(T)$ – единственное начальное состояние автомата \tilde{L}^R , т.е. $s_{\mathcal{R}}$)

назовём **базисным автоматом для данного языка L** и обозначим через $\mathcal{BA}(L)$.

Это определение конструктивно – в том смысле, что является также алгоритмом построения базисного автомата, аналогично определению бинарного отношения $\#$. Существуют и альтернативные способы построения базисного автомата – см. ниже, в разделе 2.9 представляемой диссертации. Отметим, что для заданного регулярного языка L , в соответствие с определением, строится единственный базисный автомат $\mathcal{BA}(L)$.

В разделе 2.3 рассматриваются определения прямой и обратной функций разметки. Они применяются при построении бинарного отношения $\#$, при построении базисного автомата, а также могут применяться при решении многих других задач. Впервые эти функции были исследованы в работе автора [10].

Для заданного регулярного языка L рассмотрим определяющий его произвольный НКА K , а также произвольный всюду определённый автомат (по определению являющийся детерминированным)

$$Z = (\mathcal{Q}_Z, \Sigma, \delta_Z, \{s_Z\}, F_Z).$$

Для этой пары автоматов K и Z $\varphi_{KZ}^{in} : \mathcal{Q} \rightarrow \mathcal{P}(\mathcal{Q}_Z)$ является полной функцией разметки (в дальнейшем, для краткости, функцией разметки), если для каждого $q \in \mathcal{Q}$ и $\tilde{q} \in \mathcal{Q}_Z$ множество $\varphi_{KZ}^{in}(q)$ содержит \tilde{q} в том и только том случае, если

$$\mathcal{L}_K^{in}(q) \cap \mathcal{L}_Z^{in}(\tilde{q}) \neq \emptyset. \quad (4)$$

Последнее определение можно считать и алгоритмом построения функции φ_{KZ}^{in} , поскольку $\mathcal{L}_K^{in}(q)$ и $\mathcal{L}_Z^{in}(\tilde{q})$ – регулярные языки. Отметим, что в процессе построения эквивалентного канонического автомата, выполненного специальным образом, можно получить и более простой алгоритм построения этой функции, что используется в примерах раздела 2.8 данной диссертации.

В разделе 2.4 рассмотрен подробный пример построения базисного автомата для языка, заданного регулярным выражением $(a + ab + ba)^*$. Граф переходов автомата K , одного из НКА, задающих данный регулярный язык, показан на рис. 1. Изменив на этом рисунке направления стрелок на противоположные, получим граф переходов зеркального к K автомата K^R , задающего инверсный язык L^R .

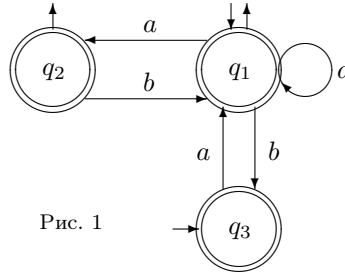


Рис. 1

Соответствующий конечный автомат в канонической форме – \tilde{L} . Графы переходов автоматов \tilde{L} и \tilde{L}^R приведены на рис. 2 и рис. 3.

В работе не рассматривается бесполезное состояние N автоматов в канонической форме – смысл от этого не меняется, а формулировка становится менее громоздка. Несложно доказываться, что для бесполезных состояний канонических автоматов не может выполняться ни $N\#X$, ни $A\#N$ для любых состояний A, X этих автоматов, поэтому бесполезное состояние не может входить в состояние базисного автомата. В графах переходов канонических автоматов бесполезные состояния для удобства также не указываются. Это делает граф более наглядным.

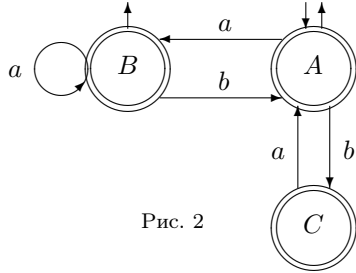


Рис. 2

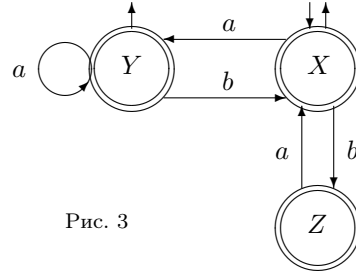


Рис. 3

В рассматриваемом нами примере из равенства $L = L^R$ следует, что графы переходов автоматов \tilde{L} и \tilde{L}^R совпадают. Состояния автомата \tilde{L}^R обозначены через X, Y и Z (вместо A, B и C).

При этом прямая и обратная функции разметки имеют значения:

$$\varphi_K^{in}(q_1) = \{A, B\}, \quad \varphi_K^{in}(q_2) = \{B\}, \quad \varphi_K^{in}(q_3) = \{A, C\};$$

$$\varphi_K^{out}(q_1) = \{X, Y\}, \quad \varphi_K^{out}(q_2) = \{X, Z\}, \quad \varphi_K^{out}(q_3) = \{Y\}.$$

Данные значения функций разметки можно получить в процессе канонизации автоматов K и K^R , а также исходя из определения. По последнему, общими для входных языков, например, являются слова:

$$aab \in \mathcal{L}_K^{in}(q_1) \cap \mathcal{L}_L^{in}(A), \quad a \in \mathcal{L}_K^{in}(q_1) \cap \mathcal{L}_L^{in}(B), \quad \mathcal{L}_K^{in}(q_1) \cap \mathcal{L}_L^{in}(C) = \emptyset ;$$

$$a \in \mathcal{L}_K^{in}(q_2) \cap \mathcal{L}_L^{in}(B), \quad \mathcal{L}_K^{in}(q_2) \cap \mathcal{L}_L^{in}(A) = \mathcal{L}_K^{in}(q_2) \cap \mathcal{L}_L^{in}(C) = \emptyset ;$$

$$ab \in \mathcal{L}_K^{in}(q_3) \cap \mathcal{L}_L^{in}(A), \quad b \in \mathcal{L}_K^{in}(q_3) \cap \mathcal{L}_L^{in}(C) \quad \mathcal{L}_K^{in}(q_3) \cap \mathcal{L}_L^{in}(B) = \emptyset ;$$

таким образом мы получаем значения прямой функции разметки.

Для получения значений обратной функции разметки ($\varphi_K^{out}(q) = \varphi_{K^R}^{in}(q)$) можно считать, что:

$$aab \in \mathcal{L}_{K^R}^{in}(q_1) \cap \mathcal{L}_{L^R}^{in}(X), \quad a \in \mathcal{L}_{K^R}^{in}(q_1) \cap \mathcal{L}_{L^R}^{in}(Y), \quad \mathcal{L}_{K^R}^{in}(q_1) \cap \mathcal{L}_{L^R}^{in}(Z) = \emptyset ;$$

и так далее, в силу $L = L^R$.

Отношение $\#$ может быть задано²⁰ следующей таблицей 1:

	X	Y	Z
A	#	#	
B	#	#	#
C		#	

Таблица 1

По определению автомата $BA(L)$,

$$\mathcal{T} = \{(A, X), (A, Y), (B, X), (B, Y), (B, Z), (C, Y)\},$$

и $BA(L)$ может быть определён следующей таблицей:

		a	b
→ ←	(A, X)	(B, Z)	(C, Y)
→	(A, Y)	(B, X), (B, Y)	
←	(B, X)	(B, Z)	(A, Y)
	(B, Y)	(B, X), (B, Y)	
	(B, Z)		(A, X)
	(C, Y)	(A, X), (A, Y)	

Таблица 2

где стартовые состояния помечены символом \rightarrow , а финальные – символом \leftarrow .

Не останавливаясь подробно на всём процессе построения этой таблицы, рассмотрим только состояние (B, X) . Оно не является начальным, так как $\alpha((B, X)) = B$ – не начальное состояние автомата \tilde{L} . Но (B, X) является финальным в силу того, что $\beta((B, X)) = X$ – начальное состояние \tilde{L}^R . Состояние (B, X) имеет две исходящих дуги, а именно:

- в состоянии (B, Z) , помеченная a , так как \tilde{L} имеет дугу из B в B , помеченную a , и \tilde{L}^R имеет дугу из Z в X , помеченную также a ;
- в состоянии (A, Y) , помеченная b , так как \tilde{L} имеет дугу из B в A , помеченную b , и \tilde{L}^R имеет дугу из Y в X , помеченную также b .

Остальные начальные и финальные состояния автомата, определяемого таблицей 2, а также его дуги, выбираются тем же самым путём. Граф переходов полученного базисного автомата показан на рис. 4.

²⁰ Сам алгоритм построения отношения $\#$ описан, например, в следующей статье: Melnikov B. *A new algorithm of the state-minimization for the nondeterministic finite automata* // The Korean J. of Comp. and App. Math. – V. 6. – No. 2. – 1999. – 277–290.

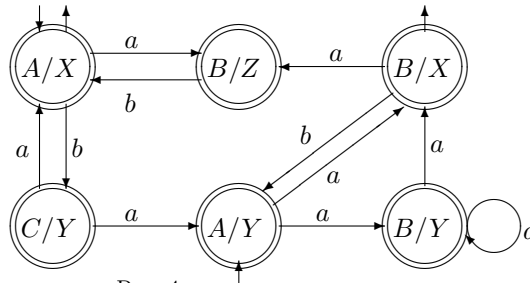


Рис. 4

Таким образом нами построен базисный автомат, задающий регулярный язык, определяемый регулярным выражением $(a + ab + ba)^*$. Этот пример используется в работе и далее.

В разделе 2.5 доказывается основное утверждение о базисном автомате – его эквивалентность определяемому им языку.

Утверждение 1 Автомат $\mathcal{BA}(L)$ эквивалентен автомату \tilde{L} .

В разделе 2.6 доказывается однозначность базисного автомата²¹, т. е. единственность последовательности его состояний при принятии им некоторого слова, принадлежащего рассматриваемому регулярному языку. Произвольный недетерминированный конечный автомат, читая слово, проходит последовательность состояний. Если такая последовательность единственна, то соответствующий автомат называется, аналогично другим подобным случаям в теории формальных языков, однозначным (unambiguous). В противном случае, когда некоторому слову языка автомата соответствует две или более последовательности состояний, – неоднозначным (ambiguous). Однозначные автоматы могут быть удобнее при использовании для решения различных проблем. Например, существует аналогия между последовательностью подмножеств контекстно-свободных языков, а значит соответствующих магазинных (push-down) автоматов, и следующей последовательностью подмножеств множества конечных автоматов: произвольные автоматы, однозначные автоматы, детерминированные автоматы, автоматы в канонической форме.

В связи с последним в работе формулируется и доказывается следующее утверждение для базисного автомата.

Утверждение 2 Базисный автомат для данного регулярного языка является однозначным.

Эквивалентная формулировка этого утверждения такова.

Утверждение 3 Каждое слово языка автомата $\mathcal{BA}(L)$ определяет единственную последовательность состояний $\mathcal{BA}(L)$.

В разделе 2.7 доказывается свойство допускающего пути любого автомата для заданного регулярного языка, т. е. свойство автомата $\mathcal{BA}(L)$, связывающее его с произвольным автоматом K , допускающим рассматриваемый регулярный язык.

Предположим, что слово u принадлежит языку L и число букв u равно n .

²¹ Про однозначные и неоднозначные автоматы см., например, в следующей статье: Мельников Б. Однозначные конечные автоматы // Известия ВУЗов. Серия «Технические науки» (Поволжский регион). – 2002. – №2. Важно отметить, что сформулированные в ней критерии однозначности используют функции разметки состояний, и, следовательно, напрямую связаны с вопросами, рассматриваемыми в данной диссертации.

Рассмотрим единственную последовательность состояний автомата $\mathcal{BA}(L)$, такую что $\mathcal{BA}(L)$ читает буквы слова u и проходит свои состояния в следующем порядке: $(A_0, X_0), (A_1, X_1), (A_2, X_2), \dots, (A_n, X_n)$.

Рассмотрим также одну из последовательностей состояний, которые проходит автомат K при чтении слова u (для простоты будем рассматривать автоматы без ε -переходов):

$$q_0, q_1, q_2, \dots, q_n.$$

Утверждение 4 $\varphi_K^{in}(q_i) \ni A_i$ и $\varphi_K^{out}(q_i) \ni X_i$ для любого $i = 0, 1, \dots, n$.

Указанные свойства связывают базисный автомат с любым недетерминированным автоматом K , допускающим рассматриваемый регулярный язык, следующим образом: базисный автомат «моделирует» работу автомата K . Эти свойства делают новый автомат более удобным для использования в некоторых специальных задачах, связанных с теорией регулярных языков, что указано в разделе 1.1 представляемой диссертации и подробно показано в главах, следующих за данной главой²². Можно сказать, что примеры выполнения этих свойств нами приведены на основе автомата, рассмотренного ранее (в разделе 2.4).

В разделе 2.8 рассматриваются ещё два примера построения базисного автомата. Эти примеры демонстрируют выполнение свойств, описанных в предыдущих разделах.

В разделе 2.9 приведён альтернативный алгоритм построения дуг базисного автомата. Кратко опишем его, применяя следующие обозначения. Дуги автоматов \tilde{L} и \tilde{L}^R будем обозначать заглавными греческими буквами, не совпадающими по написанию с латинскими – до буквы Ξ для автомата \tilde{L} и начиная с буквы Π для автомата \tilde{L}^R , а также для зеркального к последнему автомата $(\tilde{L}^R)^R$. Для некоторой конкретной дуги Γ , идущей из вершины A в вершину B и имеющей пометку $a \in \Sigma$, будем писать $\alpha^a(\Gamma) = A$ и $\beta^a(\Gamma) = B$.

Выберем некоторую букву $a \in \Sigma$; описанную здесь процедуру мы проделаем для каждой буквы алфавита. Пусть δ_Q^a – помеченные буквой a дуги автомата \tilde{L} (т.е. элементы множества δ_Q), а δ_R^a – помеченные буквой a дуги автомата \tilde{L}^R (т.е. элементы множества δ_R). Аналогично сказанному выше, будем таким же образом обозначать соответствующее множество дуг автомата $(\tilde{L}^R)^R$.

Рассмотрим бинарное отношение $\#^a \subseteq \delta_Q^a \times \delta_R^a$, определённое следующим образом. Для некоторых $\Gamma \in \delta_Q^a$ и $\Omega \in \delta_R^a$ полагаем $\Gamma \#^a \Omega$ тогда и только тогда, когда для некоторого слова $w \in L$ имеем представление $w = uav$, и при этом:

$$\begin{aligned} u &\in \mathcal{L}_{\tilde{L}}^{in}(\alpha^a(\Gamma)), \quad u \in \mathcal{L}_{(\tilde{L}^R)^R}^{in}(\alpha^a(\Omega)), \\ v &\in \mathcal{L}_{\tilde{L}}^{out}(\beta^a(\Gamma)), \quad v \in \mathcal{L}_{(\tilde{L}^R)^R}^{out}(\beta^a(\Omega)). \end{aligned} \quad (5)$$

Приведённое определение отношения $\#^a$ с помощью (5) можно рассматривать как алгоритм его построения.

Утверждение 5 Пусть $\Gamma \in \delta_Q$ и $\Omega \in \delta_R$ – некоторые дуги канонических автоматов для языков L и L^R соответственно. Тогда в базисном автомате $\mathcal{BA}(L)$ имеется дуга

$$\delta_T \left((\alpha^a(\Gamma), \beta^a(\Gamma)), a \right) \ni (\alpha^a(\Omega), \beta^a(\Omega)) \quad (6)$$

²² Ранее эти свойства нашли применение в различных работах, связанных с алгоритмами эквивалентного преобразования НКА, – см., например, практически все публикации автора представляемой диссертации. А среди последних вариантов применения можно отметить следующие статьи:

- Melnikov B. *On an expansion of nondeterministic finite automata* // J. of Applied Mathematics and Computing, Springer Berlin/Heidelberg. – Vol. 24, No. 1–2. – 2007. – 155–165.
- Мельников Б., Сайфуллина М. *О некоторых алгоритмах эквивалентного преобразования недетерминированных конечных автоматов* // Известия ВУЗов. Математика. – 2009. – №3. – 67–71.

тогда и только тогда, когда $\Gamma \#^a \Omega$.

Глава 3 («Свойства базисных конечных автоматов»)

3-я глава посвящена свойствам базисных конечных автоматов и некоторым алгоритмам работы с ними. Отметим, что все доказанные в 3-й главе утверждения применяются во всех трёх вышеназванных задачах минимизации – причём как в точных, так и в эвристических алгоритмах.

В разделе 3.1 формулируются и доказываются некоторые используемые далее в работе свойства функции разметки²³.

Утверждение 6

$$\mathcal{L}_K^{out}(q) \subseteq \bigcup_{\tilde{q} \in \varphi_K^{in}(q)} \mathcal{L}_{\tilde{L}}^{out}(\tilde{q}).$$

Утверждение 7

$$\mathcal{L}_K^{in}(q) \subseteq \left(\bigcup_{\tilde{q} \in \varphi_K^{out}(q)} \mathcal{L}_{\tilde{L}^R}^{out}(\tilde{q}) \right)^R.$$

Утверждение 8

$$\mathcal{L}_K^{in}(q) \cdot \left(\bigcap_{\tilde{q} \in \varphi_K^{in}(q)} \mathcal{L}_{\tilde{L}}^{out}(\tilde{q}) \right) \subseteq \mathcal{L}(K).$$

Утверждение 9

$$\left(\bigcap_{\tilde{q} \in \varphi_K^{out}(q)} \mathcal{L}_{\tilde{L}^R}^{out}(\tilde{q}) \right)^R \cdot \mathcal{L}_K^{out}(q) \subseteq \mathcal{L}(K).$$

В разделе 3.2 приводится вариант объединения тех состояний недетерминированного автомата (1), которые имеют одинаковые значения функций φ^{in} или φ^{out} . Это делается аналогично тому, как два эквивалентных состояния могут быть объединены в случае детерминированных автоматов.

Определение 2 Для некоторой пары состояний НКА (1), т.е. для $q_1, q_2 \in Q$, определим автомат $\mathcal{J}^{q_1 q_2}(K)$, граф переходов которого получается из графа переходов автомата K следующим образом.

- Для каждого состояния $r \in Q$, соответствующее множество дуг $\gamma(q_1, r)$ заменяется на множество $\gamma(q_1, r) \cup \gamma(q_2, r)$.
- Аналогично, для каждого состояния $r \in Q$, множество дуг $\gamma(r, q_1)$ заменяется на множество $\gamma(r, q_1) \cup \gamma(r, q_2)$.²⁴
- Состояние q_2 отбрасывается; соответствующие элементы функции переходов γ также отбрасываются.

²³ Они были приведены без доказательств в статье Melnikov В. *A new algorithm of the state-minimization for the nondeterministic finite automata* // The Korean J. of Comp. and App. Math. – 1999. – V. 6. – No. 2. – 277–290. В представляемой диссертации приводятся соответствующие доказательства.

²⁴ Поэтому, если имеются какие-либо из следующих четырёх случаев: $q', q'' \in \{q_1, q_2\}$, тогда каждая дуга $\gamma(q', q'') \ni a$ заменяется на петлю $\gamma(q_1, q_1) \ni a$.

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

Утверждения 10 и 11 (см. [19]) определяют первый алгоритм объединения состояний²⁵.

Утверждение 10 Пусть для автомата (1) и некоторых его состояний $q_1, q_2 \in Q$ выполнено равенство:

$$\varphi_K^{in}(q_1) = \varphi_K^{in}(q_2). \quad (7)$$

Тогда для автомата $K_1 = \mathcal{J}^{q_1 q_2}(K)$ выполняется следующее равенство:

$$\mathcal{L}(K) = \mathcal{L}(K_1). \quad (8)$$

Кроме того,

$$\varphi_K^{in}(q_1) = \varphi_{K_1}^{in}(q_1). \quad (9)$$

Утверждение 11 Пусть для автомата (1) и некоторых его состояний $q_1, q_2 \in Q$ выполнено следующее равенство:

$$\varphi_K^{out}(q_1) = \varphi_K^{out}(q_2). \quad (10)$$

Тогда для автомата $K_1 = \mathcal{J}^{q_1 q_2}(K)$ выполняется равенство

$$\mathcal{L}(K) = \mathcal{L}(K_1); \quad \text{кроме того, } \varphi_K^{out}(q_1) = \varphi_{K_1}^{out}(q_1).$$

Далее, в разделе 3.3, приведены примеры применения алгоритма объединения состояний недетерминированного автомата.

В разделе 3.4 описываются свойства входных и выходных языков базисного автомата, которые могут также быть названы свойствами таблицы соответствующих состояний. В частности, это свойства, связывающие входные и выходные языки базисного автомата с входными и выходными языками состояний канонических автоматов (канонического \tilde{L} для заданного регулярного языка, канонического к зеркальному \tilde{L}^R , а также автомата $(\tilde{L}^R)^R$).

Доказано, что канонический автомат \tilde{L}^R содержит не более $2^n - 1$ состояний (где n – число состояний автомата \tilde{L})²⁶. Этим же числом ограничивается число возможных столбцов таблицы отношения $\#$, в которой n строк. Кроме того, доказано, что в таблице соответствия состояний не может быть одинаковых строк и столбцов.

Следующие утверждения 12–15 доказываются на основе определения базисного автомата.

Утверждение 12 Пусть задан регулярный язык L , его канонический автомат \tilde{L} и некоторое состояние A автомата \tilde{L} . Тогда для каждого состояния X автомата \tilde{L}^R выполнено следующее:

$$\mathcal{L}_{\tilde{L}}^{in}(A) = \mathcal{L}_{\mathcal{BA}(L)}^{in}((A, X)).$$

²⁵ Аналогично утверждениям 6 и 8, эти утверждения были приведены без доказательств в статье [19]. Доказательства были впоследствии опубликованы в следующей статье: Melnikov V. *Once more on the edge-minimization of nondeterministic finite automata and the connected problems* // Fundamenta Informaticae – 2010. – V. 104. – No. 3. – 267–283. В представляемой диссертации соответствующие доказательства не рассматриваются – однако рассматриваются примеры применения, связанные с основным объектом диссертации – базисными конечными автоматами.

²⁶ В представляемой диссертации приведено более полное доказательство, чем доказательство, приведённое в вышеупомянутой диссертации Зузановой. То же самое относится к ещё двум утверждениям разделов 3.4 и 3.5.

Утверждение 13 Пусть задан регулярный язык L , канонический для L^R автомат $(\widetilde{L^R})^R$, и X – некоторое состояние $\widetilde{L^R}$.²⁷ Тогда для каждого состояния A автомата \widetilde{L} выполнено следующее:

$$\mathcal{L}_{(\widetilde{L^R})^R}^{out}(X) = \left(\mathcal{L}_{\widetilde{L^R}}^{in}(X) \right)^R = \mathcal{L}_{\mathcal{BA}(L)}^{out}((A, X)).$$

Далее, как и выше, применяются обозначения (в частности – \mathcal{Q} и \mathcal{R}), введённые ранее в (2) и (3).

Утверждение 14 Пусть задан регулярный язык L и его канонический автомат \widetilde{L} . Тогда для некоторого состояния A автомата \widetilde{L} выполнено следующее:

$$\mathcal{L}_{\widetilde{L}}^{out}(A) = \bigcup_{X \in \mathcal{Q}} \mathcal{L}_{\mathcal{BA}(L)}^{out}((A, X)).$$

Утверждение 15 Пусть задан регулярный язык L , и $\widetilde{L^R}$ – канонический автомат, определяющий L^R . Тогда для некоторого состояния X автомата $\widetilde{L^R}$ выполнено следующее:

$$\left(\mathcal{L}_{\widetilde{L^R}}^{out}(X) \right)^R = \mathcal{L}_{(\widetilde{L^R})^R}^{in}(X) = \bigcup_{A \in \mathcal{R}} \mathcal{L}_{\mathcal{BA}(L)}^{in}((A, X)).$$

Утверждение 16 Пусть канонический автомат \widetilde{L} для данного регулярного языка L имеет по крайней мере два различных состояния²⁸, и A, B – некоторая пара таких состояний. Тогда существует состояние X канонического автомата $\widetilde{L^R}$, такое что базисный автомат для языка L содержит в точности одно состояние из следующих двух: (A, X) и (B, X) .

Утверждение 17 Пусть канонический автомат $\widetilde{L^R}$, определяющий язык L^R , имеет по крайней мере 2 различных состояния, и X, Y – некоторая пара таких состояний. Тогда существует состояние A канонического автомата \widetilde{L} , такое что базисный автомат для L содержит ровно 1 состояние из следующих: (A, X) и (A, Y) .

В разделе 3.5 показано, что при выполнении всех сформулированных выше ограничений соответствующий регулярный язык возможен для любой таблицы отношения $\#$. Последнее доказательство является конструктивным – т.е. в работе приведён алгоритм построения конкретного базисного автомата, имеющего заданную таблицу отношения $\#$.

Глава 4 («Применение базисных автоматов в задачах минимизации недетерминированных конечных автоматов»)

4-я глава посвящена применению базисных автоматов в различных задачах минимизации недетерминированных конечных автоматов, в первую очередь – в «точных» (неэвристических) алгоритмах.

В разделе 4.1 определяются специальные объекты, также связанные с базисными автоматами. Это – блоки и псевдоблоки, состоящие из пары множеств вершин. Блоки и псевдоблоки используются в дальнейших построениях (во всех разделах 4-й и 5-й глав), определения иллюстрируются примерами.

²⁷ Поэтому X может также рассматриваться как состояние автомата $(\widetilde{L^R})^R$.

²⁸ Как и ранее, возможное бесполезное состояние канонического автомата не рассматриваем.

Определение 3 Пара непустых подмножеств $Q \subseteq \mathcal{Q}$ и $R \subseteq \mathcal{R}$ образует псевдоблок, если выполняется следующее условие: $(\forall A \in Q, X \in R) (A \# X)$.

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 будем говорить, что пара Q и R образует блок.

При этом здесь, как и ранее, не будет рассматриваться бесполезное состояние N , так как согласно вышеизложенному ни одно из рассматриваемых подмножеств состояний (Q и R) не может его содержать. Кроме того, очевидно, что множество блоков является подмножеством множества псевдоблоков.

Для каждого псевдоблока \mathcal{B} через $\alpha(\mathcal{B})$ обозначим соответствующие ему состояния $Q \subseteq \mathcal{Q}$, а через $\beta(\mathcal{B})$ – соответствующие состояния $R \subseteq \mathcal{R}$. При этом будем также обозначать данный псевдоблок записью $\mathcal{B}(Q, R)$.

Также в разделе 4.1 формулируется *необходимое* условие минимальности автомата по числу вершин²⁹. Это условие было получено на основе развитой автором представляемой диссертации теории базисных автоматов и может рассматриваться как одно из приложений этой теории.

Определение 4 Пусть \mathcal{B}_1 и \mathcal{B}_2 – некоторые псевдоблоки. Будем говорить, что \mathcal{B}_2 покрывает \mathcal{B}_1 , если

$$\alpha(\mathcal{B}_1) \subseteq \alpha(\mathcal{B}_2) \quad \text{и} \quad \beta(\mathcal{B}_1) \subseteq \beta(\mathcal{B}_2).$$

Определение 5 Назовём некоторое множество псевдоблоков полным, если для каждого элемента $T \in \mathcal{Q} \times \mathcal{R}$ существует некоторый псевдоблок \mathcal{B} этого множества, такой что $T \in \mathcal{B}$.³⁰

Необходимое условие минимальности автомата по числу вершин заключается в том, что множество его вершин образует полное множество псевдоблоков. При таком множестве блоков (т.е. при выполнении этого необходимого условия) имеется следующая формулировка *достаточного* условия минимальности автомата по числу вершин: его множество дуг включает все возможные дуги, удовлетворяющие ограничениям, сформулированным в следующем разделе 4.2.

Возможные дуги конечного автомата для заданного регулярного языка рассматриваются в разделе 4.2. Множество возможных дуг *любого* недетерминированного конечного автомата, определяющего заданный регулярный язык, описывается на множестве псевдоблоков (вместо вершин заданного автомата)³¹. Описанное множество возможных дуг применяется к следующей проблеме, рассматриваемой в данной главе диссертации, – проблеме дуговой минимизации.

В разделе 4.3 рассматривается первый вариант алгоритма дуговой минимизации недетерминированного конечного автомата. Автомат, полученный в результате выполнения данного алгоритма, имеет минимальное число дуг среди всех автоматов, определяющих данный язык. Как уже отмечалось во введении, такие автоматы могут быть

²⁹ См. также следующие статьи:

- Melnikov B. A new algorithm of the state-minimization for the nondeterministic finite automata. – The Korean Journal of Computational and Applied Mathematics. – Vol. 6, No. 2. – 1999. – 277–290.
- Melnikov B. Once more about the state-minimization of the nondeterministic finite automata. – The Korean Journal of Computational and Applied Mathematics. – Vol. 7, No. 3. – 2000. – 655–662.

³⁰ Данная формулировка несколько отличается от формулировок определений, приведённых в статьях, упомянутых в предыдущей сноске.

³¹ Подробнее см. следующую статью: Melnikov B., Sciarini-Guryanova N. Possible edges of a finite automaton defining the given regular language // The Korean J. of Comp. and App. Math. – V. 9. – No. 2. – 2002. – 475–485.

удобны при решении различных практических задач, возникающих в теории регулярных языков³².

Алгоритм 1

Вход: заданный регулярный язык L .

Выход: конечный автомат, определяющий L и имеющий минимально возможное количество дуг.

Метод.

- Зафиксировать некоторое множество псевдоблоков на множестве $Q \times R$. При этом можно считать, что выполняются следующие два условия (т.е. выбирать только такие множества псевдоблоков, которые удовлетворяют им обоим)³³:
 - 1) это множество является полным;
 - 2) ни один из псевдоблоков множества не покрывает никакой другой³⁴.
- Рассматривая каждый из этих псевдоблоков как вершину искомого автомата, определить среди них входные и выходные вершины.
- Определить множество возможных дуг между вершинами (пусть M).
- Перебрав все возможные множества псевдоблоков (удовлетворяющих сформулированным выше условиям), а для каждого из таких множеств – все подмножества построенного нами множества M , определить те из них, при которых полученный автомат:
 - действительно определяет язык L ;
 - имеет минимально возможное число дуг.

Также в разделе 4.3 приводятся соответствующие примеры.

В разделе 4.4 рассматривается второй вариант алгоритма дуговой минимизации конечного автомата. Эта версия напрямую использует базисный автомат для данного регулярного языка в описании алгоритма, – в то время как в первом случае, в разделе 4.3, базисный автомат используется в доказательстве корректности приведённого алгоритма.

Алгоритм 2

Вход: заданный регулярный язык L .

Выход: конечный автомат, определяющий L и имеющий минимально возможное количество дуг.

Метод.

- Построить множество блоков.
- Рассматривая каждый из блоков как вершину искомого автомата, определить среди них входные и выходные вершины.
- Построить множество дуг δ' следующим образом. Для некоторых двух вершин q_1 и q_2 (допускаем возможность $q_1 = q_2$) и некоторой буквы $a \in \Sigma$ считаем

³² Алгоритмы этого и следующего разделов приведены в статье [18] соискателя данной диссертации. Русский вариант алгоритма приводится согласно следующей вышеупомянутой монографии: Мельников Б. Недетерминированные конечные автоматы // Тольятти. – Изд-во ТГУ, 2009.

³³ Заранее отметим, что множество блоков, рассматриваемых как псевдоблоки, этим условиям удовлетворяет.

³⁴ В том числе – нет двух псевдоблоков с одинаковыми значениями α либо β .

$\delta'(q_1, a) \ni q_2$, если для некоторых $A \in \alpha(q_1)$, $X \in \beta(q_1)$, $B \in \alpha(q_2)$ и $Y \in \beta(q_2)$ в базисном автомате существует дуга

$$\delta_T((A, X), a) \ni (B, Y). \quad 35 \quad (11)$$

- Перебирая подмножества множества δ' , определить то из них (искомое множество δ), при котором полученный автомат:
 - действительно определяет язык L ;
 - имеет минимально возможное число дуг.

Глава 5 («Заключение»)

В заключении работы приводятся её основные результаты.

3 Основные результаты диссертации

- Определены прямая и обратная функций разметки, применяемые во многих задачах, связанных с недетерминированными конечными автоматами. Исследованы основные свойства этих функций.
- Определён инвариант регулярного языка – базисный конечный автомат, являющийся альтернативой другому инварианту регулярного языка – каноническому конечному автомату.
- Описаны два алгоритма построения базисного конечного автомата, доказана корректность этих алгоритмов.
- Исследованы свойства базисного конечного автомата:
 - доказана эквивалентность базисного автомата и канонического;
 - доказана однозначность базисного автомата;
 - доказано утверждение о связи любого допускающего пути любого автомата, определяющего заданный регулярный язык, с соответствующим путём базисного автомата – т.е. непустое пересечение значений функций разметки всех соответствующих вершин;
 - доказаны утверждения о свойствах таблицы состояний конечного автомата;
 - доказаны утверждения о свойствах функций разметки.
- Доказанные свойства базисного автомата применены для описания новых вариантов алгоритмов вершинной минимизации НКА – причём как точных, так и псевдооптимальных алгоритмов реального времени.
- Впервые описаны два алгоритма дуговой минимизации недетерминированных конечных автоматов, доказана корректность этих алгоритмов.

³⁵Построение автомата с функцией переходов δ' в данном алгоритме можно *неформально* описать следующим образом. Для заданного языка строим базисный автомат, после чего делаем в нём копии всех вершин – в количестве, равном количеству вхождений пары функций разметки состояний вершины в блоки. Далее объединяем «размноженные» вершины согласно описанию множества блоков. Таким образом мы описываем ещё один возможный вариант *применения* базисных автоматов. Подробнее см. процитированную выше диссертацию Зузановой.

Список литературы

- [1] Melnikov B., Melnikova A. *A new algorithm of constructing the basis finite automaton* // Informatika (Lithuania). – Vol. 13. – No. 3. – Lithuania. – 2002. – 299–310. **(Рецензируемый журнал, рекомендованный ВАК.)**
- [2] Мельникова А.А. *Базисные автоматы в решении проблемы оптимизации* // Вестник Тамбовского университета. Сер. Естественные и техн. науки. – Т. 12. – Вып. 4. – Тамбов: Изд-во ТГУ. – 2007. – 492–494. **(Рецензируемый журнал, рекомендованный ВАК.)**
- [3] Мельникова А.А. *Применение свойств конечных автоматов в алгоритмах вершинной минимизации* // Вестник Тамбовского университета. Сер. Естественные и техн. науки. – Т. 14. – Вып.4. – Тамбов: Изд-во ТГУ. – 2009. – 763–765. **(Рецензируемый журнал, рекомендованный ВАК.)**
- [4] Мельникова А.А. *К вопросу о сходимости значений динамических функций риска* // Вестник транспорта Поволжья. – Т. 24. – Вып.4. – Самара: Изд-во СамГУПС. – 2010. – 11–15. **(Рецензируемый журнал, рекомендованный ВАК.)**
- [5] Мельников Б.Ф., Мельникова А.А. *Многоаспектная минимизация недетерминированных конечных автоматов. Часть I. Вспомогательные факты и алгоритмы* // Известия вузов. Поволжский регион. Физико-математические науки. – № 4 (20). – 2011. – 59–69. **(Рецензируемый журнал, рекомендованный ВАК.)**
- [6] Мельников Б.Ф., Мельникова А.А. *Многоаспектная минимизация недетерминированных конечных автоматов. Часть II. Основные алгоритмы* // Известия вузов. Поволжский регион. Физико-математические науки. – № 1 (21). – 2012. – 31–43. **(Рецензируемый журнал, рекомендованный ВАК.)**
- [7] Мельникова А.А. *Некоторые свойства базисного автомата* // Вестник Воронежского государственного университета. Серия «Физика. Математика». – № 2. – 2012. – 184–189. **(Рецензируемый журнал, рекомендованный ВАК.)**
- [8] Мельников Б.Ф., Вахитова А.А. *Звёздная высота конечного автомата* // Учёные записки УлГУ. Фундаментальные проблемы математики и механики. – Вып.3. – Ульяновск: Изд-во УлГУ. – 1997. – 51–57.
- [9] Вахитова А.А. *Базисный конечный автомат Рабина-Скотта как инвариант регулярного языка* // Труды научн. конф. «Математическое моделирование физических, экономических, социальных систем и процессов». – Ульяновск: Изд-во УлГУ. – 1998. – 13–14.
- [10] Вахитова А.А. *Об одном алгоритме построения функции разметки конечного автомата* // Тезисы докладов Международной алгебраической конференции памяти А.Г. Куроша – М. – Изд-во МГУ. – 1998. – 152–154.
- [11] Мельникова А.А. *Свойства базисных конечных автоматов* // Труды второй научн. конф. «Математическое моделирование физических, экономических, социальных систем и процессов». – Ульяновск: Изд-во УлГУ. – 1999. – С. 11.
- [12] Мельникова А.А. *Базисные конечные автоматы Рабина-Скотта* // Материалы VII международного семинара «Дискретная математика и её приложения». Часть II – М. – Изд-во МГУ. – 2001. – 170–172.
- [13] Мельникова А.А. *Свойства базисных конечных автоматов как инварианта регулярного языка* // Тезисы докладов Межд. научно-практ. конф. «Методы и алгоритмы прикладной математики в технике, медицине и экономике» – Новочеркасск: Изд-во ЮРГТУ (НПИ). – 2001. – С. 49.

- [14] Мельников Б.Ф., Мельникова А.А. *Новый алгоритм построения базисных конечных автоматов* // Тезисы докладов XIII Межд. конф. «Проблемы теоретической кибернетики» – М. – Изд-во МГУ. – 2002. – С.124.
- [15] Мельникова А.А. *Варианты некоторых вспомогательных функций для эвристических алгоритмов* // Материалы межд. науч.-практ. конф. «Проблемы математического образования и культуры» – Тольятти: Изд-во ТГУ. – 2004. – 132–135.
- [16] Melnikov B., Vakhitova A. *Some more on the finite automata* // The Korean Journal of Computational and Applied Mathematics. – Vol. 5. – No. 3. – Korea. – 1998. – 495–506.
- [17] Vakhitova A. *The basis automaton for the given regular language* // The Korean Journal of Computational and Applied Mathematics. – Vol. 6. – No. 3. – Korea. – 1999. – 617–624.
- [18] Melnikov B., Melnikova A. *Edge-minimization of non-deterministic finite automata* // The Korean Journal of Computational and Applied Mathematics. – Vol. 8. – No. 3. – Korea. – 2001. – 469–479.
- [19] Melnikov B., Melnikova A. *Some properties of the basis finite automaton* // The Korean Journal of Computational and Applied Mathematics. – Vol. 9. – No. 1. – Korea. – 2002. – 135–150.
- [20] Мельникова А. *Пример использования базисных конечных автоматов* // Развитие и перспективы вузовской науки и образования в современных условиях. Сборник научных статей: в 2-х частях. – 1 часть. – Димитровград: ДИТИ. – 2012. – 110–114.