Утверждение 3. *Пусть $F(t)$ – M_k-пороговая функция, тогда существует такое задание $F(t)$ многочленом $f(t)$, что сферы с центрами в вершинах n-мерного единичного куба и радиуса*

$$\frac{1}{5^{\frac{k}{2}}\sqrt{\left(\sum_{j=0}^{k} c_n^j\right)^{\sum_{j=0}^{k} c_n^j+1}}}$$

не пересекают повехность, задаваемую уравнением $f(t)=0$.

Автор благодарит Алешина С.В. и Зуева Ю.А. за предоставленную информацию.

Работа выполнена при поддержке гранта РФФИ №99-01-00317.

Список литературы

[1] Черников С.Н. Линейные неравенства. М.: Наука, 1968.

[2] Вапник В.Н., Червоненкис А.Я. Теория распознавания образов. М.: Наука, 1974.

[3] Алешин С.В. Распознавание динамических образов. М.: изд-во МГУ, 1996.

[4] Зуев Ю.А. Асимптотика логарифма числа пороговых функций алгебры логики // Доклады АН СССР. 1989. Т. 306. Вып. 3. С. 528-530.

[5] Ирматов А.А. О числе пороговых функций. Дискретная математика. 1993. Т. 5. №3. С. 40-43.

[6] Hastad J. On the size of weights for threshold gates // SIAM J. Discrete Math. 1994. V. 7. N. 3. P. 484-492.

О метрической сложности событий, представимых полиномиальными автоматами

А.Н. Руденко, А.С. Строгалов

В статье рассматривается асимптотическое поведение функции роста $G(N)$ полиномиального автомата-акцептора, представляющей для заданного автомата число слов длины не более N, распознаваемых данным автоматом. Для этой функции находится простейший асимптотически эквивалентный ей полином (моном). Описывается класс таких простейших полиномов, в точности состоящий из тех полиномов, каждый из которых является асимптотически эквивалентным некоторой функции роста полиномиального автомата.

Пусть $A = \{a_1, a_2, ..., a_n\}$ – конечный алфавит, и A^* – множество всех слов (конечных последовательностей) над A; подмножество $L \subseteq A^*$ называется языком L над алфавитом A. Мы рассматриваем автоматы-акцепторы (распознаватели), понимаемые как пятерка $M = \langle A, Q, \varphi, q_0, Q_F \rangle$; где A, Q – конечные алфавиты, входной и состояний соответственно, $Q_F \subseteq Q$ называется множеством финальных состояний, q_0 – начальное состояние, а $\varphi : A \times Q \to Q$ – функция переходов автомата. Все неопределяемые далее в тексте понятия взяты из [1].

Язык $L_M \subseteq A^*$ распознается автоматом M, если при подаче любого слова $\widetilde{\ell} \in L_M$ на вход автомата он переходит из состояния q_0 в одно из состояний Q_F. Иными словами, $\varphi(q_0, \widetilde{\ell}) \in Q_F$, где $\varphi(q_0, \widetilde{\ell})$ – стандартное распространение функции $\varphi : Q \times A \to Q$ на множество слов $\widetilde{\ell} \in A^*$ [1]. Представим это множество слов в виде $L_M = \bigcup_i L_M(i)$, где $L_M(i)$ – множество слов длины i в языке L_M.

Определение 1. Функцией роста $G_M(N)$ автомата M называется величина $G_M(N) = \sum\limits_{i \leqslant N} |L_M(i)|$. Очевидно, что $G_m : \mathbb{N}^+ \to \mathbb{N}^+$ есть целозначная функция натурального ряда.

Содержательный смысл $G_M(N)$ очевиден – это число слов длины не большей N, распознаваемых автоматом M. В дальнейшем, когда будет ясно, о каком автомате идет речь, мы будем опускать нижний индекс и будем писать $G(N)$.

Определение 2. Функция $F(N)$ есть асимптотическая функция для автомата M (для краткости, асимптотика автомата M), если $F(N) \sim G(N)$, где $G(N)$ есть функция роста автомата M. (Здесь $F(N) \sim G(N)$ означает что $F(N) = G(N)(1 + o(1))$).

В работе [2] показано, что автоматы-распознаватели разбиваются на два класса – автоматы, распознающие язык экспоненциального роста, то есть $G_M(N)$ растет не медленнее экспоненты вида $F_M(N) = k \cdot \alpha^N; k, \alpha \in \mathbb{Q}, \alpha > 1$, и распознающие язык полиномиального роста, то есть $G_M(N)$ растет не быстрее полинома вида $F_M(N) = \alpha N^n; \alpha \in \mathbb{Q}^+$, а n – натуральное число.

Не ограничивая общности, мы будем рассматривать языки над алфавитом $\{0,1\}$. В случае алфавита A мощности большей двух мы можем закодировать каждый символ алфавита A двоичной цепочкой некоторой длины k из нулей и единиц. При этом слова языка L_1 над алфавитом A очевидным образом будут взаимно однозначно соответствовать некоторому множеству цепочек, состоящих из нулей и единиц, то есть некоторому языку L_2 над алфавитом $\{0,1\}$, очевидно регулярному, и число слов длины не большей N языка L_1 над алфавитом A будет равно числу слов длины не большей kN языка L_2 над алфавитом $\{0,1\}$. Тогда, если $G_1(N)$ и $G_2(N)$ – функции роста языков L_1 и L_2 соответственно, $F_1(N)$ и $F_2(N)$ – соответствующие асимптотики, то

$$F_1(N) \sim G_1(N) = G_2(k \cdot N) \sim F_2(k \cdot N)$$

и нахождение асимптотики языка L_1 над алфавитом A сводится к нахождению асимптотики языка L_2 над алфавитом $\{0,1\}$.

Определение 3. Последовательность состояний диаграммы автомата

$$q_{i_0} \to q_{i_1} \to \ldots \to q_{i_{k-1}} \to q_{i_0},$$

где $q_{i_t} \neq q_{i_s}, t \neq s,\ t, s \in \{0, 1, \ldots, k-1\}$, называется циклом длины K $(K \geqslant 1)$.

В работе [2] доказано, что критерием полиномиальности автомата является отсутствие в диаграмме приведенного (минимального) автомата двух различных пересекающихся циклов, то есть циклов, имеющих хотя бы одно общее состояние диаграммы, и существует путь, ведущий из хотя бы одного состояния этих циклов в множество финальных состояний. Заметим, что из этого следует наличие в полиномиальном автомате поглощающего нефинального состояния.

В работе [3] показано, что любой неотрицательный целозначный полином является функцией роста некоторого автомата. Однако нахождение коэффициентов данного полинома для заданного автомата связано, как правило, с большими вычислительными трудностями. Мы докажем, что можно пренебречь степенями полинома, кроме старшей, с целью построения автомата, построение которого не вызывает особых трудностей, и функция роста которого позволяет описать ее поведение на бесконечности и тем самым дает количественный критерий для сравнения поведения двух полиномиальных автоматов.

Далее будем изображать фрагменты диаграммы автомата, опуская изображения поглощающего состояния и ребер, ведущих в него, так как они при необходимости всегда легко восстанавливаются. Будем говорить далее о диаграмме автомата, имея в виду сказанное выше («фрагмент диаграммы»).

Определение 4. Любой автомат с единственным финальным состоянием $q_n = q_F$, диаграмма которого имеет вид, изображеный на рис. 1, называется элементарным автоматом.

Название «элементарный» выбрано в связи с тем, что диаграмма любого полиномиального автомата содержит поддиаграмму вида, изображенного на рис. 1, и любой полиномиальный автомат эквивалентен, в смысле обычной эквивалентности автоматов (см. [1]),

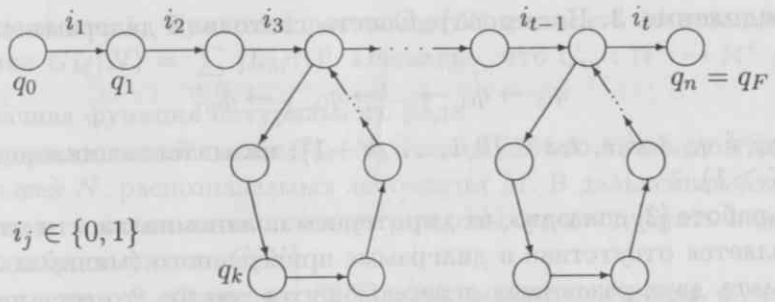

Рис. 1. Элементарный автомат.

объединению некоторых элементарных автоматов.

Ввиду того, что целью настоящей статьи является нахождение асимптотического поведения функций роста автоматов, мы можем заменить рассматриваемый автомат эквивалентным ему в асимптотическом смысле автоматом, но имеющим более удобную для нахождения асимптотики структуру диаграммы.

Рассмотрим важный для дальнейших рассуждений пример асимптотически эквивалентного преобразования автомата, диаграмма которого изображенна на рис. 1. Заменим его асимптотически эквивалентным автоматом B следующим образом: финальное состояние заменим цепочкой последовательных состояний длины $l+1$, а новым финальным состоянием сделаем последнее состояние этой цепочки. Процедура замены изображена на рис. 2.

В силу эквивалентности $aN^n \sim a(N \pm l)^n$, асимптотика преобразованного автомата останется эквивалентной асимптотике исходного автомата. (Здесь aN^n и $a(N \pm l)^n$ есть соответственно асимптотики исходного и преобразованного автоматов.)

Прежде чем переходить к нахождению асимптотики функций роста элементарных автоматов, докажем следующую лемму.

Лемма 1 (Руденко А.Н.). *Пусть* $\{a_i\}$ *и* $\{b_i\}$ – *асимптотически эквивалентные последовательности* $(a_i, b_i > 0, i \in \mathbb{N}\backslash\{0\})$. *Если*

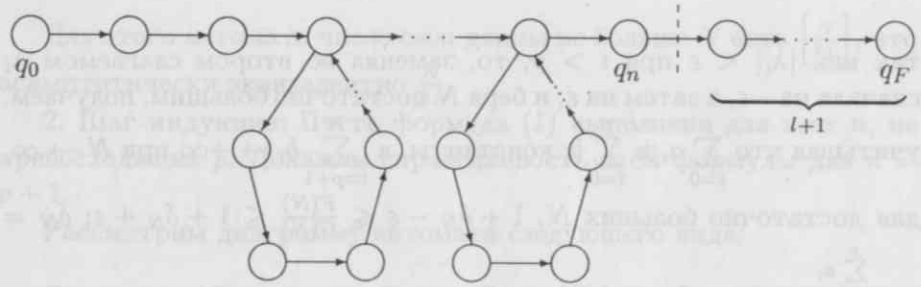

Рис. 2. Автомат B.

$\sum\limits_{i=0}^{N} a_i \to +\infty$ *при* $N \to \infty$, *то*

$$F(N) = \sum_{i=0}^{N} a_i \sim G(N) = \sum_{i=0}^{N} b_i \tag{1}$$

Доказательство. Так как последовательности $\{a_i\}$ и $\{b_i\}$ асимптотически эквивалентны, то $a_i = b_i(1+\lambda_i)$, где $\lambda_i \to 0, i \to +\infty$ (а следовательно и $\sum\limits_{i=0}^{N} b_i \to +\infty$, так как $\sum a_i = \sum(b_i + \lambda_i b_i)$, где $\lambda_i \to 0$), отсюда $a_i \sim b_i$ равносильно $\frac{a_i}{b_i} \to 1$ при $i \to +\infty$.

Аналогично $F(N) \sim G(N)$ равносильно тому, что

$$\frac{F(N)}{G(N)} \to 1, \text{ при } N \to +\infty. \tag{2}$$

Поэтому вместо (1) достаточно доказать (2).

Возьмем произвольное малое положительное число $\varepsilon > 0$ и $p \in \mathbb{N}\backslash\{0\}$ такое, что $|\lambda_i| < \varepsilon$ при $i > p$ (λ_i из равенства $a_i = b_i(1+\lambda_i)$). Тогда:

$$\frac{F(N)}{G(N)} = \frac{\sum\limits_{i=0}^{N} a_i}{\sum\limits_{i=0}^{N} b_i} = \frac{\sum\limits_{i=0}^{p} a_i + \sum\limits_{i=p+1}^{N} a_i}{\sum\limits_{i=0}^{p} b_i + \sum\limits_{i=p+1}^{N} b_i} = \frac{\sum\limits_{i=0}^{p} a_i}{\sum\limits_{i=0}^{p} b_i + \sum\limits_{i=p+1}^{N} b_i} + \frac{\sum\limits_{i=p+1}^{N} b_i + \sum\limits_{i=p+1}^{N} b_i \cdot \lambda_i}{\sum\limits_{i=p+1}^{N} b_i + \sum\limits_{i=0}^{p} b_i}$$

так как $|\lambda_i| < \varepsilon$ при $i > p$, то, заменяя во втором слагаемом λ_i сначала на $-\varepsilon$, а затем на ε, и беря N достаточно большим, получаем, учитывая что $\sum\limits_{i=0}^{p} a_i$ и $\sum\limits_{i=0}^{p} b_i$ константы, а $\sum\limits_{i=p+1}^{N} b_i \to +\infty$ при $N \to \infty$, для достаточно больших N, $1 + \delta_N - \varepsilon \leqslant \frac{F(N)}{G(N)} \leqslant 1 + \delta_N + \varepsilon$; $\delta_N = \frac{\sum\limits_{i=0}^{p} a_i}{\sum\limits_{i=0}^{p} b_i + \sum\limits_{i=p+1}^{N} b_i}$, где $\delta_N \to 0, N \to +\infty$.

В силу произвольности ε мы получаем, что $\frac{F(N)}{G(N)} \to 1$.
Лемма 1 доказана.

Докажем, что автомат с диаграммой на рис. 1, имеет асимптотику

$$F(N) = \frac{N^n}{k_1 \cdot k_2 \cdot \ldots \cdot k_n \cdot n!}, \qquad (3)$$

где n – число циклов, а k_i – длина i-го цикла. Доказательство ведется индукцией по числу циклов n диаграммы состояний.

1. База индукции: $n = 1$. Рассмотрим автомат, диаграмма которого изображена на рис. 3а.

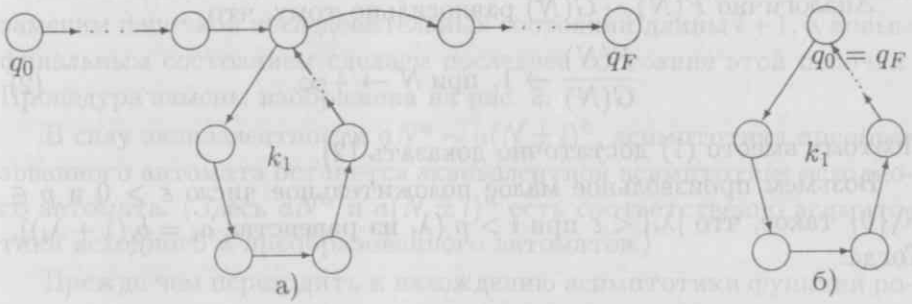

Рис. 3. База индукции.

Выше было показано, что такой автомат асимптотически эквивалентен автомату с диаграммой на рис. 3б.

Для этого автомата число слов длины не больше N есть $\left[\frac{N}{k_1}\right]$, что асимптотически эквивалентно $\frac{N}{k_1}$.

2. Шаг индукции: Пусть формула (1) выполнена для всех n, не превосходящих p. Докажем справедливость этой формулы для $n = p + 1$.

Рассмотрим диаграмму автомата следующего вида:

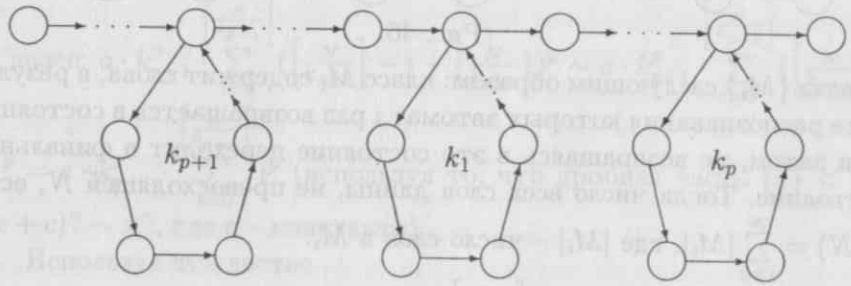

Рис. 4.

Занумеруем циклы так, как это показано на диаграмме (рис. 4). Здесь $k_1, \ldots, k_p, k_{p+1}$ – длины соответствующих циклов.

Автомат с этой диаграммой мы можем заменить асимптотически эквивалентным ему автоматом (рис. 4а), и дальнейшие рассуждения продолжить для него.

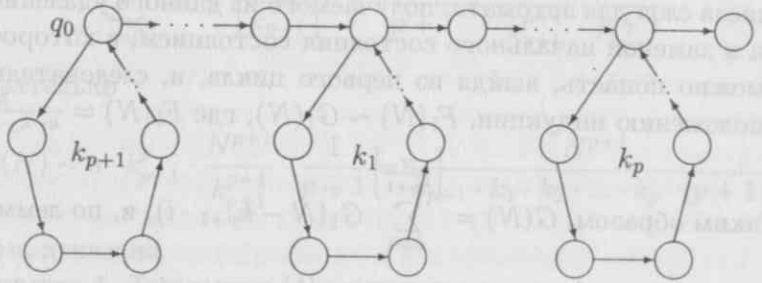

Рис. 4а.

Все слова длины меньшей либо равной N, распознаваемые автоматом, изображенным на рис. 4а, разобьем на непересекающиеся

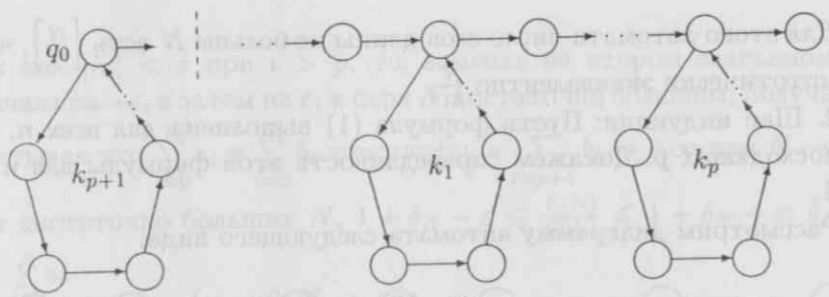

Рис. 4б.

классы $\{M_i\}$ следующим образом: класс M_i содержит слова, в результате распознавания которых автомат i раз возвращается в состояние q_0 и затем, не возвращаясь в это состояние переходит в финальное состояние. Тогда число всех слов длины, не превосходящей N, есть $G(N) = \sum\limits_{i=0}^{\infty} |M_i|$, где $|M_i|$ – число слов в M_i.

Очевидно, что при $i > \left[\frac{N}{k_{p+1}}\right]$ имеем $|M_i| = 0$, поэтому

$$G(N) = \sum_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} |M_i|.$$

Из определения M_i следует, что любые два слова из M_i имеют одинаковое начало длины $k_{p+1} \cdot i$, а так как эти слова длины не большей N, то они могут отличаться лишь остатками слов длины не более чем $N - k_{p+1} \cdot i$, то есть таких слов $G_1(N - k_{p+1} \cdot i)$, где G_1 есть функция числа слов для автомата, получаемого из данного удалением 1-го цикла и заменой начального состояния состоянием, в которое впервые можно попасть, выйдя из первого цикла, и, следовательно, по предположению индукции, $F_1(N) \sim G_1(N)$, где $F_1(N) = \frac{N^p}{k_1 \cdot k_2 \cdot \ldots \cdot k_p \cdot p!}$

Таким образом, $G(N) = \sum\limits_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} G_1(N - k_{p+1} \cdot i)$, и, по лемме 1

$$G(N) = \sum_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} G_1(N - k_{p+1} \cdot i) \sim \sum_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} F_1(N - k_{p+1} \cdot i) = \sum_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} a \cdot (N - k_{p+1} \cdot i)^p,$$

где $a = \frac{1}{k_1 \cdot k_2 \cdot \ldots \cdot k_p \cdot p!}$.

Но $\sum\limits_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} a \cdot (N - k_{p+1} \cdot i)^p = a \cdot k_{p+1}^p \cdot \sum\limits_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} (\frac{N}{k_{p+1}} - i)^p =$

$= a \cdot k_{p+1}^p \cdot \sum\limits_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} (\left[\frac{N}{k_{p+1}}\right] - i + \{\frac{N}{k_{p+1}}\})^p,$

и далее: $a \cdot k_{p+1}^p \cdot \sum\limits_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} (\left[\frac{N}{k_{p+1}}\right] - i + \{\frac{N}{k_{p+1}}\})^p \sim a \cdot k_{p+1}^p \cdot \sum\limits_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} (\left[\frac{N}{k_{p+1}}\right] - i)^p = a \cdot k_{p+1}^p \cdot \sum\limits_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} i^p$ (используя то, что дробная часть $\{x\} < 1$, и $(x + c)^n \sim x^n$, где c – константа).

Используя тождество

$$(n+1) \sum_{i=0}^{N} i^n = N^{n+1}(1 + o(1)),\ N \to +\infty, \tag{4}$$

несложно доказываемое по индукции (см. ниже замечание 1), получаем, что

$$\sum_{i=0}^{\left[\frac{N}{k_{p+1}}\right]} i^p \sim \left[\frac{N}{k_{p+1}}\right]^{p+1} \cdot \frac{1}{p+1} \sim \frac{N^{p+1}}{k_{p+1}^{p+1}} \cdot \frac{1}{p+1}.$$

Следовательно

$$G(N) \sim a \cdot k_{p+1}^p \cdot \frac{N^{p+1}}{k_{p+1}^{p+1}} \cdot \frac{1}{p+1} = \frac{N^{p+1}}{k_{p+1} \cdot k_1 \cdot k_2 \cdot \ldots \cdot k_p \cdot (p+1)!}$$

Теорема доказана.

Замечание 1. Тождество (4) можно получить по индукции следующим образом:

(4) верно для $n = 1$; пусть верно для $(n-1)$. Тогда из тождества

$\sum_{i=0}^{N}(i+1)^{n+1} = \sum_{i=0}^{N} i^{n+1} + (N+1)^{n+1}$, раскрывая скобки, получаем

$$\sum_{i=0}^{N} i^{n+1} + \sum_{i=0}^{N}(n+1)\cdot i^{n} + ... = \sum_{i=0}^{N} i^{n+1} + N^{n+1} + ...,$$

где слева многоточием, в силу предположения индукции, обозначен член $o(N^{n+1})$, а справа также присутствует член $o(N^{n+1})$ в силу разложения бинома. Сокращая первые слагаемые слева и справа, получаем (4) для n. Следовательно, по индукции равенство (4) доказано, а следовательно, доказана и теорема.

Рассмотрим произвольный инициальный полиномиальный автомат (то есть инициальный автомат-акцептор с полиномиальной функцией роста) и рассмотрим два цикла A и B в диаграмме этого автомата. Будем говорить, что цикл A предшествует циклу B, если из некоторого состояния цикла A можно перейти в состояние цикла B. Очевидно, таким образом мы вводим на множестве циклов полиномиального автомата частичный порядок по отношению достижимости состояний.

Примем следующее соглашение: будем считать состояния, в которые можно попасть, выйдя из последних (в смысле описанного выше упорядочения) циклов, циклами нулевой длины, и включим их в ранее полученное частично упорядоченное множество. Очевидно, оно по прежнему сохранит свойство частичной упорядоченности и останется конечным. Назовем цепью в автомате множество его циклов (включая нулевые), являющееся линейно упорядоченным, в смысле введенного выше упорядочения, и такое, что для любого цикла, кроме последнего, один из непосредственно следующих за ним также принадлежит этому множеству. Число циклов в цепи называется ее длиной. Длину максимальной цепи назовем длиной автомата. Выполним следующую процедуру преобразования автомата: рассмотрим последние циклы (включая нулевые) в автомате, очевидно, их конечное число. Для каждого из них возможны два случая:

1. Существует более одного входа в данный цикл, тогда проведем следующее преобразование автомата. Для каждого из входов в данный цикл построим экземпляр автомата, получаемый следующим образом.

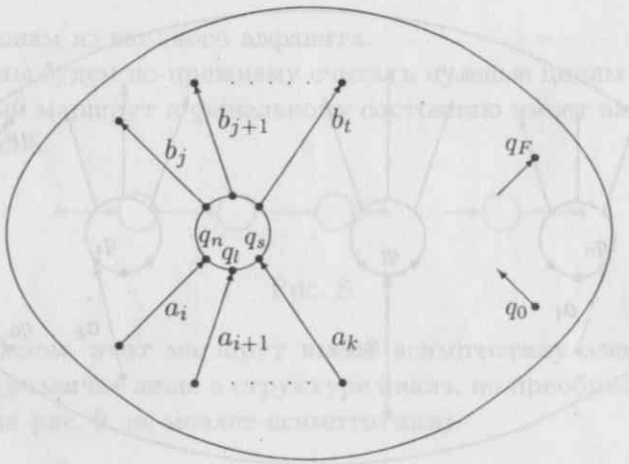

Рис. 5.

Множество его состояний есть множество состояний исходного автомата, следующих за состоянием, являющимся точкой входа, (включая и его само). Начальное состояние есть состояние входа в данный цикл, ребра индуцируются соответствующими ребрами исходного автомата.

Далее, на каждый из этих входов «приклеим» соответственно полученный автомат. Графическое изображение данной процедуры, легко формализуемой в терминах теории графов, поясняется рис. 5,6.

2. Вход единственный: автомат не изменяем.

Очевидно, что, проделав данную процедуру для каждого из последних состояний, мы получим эквивалентный автомат той же длины, у которого последние циклы имеют в точности один вход. Проделаем ту же процедуру для циклов, непосредственно предшествующих циклам, рассматриваемым на предыдущем шаге (то есть последним циклам). Очевидно, что проделав данную процедуру некоторое ко-

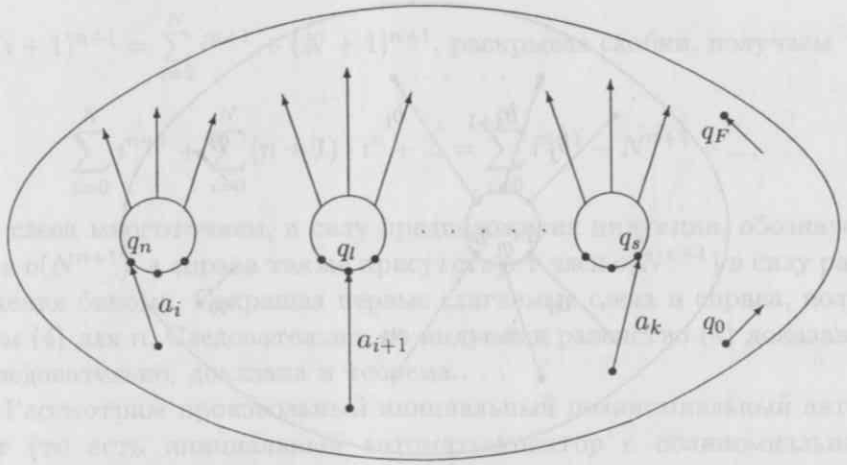

Рис. 6.

нечное число раз мы преобразуем автомат в эквивалентный ему и имеющий древовидную структуру, изображенную на рис. 7.

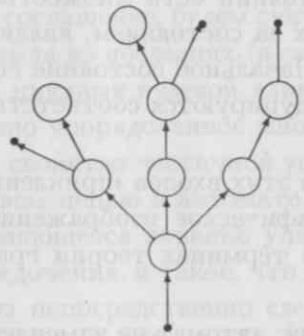

Рис. 7.

Из построения ясно, что автомат, изображенный на рис. 7, распознает тот же язык, что и исходный, и что любые два маршрута в преобразованном автомате, начинающиеся в инициальном состоянии и заканчивающиеся в различных финальных, соответствуют

разным словам из входного алфавита.

Далее мы будем по-прежнему считать нулевые циклы состояниями. Каждый маршрут к финальному состоянию имеет вид, показанный на рис. 8.

Рис. 8.

Таким образом, этот маршрут имеет асимптотику элементарного автомата (различие лишь в структуре цикла, но преобразование, показанное на рис. 9, не меняет асимптотики).

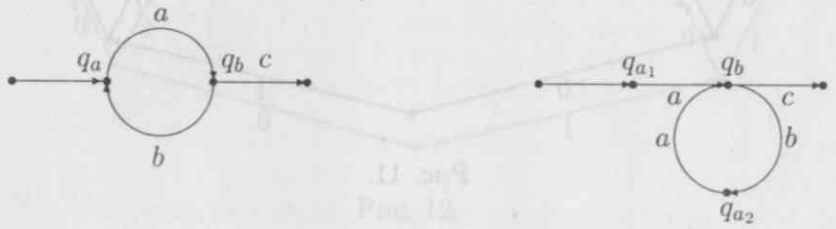

Рис. 9.

Рассматривая только маршруты с максимальным числом встречающихся циклов (так как маршруты с меньшим числом циклов не меняют асимптотики) и суммируя соответствующие им асимптотики, получаем асимптотику всего автомата. Из процесса вычислений, приведенного выше, ясно, что она имеет вид $a \cdot N^n$, где $a \in \mathbb{Q}^+$, $n = 0, 1, \ldots$.

Докажем, что для любых $p, q \in \mathbb{N}^+$, для любого $n \in \mathbb{N}$, существует автомат с асимптотикой $\frac{p}{q} \cdot N^n$.

Действительно, рассмотрим автомат, диаграмма которого изображена на рис. 10, и бинарное дерево глубины m, где $2^m \geqslant p \cdot n!$ (рис. 11).

К первым $pn!$ вершинам приклеим $pn!$ экземпляров элементарных автоматов, изображенных на рис. 10, а остальные вершины объявим финальными.

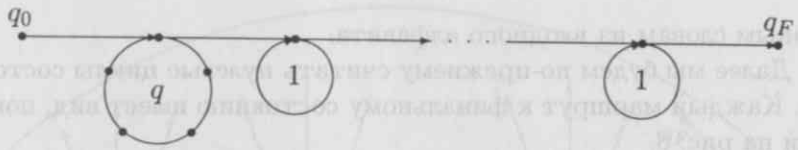

Рис. 10.

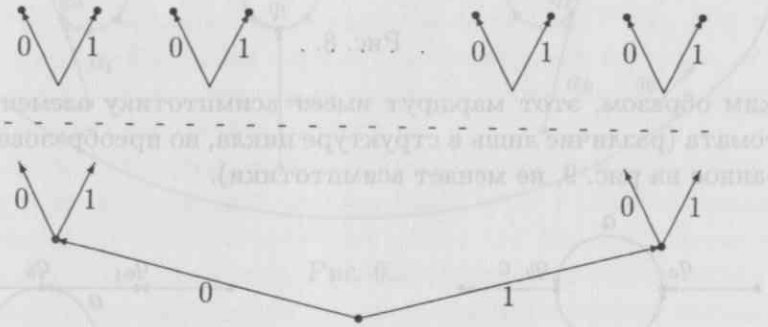

Рис. 11.

Дерево приобретет вид, изображенный на рис. 12. Очевидно, его асимптотика есть:
$p \cdot n! \cdot \frac{N^n}{p \cdot n!} = \frac{p}{q} \cdot N^n$.

Данная асимптотика автомата учитывает только главный член (то есть старшую степень). Поэтому при малых значениях N она будет давать отличие от действительного числа распознаваемых слов. Можно получить, пользуясь вышеописанным методом, асимптотику, учитывающую 2-ой, и т.д. члены. Однако при этом значительно возрастает сложность вывода асимптотических формул, и фактически мы возвращаемся к конструкциям, предложенным в работе [3]. Учитывая это, при небольших N можно использовать достаточно простой способ, использующий стандартный аппарат производящих функций (используя операции типа свертки), позволяющий точно вычислять число слов длины N (а следовательно и длины не более N) для заданных автоматов и их последовательного соединения (то есть объединения финального состояния одного автомата с

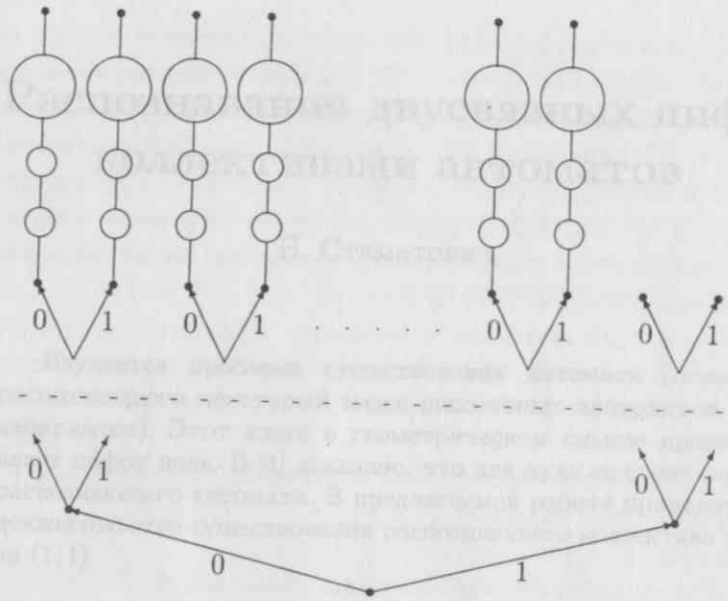

Рис. 12.

начальным состоянием другого).

Список литературы

[1] Кудрявцев В.Б., Алешин С.В., Подколзин А.С. Введение в теорию конечных автоматов. М.: Наука, 1985.

[2] Строгалов А.С. Об ε-моделировании конечных автоматов // Труды Всесоюзного семинара по дискретной математике. М.: Изд-во МГУ, 1986.

[3] Строгалов А.С. О регулярных языках с полиномиальным числом слов // Дискретная математика. 1990. Т. 2. Вып. 3.