ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Март—апрель 2012. Том 19, № 2. С. 3–18

УДК 510.53

АЛГОРИТМИЧЕСКАЯ РАЗРЕШИМОСТЬ ЗАДАЧ О ПОВЕДЕНИИ АВТОМАТОВ НА СВЕРХСЛОВАХ *)

М. Н. Вялый, А. А. Рубцов

Аннотация. Работа посвящена двум алгоритмическим задачам, связанным с анализом поведения конечного автомата при чтении сверхслова (бесконечной последовательности): достигает ли автомат принимающего состояния и достигает ли он принимающего состояния бесконечно часто. Первая задача возникает при анализе моделей обобщённого недетерминизма, а вторая — при анализе разрешимости монадических теорий второго порядка. Получены новые условия разрешимости для этих задач. Доказано, что всякая задача регулярной реализуемости (проверки выполнимости некоторого регулярного свойства на заданном множестве слов) алгоритмически эквивалентна некоторой задаче о достижении автоматом принимающего состояния при чтении сверхслова.

Ключевые слова: сверхслово, регулярный язык, алгоритмическая разрешимость, монадическая теория.

Введение

Рассматривается поведение автоматов, читающих бесконечную последовательность (сверхслово), и алгоритмические задачи, возникающие при анализе этого поведения.

С каждым сверхсловом связана алгоритмическая задача *префиксной реализуемости*: по описанию регулярного языка проверить, существует ли префикс сверхслова, принадлежащий данному языку. Сверхслово, для которого задача префиксной реализуемости разрешима, будем называть *префиксно разрешимым*.

Если изменить формулировку и интересоваться бесконечностью пересечения регулярного языка и множества префиксов данного сверхслова, то придём к другой алгоритмической задаче, которую будем называть

*) Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проекты 09–01–00709 (первый автор) и 11–01–00398 (второй)).

задачей *Бюхи-реализуемости.* Сверхслова, для которых эта задача разрешима, будем называть *разрешимыми по Бюхи.*

В [7] обнаружена связь между задачей Бюхи-реализуемости и разрешимостью монадических теорий второго порядка. Дальнейшие результаты о разрешимости этих теорий основывались на этой связи (см., например, [4, 8]).

В нашей работе установлена сводимость задачи префиксной реализуемости к задаче Бюхи-реализуемости, тем самым разрешимые по Бюхи сверхслова являются префиксно разрешимыми.

Обратной сводимости не существует: приведён пример префиксно разрешимого сверхслова, не разрешимого по Бюхи.

Получены новые условия разрешимости для задач префиксной реализуемости и Бюхи-реализуемости, которым удовлетворяет широкий класс сверхслов. Из эквивалентности разрешимости по Бюхи и разрешимости монадических теорий второго порядка получены новые примеры разрешимых монадических теорий.

Эти условия разрешимости не требуют свойств почти периодичности, как в [3, 4], или морфичности, как в [8], или k-лексикографической автоматности, как в [6]. Вместо этого требуется, чтобы сверхслово получалось некоторой эффективной процедурой из сверхслова (быть может, в бесконечном алфавите), среди факторов которого встречаются все слова.

Обобщением задачи префиксной реализуемости является задача регулярной реализуемости, под которой понимаем проверку непустоты пересечения некоторого фиксированного языка (фильтра) и заданного регулярного языка. Задачи регулярной реализуемости возникают при анализе обобщённых моделей недетерминизма [1, 11]. В [2] найдена связь между задачами регулярной реализумости и проблемой Сколема о нулях линейных рекуррентных последовательностей.

С учётом упомянутых выше условий разрешимости для задач префиксной реализуемости доказано, что любая задача регулярной реализуемости эквивалентна задаче префиксной реализуемости для некоторого сверхслова относительно сводимостей по Тьюрингу.

1. Формулировки задач реализуемости

Пусть L — язык в конечном алфавите Σ, который будем называть *фильтром. Задача регулярной L-реализуемости* — это следующая массовая задача. Дано описание некоторого регулярного языка R в алфавите Σ и требуется проверить, что $L \cap R \neq \varnothing$. Задачу регулярной реализуемости для фильтра L будем обозначать через $\mathrm{RR}(L)$.

Если нас интересует только разрешимость задачи регулярной реализуемости, то нет разницы между различными способами задания регулярного языка — описанием детерминированного автомата, принимающего язык, недетерминированного автомата или регулярного выражения.

Всюду далее, если не оговорено противное, под автоматом понимается детерминированный конечный автомат. Будем обозначать через $\delta_A(u,q)$ функцию переходов автомата A, продолженную на множество слов во входном алфавите.

Сверхсловом в алфавите Σ назовём бесконечную последовательность $W = a_1a_2\dots a_n\dots$ символов алфавита Σ. Через $W[n,m]$ обозначим слово $a_n \dots a_m$ — *фактор* сверхслова. Факторы $W[1,n]$ будем называть *префиксами*. Для краткости записи используем также сокращение $W[n] = W[n,n] = a_n$.

Сверхслово $W = a_1a_2\dots a_n\dots$ назовём *вычислимым*, если вычислима функция $n \mapsto a_n$. В дальнейшем рассматриваем только вычислимые сверхслова, не оговаривая это в формулировках утверждений.

Множество префиксов сверхслова W обозначим через $\mathrm{Pref}(W)$. Задачу регулярной реализуемости для $\mathrm{Pref}(W)$ будем называть *задачей префиксной реализуемости* $\mathrm{R}_p(W) = \mathrm{RR}(\mathrm{Pref}(W))$ для сверхслова W и говорить, что W префиксно разрешимо, если $\mathrm{R}_p(W)$ разрешима.

Задачей Бюхи-реализуемости $\mathrm{R}_p^\infty(W)$ для сверхслова W будем называть проверку того, что $|R \cap \mathrm{Pref}(W)| = \infty$. Слово, для которого эта задача разрешима, назовём *разрешимым по Бюхи*.

Если в определениях префиксной реализуемости заменить множество префиксов множеством факторов сверхслова, то получим соответственно задачи *факторной реализуемости* $\mathrm{R}_f(W)$ (разрешимые для факторно разрешимых сверхслов) и ∞*-факторной реализуемости* $\mathrm{R}_f^\infty(W)$ (разрешимые для ∞-факторно разрешимых слов).

Между введёнными задачами имеются естественные сводимости.

Утверждение 1. $\mathrm{R}_f(W) \leqslant_{\mathrm{m}} \mathrm{R}_p(W) \leqslant_{\mathrm{m}} \mathrm{R}_p^\infty(W)$, *где* $\leqslant_{\mathrm{m}}$ *обозначает* m*-сводимость.*

Доказательство. Факторные задачи сводятся к аналогичным префиксным вариантам отображением языка R в язык $\Sigma^* R$.

Сводимость $\mathrm{R}_p(W) \leqslant_{\mathrm{m}} \mathrm{R}_p^\infty(W)$ отображает описание автомата A на входе задачи $\mathrm{R}_p(W)$ в описание автомата $\widetilde{A}$, который отличается от A лишь тем, что каждое принимающее состояние A становится поглощающим: если q — принимающее состояние, то для любого $a \in \Sigma$ выполнено $\delta_{\widetilde{A}}(a,q) = q$. Утверждение 1 доказано.

2. Бюхи- и префиксная реализуемость

Опишем связь задач Бюхи-реализуемости $\mathrm{R}_p^\infty(W)$ и разрешимости монадической теории второго порядка $\mathrm{MT}(\mathbb{N}, <, W)$, $W \in \Sigma^\infty$. Это расширение теории первого порядка натуральных чисел с отношением порядка и одноместными предикатами $a(n)$ для каждого $a \in \Sigma$. В монадической теории добавляются монадические переменные по одноместным предикатам на $\mathbb{N}$. При интерпретации формул теории $\mathrm{MT}(\mathbb{N}, <, W)$ полагаем предикат $a(n)$ истинным, если и только если $W[n] = a$ (отношение порядка интерпретируется естественным образом).

Для формулы φ теории $\mathrm{MT}(\mathbb{N}, <, W)$ через $L^\infty(\varphi)$ обозначим множество сверхслов, для которых φ истинна. Оказывается, что для любой формулы φ множество $L^\infty(\varphi)$ является регулярным ω-языком.

Под регулярным ω-языком мы понимаем множество сверхслов, принимаемых некоторым автоматом. В отличие от слов конечной длины классы регулярных ω-языков различаются для разных конструкций автоматов. Эти конструкции различаются выбором типа автомата (детерминированный или недетерминированный) и правилом, по которому автомат принимает сверхслово.

Ходом автомата A с множеством состояний Q на сверхслове W называется сверхслово в алфавите Q

$$\rho = q_0 q_1 \dots q_n \dots, \tag{1}$$

в котором q_{n+1} получается из q_n при чтении n-го символа сверхслова. Для детерминированных автоматов ход определяется по сверхслову однозначно.

Предельным множеством $\lim_\rho A$ на заданном ходе автомата A назовём множество тех состояний, которые встречаются в последовательности (1) бесконечно часто.

Автомат Бюхи принимает сверхслово W тогда и только тогда, когда существует такой ход ρ автомата на этом сверхслове, на котором одно из принимающих состояний попадает в $\lim_\rho A$.

В *автомате Мюллера* множество финальных состояний заменяется семейством финальных макросостояний $\mathcal{F} \subseteq 2^Q$, где Q — множество состояний. Автомат Мюллера принимает сверхслово W тогда и только тогда, когда существует такой ход ρ автомата на этом сверхслове, что $\lim\limits_\rho A \in \mathcal{F}$.

Через $L^\infty(A)$ будем обозначать множество сверхслов, принимаемых автоматом A, и называть это множество *регулярным ω-языком*. Оказывается [9, 10], что классы ω-языков, принимаемых недетерминированными автоматами Бюхи и автоматами Мюллера (как детерминированными, так и недетерминированными), совпадают. При этом построение эквивалентных автоматов эффективно.

Теорема [7]. *Существует алгоритм, который по недетерминированному автомату Бюхи A строит такую формулу φ монадической теории второго порядка, что $L^\infty(A) = L^\infty(\varphi)$. Существует также обратный алгоритм, который строит по формуле φ автомат Бюхи, принимающий в точности сверхслова из $L^\infty(\varphi)$.*

Меньший класс сверхслов принимают детерминированные автоматы Бюхи. Однако если нас интересует разрешимость теории $\mathrm{MT}(\mathbb{N}, <, W)$, то разницы между детерминированными автоматами Бюхи и остальными видами ω-автоматов нет.

Утверждение 2. *Проверка того, что данный недетерминированный автомат Бюхи принимает сверхслово W, сводится по Тьюрингу к проблеме $\mathrm{R}_p^\infty(W)$ Бюхи-реализуемости.*

Доказательство. По недетерминированному автомату Бюхи N построим эквивалентный ему детерминированный автомат Мюллера M. По автомату Мюллера M построим семейство детерминированных автоматов Бюхи D_F, где F принадлежит семейству финальных макросостояний автомата M. Состояниями D_F являются макросостояния (подмножества состояний) автомата Мюллера M, а переходы индуцируются переходами в автомате Мюллера: $\delta_D(a, S) = \{q : q = \delta_M(a, q'),\ q' \in S\}$. Принимающим состоянием является макросостояние F.

Автомат M принимает сверхслово W тогда и только тогда, когда какое-то макросостояние из семейства финальных макросостояний автомата M встречается на ходе автомата бесконечно часто. Но это равносильно тому, что для одного из автоматов D_F ответ в задаче Бюхи-реализуемости положительный. Утверждение 2 доказано.

Обратное утверждение очевидно, так как детерминированный автомат Бюхи — частный случай недетерминированного, а задача Бюхи-реализуемости совпадает с проверкой того, что детерминированный автомат Бюхи принимает сверхслово.

Таким образом, разрешимость по Бюхи сверхслова W равносильна разрешимости теории $\mathrm{MT}(\mathbb{N}, <, W)$. Значит, в силу утверждения 1 из разрешимости теории $\mathrm{MT}(\mathbb{N}, <, W)$ для сверхслова W следует его префиксная разрешимость.

В частности, из результатов [3, 4] вытекает префиксная разрешимость эффективно обобщённо почти периодических сверхслов, а из результатов [8] — префиксная разрешимость морфических сверхслов.

Обратной сводимости не существует.

Теорема 1. *Существует префиксно разрешимое сверхслово W, для которого задача Бюхи-реализуемости неразрешима.*

Доказательство. Для построения искомого сверхслова зафиксируем вычислимую нумерацию машин Тьюринга и вычислимую в обе стороны биекцию между множеством конечных автоматов и натуральными числами.

Сверхслово $W \in \{0,1\}^\infty$ имеет вид $w_1u_1w_2u_2\ldots w_nu_n\ldots$, где слова w_n и u_n являются конкатенациями блоков: слов вида $b_m = 10^m1$ (число m будем называть *рангом блока*).

Обозначим через T_n множество номеров k (в выбранной нумерации) машин Тьюринга таких, что $k \leqslant n$ и k-я машина Тьюринга, запущенная на пустом входе, не останавливается после n тактов работы. Из определения ясно, что множество T_n конечно. Слово w_n является конкатенацией блоков всех рангов $j \in T_n$ в порядке возрастания j. Слово u_n зависит от того, может ли автомат A_n (в выбранной нумерации автоматов) пройти через принимающее состояние при чтении продолжения слова $v_n = w_1u_1w_2u_2\ldots w_n$ какой-нибудь конкатенацией блоков, отвечающих всем машинам, кроме тех, которые уже остановились к этому времени.

Более формально, определим множество слов

$$S_n = \{b_k : (k > n) \vee (k \in T_n)\}$$

и язык S_n^*. Обозначим через q состояние автомата A_n после чтения слова v_n, а через A_n^q — автомат, отличающийся от A_n только начальным состоянием, которое у него равно q.

Если $L\big(A_n^q\big)\Sigma^* \cap S_n^* \neq \varnothing$, то в качестве u_n выбираем лексикографически наименьшее непустое слово из языка $L\big(A_n^q\big)\Sigma^* \cap S_n^*$. В противном случае полагаем $u_n = \varepsilon$ (пустое слово).

Построение сверхслова W закончено. Проверим его вычислимость. Вычислимость слов w_n не вызывает сомнений. Для проверки вычислимости u_n докажем, что язык S_n^* регулярен. В самом деле, в S_n не входит лишь конечное число блоков. Язык L, составленный из всех конкатенаций блоков и пустого слова, регулярен. Поэтому $S_n^* = L \setminus \Big(\bigcup\limits_{b \notin S_n} LbL\Big)$. Отсюда следует регулярность S_n^* в силу замкнутости класса регулярных языков относительно теоретико-множественных операций. По тем же причинам язык $L\big(A_n^q\big)\Sigma^* \cap S_n^*$ также регулярный. Построение лексикографически наименьшего слова в регулярном языке сводится к нахождению расстояния между вершинами ориентированного графа.

Докажем разрешимость задачи префиксной реализуемости $\mathrm{R}_p(W)$. Поскольку в конструкции используется вычислимая биекция между автоматами и натуральными числами, по автомату A можно определить его номер n. Из построения следует, что слова u_k, w_k при $k > n$ составлены только из блоков, принадлежащих множеству S_n (множество остановившихся машин не уменьшается с ростом числа тактов работы). Слово u_n форсирует проход через принимающее состояние при чтении таких блоков, если это вообще возможно. Поэтому если автомат A не побывал в принимающем состоянии после чтения префикса $v_n u_n$ сверхслова W, то он никогда не попадет в принимающее состояние при чтении сверхслова W.

С другой стороны, из построения следует, что блок b_n встречается в сверхслове W бесконечно часто тогда и только тогда, когда n-я машина не останавливается на пустом входе. Отсюда следует неразрешимость задачи $\mathrm{R}_f^\infty(W)$, а в силу утверждения 1 и неразрешимость задачи $\mathrm{R}_p^\infty(W)$. Теорема 1 доказана.

3. Разрешающие слова и языки

В этом разделе описан новый метод построения сверхслов, для которых разрешимы задачи префиксной и Бюхи-реализуемости. Построение основано на понятии разрешающего языка. Это понятие аналогично конгруэнтности, используемой в работах по разрешимости задач о ω-автоматах, начиная с работы Бюхи [7].

Слово $w_A \in \Sigma^*$ будем называть *разрешающим* для автомата A с входным алфавитом Σ, если при чтении w_A из любого состояния A либо пройдёт через принимающее состояние, либо попадёт в такое состояние q, из которого нет пути ни в одно из принимающих состояний A (такие состояния q будем называть *тупиковыми*).

Язык, состоящий из всех разрешающих слов для автомата A, будем называть *разрешающим языком* и обозначать через $\mathcal{D}(A)$.

Разрешающее слово форсирует ответ в задаче префиксной реализуемости в том смысле, что если w_A является фактором сверхслова W, то ответ на вопрос задачи префиксной реализуемости $\mathrm{R}_p(W)$ на входе A определяется после чтения префикса, заканчивающегося w_A. Оказывается, что для любого автомата разрешающий язык непуст и регулярен.

Утверждение 3. *Для всякого автомата A существует хотя бы одно разрешающее слово w_A.*

Доказательство. Приведём алгоритм построения некоторого разрешающего слова и докажем его корректность.

Построение. Пусть автомат A содержит n состояний, занумеруем их. Если из i-го состояния автомата существует путь в принимающее состояние q_f, то выберем такое слово u_i, что $\delta_A(u_i, q_i) = q_f$. Если же ни одного такого пути не найдётся, то q_i — тупиковое состояние и полагаем $u_i = \varepsilon$.

Для построения разрешающего слова нам понадобятся вспомогательные слова w_i, $i = 1, \ldots, n$:

$$w_1 = u_1, \quad w_i = w_{i-1}u_j, \quad i > 1, \tag{2}$$

где $q_j = \delta_A(w_{i-1}, q_i)$.

Слово w_n и является искомым разрешающим словом.

Корректность. Пусть автомат читает w_n, начиная из i-го состояния. Рассмотрим его работу на слове w_i, которое по построению является префиксом w_n. Из (2) получаем $\delta_A(w_i, q_i) = \delta_A(u_j, q_j)$. По определению u_j состояние $\delta_A(u_j, q_j)$ либо принимающее, либо тупиковое. Таким образом, слово w_n является разрешающим. Утверждение 3 доказано.

Итак, язык $\mathcal{D}(A)$ непуст. Покажем, что язык $\mathcal{D}(A)$ регулярен, а также приведём процедуру его построения.

Утверждение 4. *Разрешающий язык $\mathcal{D}(A)$ является регулярным.*

Доказательство. Пусть $A = (\Sigma, Q, \delta, q_0, F)$, где Σ — входной алфавит, Q — множество состояний, δ — функция переходов, q_0 — начальное состояние и F — множество принимающих состояний. Множество тупиковых состояний автомата A обозначим через T.

Построим автомат $A_q = (\Sigma, Q, \delta, q, F \cup T)$ для каждого состояния $q \in Q$. Пусть L_q — язык, принимаемый автоматом A_q. Тогда

$$\mathcal{D}(A) = \bigcap_{q \in Q} L_q \Sigma^*.$$

Действительно, по определению $\mathcal{D}(A)$ всякое слово $w \in \mathcal{D}(A)$ для любого q принадлежит также и $L_q\Sigma^*$, поскольку чтение некоторого префикса w_q слова w, начиная из состояния q, переводит q в множество $F \cup T$, т. е. $w_q \in L_q$. Следовательно, $\mathcal{D}(A) \subseteq \bigcap\limits_{q \in Q} L_q\Sigma^*$.

Докажем обратное включение. Пусть $w \in \bigcap_{q\in Q} L_q\Sigma^*$. Это значит, что при чтении автоматом A слова w, начиная с произвольного состояния, он либо пройдёт через принимающее состояние, либо попадёт в тупиковое. Но это и означает, что слово w разрешающее. Утверждение 4 доказано.

Сформулируем теорему о достаточном условии разрешимости задач префиксной $\mathrm{R}_p(W)$ и Бюхи-реализуемости $\mathrm{R}_p^\infty(W)$, $W \in \Sigma^\infty$.

Теорема 2. *Задача префиксной реализуемости* $\mathrm{R}_p(W)$, $W \in \Sigma^\infty$, *разрешима, если сверхслово* W *содержит в качестве фактора любое слово из* Σ^*.

Доказательство. Пусть входом задачи префиксной реализуемости является регулярный язык R, распознаваемый детерминированным автоматом A.

Решающий алгоритм очень прост: нужно моделировать работу автомата A на (вычислимом) сверхслове W до тех пор, пока не встретится принимающее или тупиковое состояние. В первом случае ответ положительный, во втором — отрицательный.

Из того, что W имеет фактор $w_A \in \mathcal{D}(A)$, следует корректность алгоритма. После чтения префикса, содержащего w_A, алгоритм остановится. Теорема 2 доказана.

Теорема 3. *Задача Бюхи-реализуемости* $\mathrm{R}_p^\infty(W)$ *разрешима, если сверхслово* W *содержит в качестве фактора любое слово из* Σ^*.

Доказательство. Пусть A — вход задачи Бюхи-реализуемости $\mathrm{R}_p^\infty(W)$. Автомат $\widetilde{A}$ отличается от A только множеством принимающих состояний, которое у него совпадает с множеством тупиковых состояний автомата A.

Докажем, что $L(\widetilde{A}) \cap \mathrm{Pref}(W) = \varnothing$ равносильно $|L(A) \cap \mathrm{Pref}(W)| = \infty$. Другими словами, докажем сводимость $\mathrm{R}_p^\infty(W) \leqslant_{\mathrm{m}} \neg\mathrm{R}_p(W)$. По теореме 2 отсюда следует разрешимость $\mathrm{R}_p^\infty(W)$.

Если $L(\widetilde{A}) \cap \mathrm{Pref}(W) \neq \varnothing$, то рано или поздно автомат A при чтении сверхслова W попадёт в тупиковое состояние. Поэтому

$$|L(A) \cap \mathrm{Pref}(W)| < \infty.$$

В случае, когда $L(\widetilde{A}) \cap \mathrm{Pref}(W) = \varnothing$, достаточно проверить, что сверхслово W содержит бесконечно много непересекающихся факторов из $\mathcal{D}(L(A))$. Тогда $|L(A) \cap \mathrm{Pref}(W)| = \infty$, поскольку при чтении каждого из этих факторов автомат A проходит либо через принимающее, либо через тупиковое состояние.

Пусть $w \in \mathcal{D}(L(A))$. По условию теоремы у сверхслова W есть фактор w. Обозначим его через $w = W[n_1, m_1] = w_1$. Далее построим бесконечную последовательность непересекающихся факторов w по следующему правилу: если $w_k = W[n_k, m_k]$, то найдём вхождение слова w^{2m_k} в сверхслово W и возьмём в качестве w_{k+1} последние $m_k - n_k + 1$ символов этого вхождения. Теорема 3 доказана.

Из утверждения 2 получаем

Следствие 1. *Теория* $\mathrm{MT}(\mathbb{N}, <, W)$ *разрешима, если сверхслово* W *содержит в качестве фактора любое слово из* Σ^*.

Приведём несколько примеров применения этих результатов.

Пример 1. Пусть $\Sigma = \{0, 1\}$, $W = 011011\ldots$ — сверхслово, состоящее из конкатенации всех двоичных слов, записанных в лексикографическом порядке. Из теоремы 2 следует, что задача $\mathrm{R}_p(W)$ разрешима, из теоремы 3 следует разрешимость по Бюхи сверхслова W, а из следствия 1 — разрешимость теории $\mathrm{MT}(\mathbb{N}, <, W)$. Другим способом разрешимость теории $\mathrm{MT}(\mathbb{N}, <, W)$ доказана в [6].

Пример 2. Действительное число α называется 2-*нормальным*, если любое двоичное слово длины n встречается в двоичной записи α с частотой 2^{-n}. В частности, двоичная запись 2-нормального слова содержит любое подслово как фактор. Поэтому применимы доказанные выше результаты.

В [5] установлена нормальность числа π в предположении некоторой гипотезы из теории динамических систем. Значит, в предположении той же гипотезы справедливы и утверждения о префиксной и Бюхи-разрешимости для двоичной записи числа π, а также разрешимость соответствующей монадической теории.

Теорема 2 накладывает сильное условие на сверхслова: любое слово должно входить как фактор. Ослабить это условие можно, рассматривая автоматы с бесконечным входным алфавитом $\Sigma_\infty = \{\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots\}$.

Оказывается, что достаточным условием разрешимости задачи префиксной реализуемости $\mathrm{R}_p(W)$, где W — сверхслово над алфавитом Σ, будет существование морфизма $\varphi\colon \Sigma_\infty^* \to \Sigma^*$, обладающего определёнными свойствами, и такого сверхслова $W_\infty \in \Sigma_\infty^\infty$, содержащего в качестве подслова всякое конечное слово из Σ_∞^*, что $W = \varphi(W_\infty)$.

Для точной формулировки этого результата нам потребуется обобщить понятие конечного автомата на случай бесконечного алфавита.

Определение 1. *Детерминированным конечным автоматом над* *бесконечным алфавитом* Σ_∞ назовём пятёрку $A_\infty = (\Sigma_\infty, Q, \delta, q_0, F)$ такую, что

(i) $\Sigma_\infty = \{\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots\}$ — бесконечный алфавит;
(ii) Q — конечное множество состояний;
(iii) $\delta\colon \Sigma_\infty \times Q \to Q$ — функция переходов;
(iv) q_0 — начальное состояние;
(v) F — множество принимающих состояний.

Как и в стандартном определении, слово $w \in \Sigma_\infty$ принимается автоматом A_∞, если после чтения w автомат A_∞ попадает в принимающее состояние. Множество слов, принимаемых автоматом A_∞, будем обозначать через $L(A_\infty)$. Язык L_∞ над бесконечным алфавитом Σ_∞ будем называть *регулярным*, если найдётся такой автомат A_∞, что $L(A_\infty) = L_\infty$.

Такое определение очевидным образом неконструктивно, в частности, функция δ может оказаться невычислимой. Будем рассматривать эффективные автоматы. Класс *эффективных автоматов* задаётся таким способом описания автоматов A_∞, что функция δ вычислима по описанию A_∞ и существует алгоритм проверки условия

$$\exists \alpha \in \Sigma_\infty : \delta(\alpha, q_1) = q_2, \quad \text{где } q_1, q_2 \in Q, \tag{3}$$

по описанию A_∞. Отметим, что определение неоднозначно — существует много различных классов эффективных автоматов.

Заметим, что проверка условия (3) позволяет вычислить отображение переходов $\Delta\colon 2^Q \to 2^Q$, определяемое как $\Delta(A) = B$, если и только если

$$B = \{q \mid \exists q_A \in A,\ \exists \alpha \in \Sigma_\infty : \delta(\alpha, q_A) = q\}.$$

Другими словами, образ $\Delta(A)$ состоит из тех состояний, в которые можно перейти из какого-нибудь состояния A чтением одного символа.

Нас интересует не столько сам язык $L(A_\infty)$, принимаемый автоматом A_∞, сколько образ этого языка при морфизме в Σ^*, где Σ — конечный алфавит. Мы предполагаем далее, что рассматриваемые морфизмы $\varphi\colon \Sigma_\infty^* \to \Sigma^*$ вычислимы. Язык $\varphi(L(A_\infty))$ будем обозначать через $L(A_\infty, \varphi)$.

Легко видеть, что сложность языка $L(A_\infty, \varphi)$ зависит не только от автомата, но и от морфизма φ.

Пример 3. Пусть $A_\infty = (\Sigma_\infty, Q, \delta, q_0, F)$ — автомат, где $Q = \{q_0, q_1\}$, $\delta(\alpha_{2k}, q_0) = q_0$, $\delta(\alpha_{2k+1}, q_0) = q_1$, $\delta(\alpha_k, q_1) = q_1$, $k \in \mathbb{N}$, $F = \{q_0\}$.

Рассмотрим морфизмы φ_1, $\varphi_2\colon \Sigma_\infty^* \to \Sigma^* = \{0,1\}^*$ такие, что

$$\varphi_1(\alpha_{2k}) = 0^k 1^k, \quad \varphi_1(\alpha_{2k+1}) = 1^k, \quad \varphi_2(\alpha_{2k}) = 0^k, \quad \varphi_2(\alpha_{2k+1}) = 1^k.$$

Тогда $L(A_\infty, \varphi_1) = \{0^k 1^k\}$ — контекстно-свободный, а $L(A_\infty, \varphi_2) = \{0^k\}$ — регулярный языки.

Заметим, что для регулярности языка $L(A_\infty, \varphi)$ достаточно, чтобы для любой пары состояний $q_i, q_j \in Q$ языки $R_{i,j} = \{\varphi(\alpha_k) \mid \delta(\alpha_k, q_i) = q_j\}$ были регулярными.

Для краткости задачу префиксной реализуемости $\mathrm{R}_p(\varphi(W_\infty))$ обозначим через $\mathrm{R}_p(W_\infty, \varphi)$.

Мы хотим найти условия на морфизм, которые гарантируют разрешимость задачи $\mathrm{R}_p(W_\infty, \varphi)$, если в W_∞ входит в качестве фактора любое слово из Σ_∞^*. С этой целью обобщим теорему 2 на случай бесконечного алфавита, а затем сведём задачу $\mathrm{R}_p(W_\infty, \varphi)$ к $\mathrm{R}_p(W_\infty)$.

Разрешающие слова для бесконечного алфавита определяются так же, как и для конечного. Легко видеть, что и для бесконечного алфавита справедливо утверждение 3 о существовании разрешающего слова, поскольку доказательство не зависит от мощности алфавита. Обобщение теоремы 2 справедливо для эффективных автоматов.

Утверждение 5. *Задача* $\mathrm{R}_p(W_\infty)$ *разрешима для эффективных автоматов, если каждое слово из* Σ_∞^* *входит в качестве фактора в* W_∞.

Доказательство. Сначала проверим для каждого состояния q существование пути в одно из принимающих состояний. Это условие равносильно тому, что орбита $\Delta^n(\{q\})$ имеет непустое пересечение с множеством принимающих состояний F, т. е. $\exists n \in \mathbb{N} : \Delta^n(\{q\}) \cap F \neq \varnothing$. Поскольку множество состояний автомата конечно, множество всех его подмножеств также конечно, поэтому орбита $\Delta^n(\{q\})$ конечна и находится алгоритмически по описанию функции переходов эффективного автомата.

Если пути из начального состояния q_0 в F нет, то даём отрицательный ответ. Если такой путь найдётся, то пометим все состояния, из которых нет пути в принимающее, как тупиковые, и будем следить за работой A_∞ на W_∞. Поскольку разрешающее слово из Σ_∞^* является фактором W_∞, рано или поздно A_∞ пройдёт через разрешающее слово и либо побывает в принимающем состоянии, либо окажется в тупиковом. Утверждение 5 доказано.

Осталось построить сводимость $\mathrm{R}_p(W_\infty, \varphi)$ к $\mathrm{R}_p(W_\infty)$. Достаточным условием существования такой сводимости оказывается разрешимость задачи регулярной реализуемости для языка L_φ, состоящего из образов букв бесконечного алфавита.

Утверждение 6. *Пусть отображение $\alpha \mapsto \varphi(\alpha)$ вычислимо и для языка $L_\varphi = \{\varphi(\alpha) \mid \alpha \in \Sigma_\infty\}$ разрешима задача регулярной реализуемости $\mathrm{RR}(L_\varphi)$. Тогда существует алгоритм, который по автомату A, задающему вход задачи префиксной реализуемости $\mathrm{R}_p(W_\infty, \varphi)$, строит такой автомат A_∞, что ответ на вопрос задачи $\mathrm{R}_p(W_\infty, \varphi)$ для языка $L(A)$ совпадает с ответом на вопрос задачи $\mathrm{R}_p(W_\infty)$ для языка $L(A_\infty)$. При этом результаты работы алгоритма образуют класс эффективных автоматов.*

Доказательство. Алгоритм по автомату $A = (\Sigma,\ Q,\ \delta,\ q_0,\ F)$ строит автомат $A_\infty = (\Sigma_\infty,\ \widetilde{Q},\ \widetilde{\delta}, \widetilde{q}_0, \widetilde{F})$ с множеством состояний $\widetilde{Q} = Q \times \{0, 1\}$. Дополнительный бит нам потребуется для того, чтобы запомнить, проходил ли автомат через принимающее состояние. Элемент $(q, b) \in \widetilde{Q}$ будем обозначать через $\widetilde{q}$, если значение дополнительного бита несущественно.

Заменив в автомате A начальное состояние на q_i, а множество принимающих состояний на $\{q_j\}$, получим автомат, принимающий регулярный язык $R_{i,j}$, который состоит из всех путей из состояния q_i в состояние q_j. Запись $R_{i,j}$ будем понимать как регулярное выражение, тем самым $R_{i,k}R_{k,j}$ — регулярное выражение для путей из q_i в q_j, проходящих через q_k.

Функцию переходов $\widetilde{\delta}$ автомата A_∞ определим следующим образом:

$$\begin{aligned} &\widetilde{\delta}(\alpha, \widetilde{q}_i) = (q_j, 1), \quad \exists q_k \in F \quad \varphi(\alpha) \in R_{i,k}R_{k,j}, \\ &\widetilde{\delta}(\alpha, \widetilde{q}_i) = (q_j, 0), \quad \varphi(\alpha) \in R_{i,j}, \quad \forall q_k \in F\, \varphi(\alpha) \notin R_{i,k}R_{k,j}. \end{aligned} \tag{4}$$

Условия (4) означают, что из состояния (q_i, a) есть переход в состояние (q_j, b) по α, если при чтении слова $\varphi(\alpha)$ автомат A переходит из состояния q_i в состояние q_j. Если при этом автомат A проходит через принимающее состояние, то вторая компонента b равна 1, иначе $b = 0$.

Определим множество принимающих состояний автомата A_∞ как множество всех состояний с битом 1, т. е. $\widetilde{F} = (q, 1) \in \widetilde{Q}$. Начальное состояние полагаем равным $(q_0, 0)$.

Автоматы A_∞, получающиеся в результате работы такого алгоритма, образуют класс эффективных автоматов. Функция переходов (4) вычислима, поскольку её значение определяется регулярными событиями. Условие (3) также проверяется эффективно, поскольку проверка перехода $\widetilde{q}_i \xrightarrow{\Sigma_\infty} \widetilde{q}_j$ сводится к проверке условий $R_{i,k}R_{k,j} \cap L_\varphi \neq \varnothing$, для которых существует алгоритм в силу разрешимости задачи регулярной реализуемости $\mathrm{RR}(L_\varphi)$.

Докажем корректность сводимости. Пусть $L(A) \cap \varphi(W_\infty) \neq \varnothing$, т. е. A проходит через принимающее состояние при чтении сверхслова $\varphi(W_\infty)$.

Тогда выберем наименьшее $m \in \mathbb{N}$ такое, что автомат A проходит через принимающее состояние на образе префикса $\varphi(W_\infty[1,m])$. Запустим автомат A_∞ на $W_\infty[1,m]$. По построению A_∞ имеем

$$\widetilde{\delta}(W_\infty[i], \widetilde{q}_j) = (\delta(\varphi(W_\infty[i]), q_j), b),$$

где b — дополнительный бит. Таким образом,

$$\widetilde{\delta}(W_\infty[1,m], \widetilde{q}_0) = (\delta(\varphi(W_\infty[1,m]), q_0), 1).$$

В обратную сторону доказательство практически такое же. Пусть после чтения префикса $W_\infty[1,m]$ автомат A_∞ оказался в принимающем состоянии, т. е. дополнительный бит равен 1. Это означает, что при чтении $\varphi(W_\infty[1,m])$ автомат A хотя бы раз прошёл через принимающее состояние. Утверждение 6 доказано.

Из утверждения 5 о разрешимости $\mathrm{R}_p(W_\infty)$ и полученной сводимости вытекает

Теорема 4. *Пусть отображение $\alpha \mapsto \varphi(\alpha)$ вычислимо и для языка $L_\varphi = \{\varphi(\alpha) \mid \alpha \in \Sigma_\infty\}$ разрешима задача регулярной реализуемости $\mathrm{RR}(L_\varphi)$. Если сверхслово W_∞ содержит в качестве фактора любое слово из Σ_∞^*, то*

(i) *задача префиксной реализуемости $\mathrm{R}_p(W_\infty, \varphi)$ разрешима;*

(ii) *задача Бюхи-реализуемости $\mathrm{R}_p^\infty(\varphi(W_\infty))$ разрешима;*

(iii) *теория $\mathrm{MT}(\mathbb{N}, <, W)$ разрешима.*

Утверждение (i) теоремы доказано выше, (ii) и (iii) доказываются аналогично теореме 3 и следствию 1, поскольку рассуждение в доказательстве теоремы 3 не зависит от мощности алфавита.

4. Эквивалентность задач регулярной и префиксной реализуемости

Применим полученные результаты для доказательства эквивалентности разрешимости задач регулярной и префиксной реализуемости. Для этого используем бесконечный алфавит, чтобы построить сверхслово W_∞, для которого будет разрешима задача префиксной реализуемости, и выберем морфизм φ так, чтобы из решения задачи $\mathrm{R}_p(W_\infty, \varphi)$ для специально подобранного регулярного языка следовало решение задачи регулярной реализуемости $\mathrm{RR}(L)$.

Используем следующие алфавиты: $\Sigma = \{0, 1\}$ — двоичный, $\Sigma_\# = \{0, 1, \#\}$ — двоичный с разделителем $\#$ и $\Sigma_\infty = \{\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_n, \ldots\}$ — бесконечный.

Пусть дан язык L над двоичным алфавитом, для которого разрешима задача регулярной реализуемости. Это означает, что L разрешим и, следовательно, перечислим. Занумеруем слова языка L и будем обозначать i-е слово через w_i. Зададим морфизм $\varphi\colon \Sigma_\infty \to \Sigma_\#$ такой, что $\varphi(\alpha_i) = w_i\#$, и обозначим $W = \varphi(W_\infty)$.

Пусть сверхслово W_∞ над алфавитом Σ_∞ таково, что любая конечная комбинация букв из Σ_∞ входит в W_∞ в качестве фактора. Тогда по теореме 4 задача $\mathrm{R}_p(W_\infty, \varphi) = \mathrm{R}_p(W)$ разрешима, если разрешима задача $\mathrm{RR}(L)$ (ясно, что разрешимость $\mathrm{RR}(L)$ равносильна разрешимости $\mathrm{RR}(L_\varphi) = \mathrm{RR}(L\#)$). Отсюда следует, что

$$\mathrm{R}_p(W) \leqslant_{\mathrm{T}} \mathrm{RR}(L),$$

где $\leqslant_{\mathrm{T}}$ означает сводимость по Тьюрингу.

Обратно, по регулярному языку R, который является входом задачи $\mathrm{RR}(L)$, построим язык $\widetilde{R} = (\Sigma^*\#)^* R\#$. Условие $\widetilde{R} \cap W \neq \varnothing$ равносильно $R \cap L \neq \varnothing$. Это означает, что

$$\mathrm{RR}(L) \leqslant_{\mathrm{m}} \mathrm{R}_p(W).$$

Авторы благодарны рецензенту, замечания которого позволили существенно улучшить качество статьи.

ЛИТЕРАТУРА

1. **Вялый М. Н.** О моделях недетерминизма для двусторонних автоматов // Тр. VIII междунар. конф. «Дискретные модели в теории управляющих систем». — М.: МаксПресс, 2009. — С. 54–60.
2. **Вялый М. Н., Тарасов С. П.** Орбиты линейных отображений и свойства регулярных языков // Дискрет. анализ и исслед. операций. — 2010. — Т. 17, № 6. — С. 20–49.
3. **Мучник Ан. А., Притыкин Ю. Л., Семенов А. Л.** Последовательности, близкие к периодическим // Успехи мат. наук. — 2009. — Т. 64, вып. 5. — С. 21–96.
4. **Семенов А. Л.** Логические теории одноместных функций на натуральном ряде // Изв. АН СССР. Сер. мат. — 1983. — Т. 47, вып. 3. — С. 623–658.
5. **Bailey D. H., Crandall R. E.** On the random character of fundamental constant expansions // Exp. Math. — 2001. — Vol. 10, N 2. — P. 175–190.
6. **Bárány V.** A hierarchy of automatic ω-words having a decidable MSO theory // Theor. Inform. Appl. — 2008. — Vol. 42. — P. 417–450.
7. **Büchi J. R.** On a decision method in restricted second-order arithmetic // Proc. Int. Congress Logic, Methodology, and Philosophy of Science. — Palo Alto: Stanford Univ. Press, 1962. — P. 1–11.

8. **Carton O., Thomas W.** The monadic theory of morphic infinite words and generalizations // Inf. Comput. — 2002. — Vol. 176, N 1. — P. 51–65.
9. **McNaughton R.** Testing and generating infinite sequences by a finite automaton // Inf. Control. — 1966. — Vol. 9. — P. 521–530.
10. **Roggenbach M.** Determinization of Büchi-automata // Automata, logics, and infinite games. — New York; Heidelberg; Berlin: Springer-Verl., 2002. — P. 43–60 (Lect. Notes Comput. Sci.; Vol. 2500).
11. **Vyalyi M. N.** On models of a nondeterministic computation // Proc. CSR 2009. — New York; Heidelberg; Berlin: Springer-Verl., 2009. — P. 334–345 (Lect. Notes Comput. Sci.; Vol. 5675).

Вялый Михаил Николаевич,
e-mail: vyalyi@gmail.com
Рубцов Александр Александрович,
e-mail: rubtsov99@gmail.com

Статья поступила
6 июня 2011 г.

Переработанный вариант —
9 сентября 2011 г.