Прогнозирование периодических сверхсобытий автоматами

Э. Э. Гасанов

В статье обобщаются результаты работы [1] на случай k-значных логик. Вводится понятие прогнозирующего автомата, который для каждого поданного ему на вход сверхслова из заданного множества, начиная с некоторого шага в каждый момент t выдает значение входного слова в момент $t+1$, то есть предугадывает входное сверхслово. Получен критерий прогнозируемости множеств сверхслов. Приведен наилучший по порядку метод построения прогнозирующего автомата для произвольного прогнозируемого множества сверхслов.

Ключевые слова: конечный автомат, прогнозирующий автомат, прогнозирование событий.

1. Введение

Процесс прогнозирования удобно рассматривать как альтернативную игру двух игроков. Имеется некоторое сверхсобытие L в алфавите $E_k = \{0, 1, \ldots, k-1\}$, и оба игрока его знают. Игра идет во времени, и каждый такт каждый игрок делает по одному ходу, причем в каждый момент времени первым ход делает первый игрок, а затем ходит второй игрок. Игра, вообще говоря, идет бесконечно долго. В начальный (первый) момент времени первый игрок, представляющий собой прогнозирующее устройство, называет некоторый символ $a(1)$ из E_k, стараясь предугадать символ, который будет назван вторым игроком. Затем ход делает второй игрок и выдает свой символ $b(1)$ из алфавита E_k так, чтобы хотя бы одно сверхслово из сверхсобытия L начиналось на символ $b(1)$, при этом второй игрок уже знает ход первого игрока. Пусть прошло t тактов времени, и за это время первый игрок сделал t ходов, которые образуют слово $\alpha_t = a(1)\ldots a(t)$, и второй игрок сделал t ходов, которые образуют слово $\beta_t = b(1)\dots b(t)$. В $(t+1)$-й момент времени первый игрок выбирает символ $a(t+1)$. После этого второй игрок выбирает символ $b(t+1)$ так, чтобы в сверхсобытии L нашлось сверхслово, начинающееся на слово β_t. Если оказывается, что $b(t+1) = a(t+1)$, то говорят, что произошло угадывание. Пусть за t тактов произошло d_t угадываний. Задача первого игрока увеличивать число угадываний d_t, а задача второго игрока уменьшать это значение. Если, начиная с некоторого момента времени первый игрок всегда правильно предугадывает ход второго игрока, то считается, что первый игрок победил, и прогнозирующее устройство построено, в противном случае побеждает второй игрок, который доказывает, что прогнозирующее устройство для сверхсобытия L построить нельзя.

В качестве прогнозирующего устройства у нас будет выступать конечный автомат. Будем говорить, что автомат прогнозирует сверхсобытие L, если для каждого входного сверхслова из этого множества автомат с некоторого момента времени начинает на своем выходе предугадывать значение входного слова в следующий момент. Множество сверхслов прогнозируемо, если существует прогнозирующий его автомат. Показано, что прогнозируемы только множества, состоящие из периодических сверхслов. Под сложностью прогнозирующего автомата понимается число его состояний. Для произвольного прогнозируемого множества получен порядок сложности прогнозирующего его автомата.

2. Основные понятия и формулировка результатов

Пусть $E_k = \{0, 1, \dots, k-1\}$ — конечный алфавит. Через E_k^* и E_k^∞ обозначим соответственно множество всех слов конечной длины и множество всех сверхслов в алфавите E_k.

Если a — сверхслово в алфавите E_k, t — натуральное число, то через $a\underset{t}{]}$ обозначим префикс длины t сверхслова a, то есть $a\underset{t}{]} = a(1)a(2)\dots a(t)$.

В соответствии с [2] будем рассматривать конечные автоматы следующего вида:

$$V = (E_k, Q, E_k, \varphi, \psi)\,,$$

где E_k — входной алфавит, Q — множество состояний, которое является конечным подмножеством некоторого фиксированного счётного множества, E_k — выходной алфавит, $\varphi : Q \times E_k \to Q$ — функция переходов, $\psi : Q \times E_k \to E_k$ — функция выходов. $V(q_0)$ — инициальный автомат с начальным состоянием q_0, $q_0 \in Q$.

Расширим функции φ и ψ на $Q \times E_k^*$, а именно, если $\alpha \in E_k^*$, $a \in E_k$, то индуктивно определим

$$\begin{aligned} \varphi(q, \alpha a) &= \varphi(\varphi(q, \alpha), a), \\ \psi(q, \alpha a) &= \psi(\varphi(q, \alpha), a). \end{aligned}$$

Через $\omega\,(V)$ обозначим количество состояний автомата V, то есть $\omega\,(V) = |Q|$.

Будем говорить, что инициальный автомат $V(q_0)$ *прогнозирует множество сверхслов* $\mathcal{A}$, $\mathcal{A} \subset E_k^\infty$, если для любого сверхслова a из $\mathcal{A}$ существует такое натуральное число N, что для любого $t \geqslant N$ выполняется $\varphi(\underset{t}{a\,]}, q_0) = a(t+1)$.

Через $\mathbb{G}\,(\mathcal{A})$ обозначим множество всех конечных инициальных автоматов, прогнозирующих множество сверхслов $\mathcal{A}$, $\mathcal{A} \subset E_k^\infty$.

Множество сверхслов $\mathcal{A}$, $\mathcal{A} \subset E_k^\infty$, называется *прогнозируемым*, если существует прогнозирующий его автомат.

Пусть $\omega\,(\mathcal{A}) = \min\limits_{V \in \mathbb{G}(\mathcal{A})} \omega\,(V)$ — минимальное количество состояний, достаточное автомату, чтобы прогнозировать множество $\mathcal{A}$, $\mathcal{A} \subset E_k^\infty$. Если $\mathbb{G}\,(\mathcal{A}) = \varnothing$, то полагаем, что $\omega\,(\mathcal{A}) = \infty$.

Если $\mathcal{A} \subset E_k^\infty$, то обозначим $\mathcal{A}^* = E_k^* \mathcal{A} = \{ba : b \in E_k^*, a \in \mathcal{A}\}$ — множество, полученное прибавлением к $\mathcal{A}$ всевозможных начальных отрезков конечной длины.

Если a — слово из E_k^*, то через $\langle a \rangle$ обозначим множество, состоящее из всех циклических сдвигов слова a. Понятно, что $|\,\langle a \rangle\,| \leqslant |a|$, причем $|\,\langle a \rangle\,| = |a|$, только если a — непериодическое.

Положим $\langle A \rangle = \bigcup\limits_{a \in A} \langle a \rangle$, где $A \subset E_k^*$.

Для сверхслова a из E_k^∞ через $\mathrm{T}(a), \mathrm{T}'(a)$ обозначим соответственно периодическую и предпериодическую части сверхслова a минимальной длины, такие что $a = \mathrm{T}'(a)(\mathrm{T}(a))^\infty$. Положим $\tau(a) = |\mathrm{T}(a)|$ и $\tau'(a) = |\mathrm{T}'(a)|$, то есть $\tau(a), \tau'(a)$ — период и предпериод минимальной длины сверхслова a.

Если $\mathcal{A}$ — некоторое множество периодических сверхслов, γ — символ из E_k, то положим $\mathrm{T}(\mathcal{A}) = \{\mathrm{T}(a) : a \in \mathcal{A}\}$, $\mathrm{T}'(\mathcal{A}) = \{\mathrm{T}'(a) : a \in \mathcal{A}\}$, $\|\mathcal{A}\|_\gamma = |\{a \in \langle \mathrm{T}(\mathcal{A})\rangle : a(1) = \gamma\}|$, $\|\mathcal{A}\| = \max\limits_{\gamma \in E_k} \|\mathcal{A}\|_\gamma$. Иными словами $\|\mathcal{A}\|_\gamma$ — суммарное количество вхождений символа γ в периодические части слов из $\mathcal{A}$.

Справедливы следующие теоремы.

Теорема 1. *Множество сверхслов $\mathcal{A}$, $\mathcal{A} \subset E_k^\infty$, прогнозируемо тогда и только тогда, когда оно состоит из периодических сверхслов с ограниченным периодом.*

Теорема 2. *Если $\mathcal{A}$, $\mathcal{A} \subset E_k^\infty$, — прогнозируемо, то*

$$|\langle \mathrm{T}(\mathcal{A})\rangle| / k \leqslant \|\mathcal{A}\| \leqslant \omega(\mathcal{A}^*) \leqslant |\langle \mathrm{T}(\mathcal{A})\rangle| .$$

3. Нижние оценки сложности прогнозирующего автомата

Оценим снизу количество состояний прогнозирующего автомата.

Лемма 1 (Нижняя оценка). *Если $\mathcal{A}$, $\mathcal{A} \subset E_k^\infty$, — некоторое множество периодических сверхслов, то*

$$\omega(\mathcal{A}) \geqslant |\langle \mathrm{T}(\mathcal{A})\rangle| / k.$$

Доказательство. Заметим, что для любых двух сверхслов $a, b \in \mathcal{A}$ либо $\langle \mathrm{T}(a)\rangle = \langle \mathrm{T}(b)\rangle$, то есть $\mathrm{T}(a)$ и $\mathrm{T}(b)$ получаются друг из друга циклическим сдвигом, либо $\langle \mathrm{T}(a)\rangle \cap \langle \mathrm{T}(b)\rangle = \varnothing$.

Если множество $\mathcal{A}$ не прогнозируемо, то $\omega(\mathcal{A}) = \infty$ и утверждение леммы справедливо.

Предположим, что множество $\mathcal{A}$ прогнозируемо и $V(q_0) = (E_k, Q, E_k, \varphi, \psi, q_0)$ — прогнозирующий его автомат. Пусть этот автомат задан диаграммой Мура. Ребро e диаграммы Мура можно задавать парой (q, γ), где q — вершина диаграммы Мура, из которой исходит ребро, γ — входная буква, приписанная ребру (будем писать $e = (q, \gamma)$).

Через $\psi(e)$ и $\varphi(e)$ будем обозначать соответственно выходную букву, соответствующую ребру e, и вершину диаграммы Мура, в которую ведет ребро e.

Пусть $a_1, a_2 \in \mathcal{A}$ такие сверхслова, что $\langle \mathrm{T}(a_1)\rangle \cap \langle \mathrm{T}(a_2)\rangle = \varnothing$.

Фиксируем $i \in \{1, 2\}$. Пусть e_i — последовательность ребер диаграммы Мура, соответствующая входному сверхслову a_i, то есть для любого $t \geqslant 1$ выполнено $e_i(t) = (\varphi(a_i\underset{t-1}{]}, q_0), a_i(t))$. Здесь $\varphi(a_i\underset{0}{]}, q_0) = q_0$. Так как a_i — периодическое и Q конечное множество, то последовательность e_i — периодическая. Пусть l_i — период последовательности e_i и $e'_i(1)e'_i(2)\dots e'_i(l_i)$ — периодическая ее часть. Пусть $e'_i(t) = (q_i(t), \gamma_i(t))$, $t = 1, 2, \dots, l_i$.

Так как слово $\gamma_i(1)\gamma_i(2)\dots\gamma_i(l_i)$ — периодическое с периодом $\tau(a_i)$, то $l_i \geqslant \tau(a_i) = |\langle \mathrm{T}(a_i)\rangle|$.

По определению $\varphi(e'_i(t)) = q_i(t+1)$ для каждого $t \in \{1, 2, \dots, l_i\}$ (здесь и далее считаем, что $q_i(l_i+1) = q_i(1)$, $\gamma_i(l_i+1) = \gamma_i(1)$, $e_i(l_i+1) = e_i(1)$). Так как $V(q_0)$ прогнозирующий автомат, то $\psi(e'_i(t)) = \gamma_i(t+1)$ для каждого $t \in \{1, 2, \dots, l_i\}$.

Покажем, что множества ребер $\{e'_1(1), e'_1(2), \dots, e'_1(l_1)\}$ и $\{e'_2(1), e'_2(2), \dots, e'_2(l_2)\}$ не пересекаются. Предположим обратное, то есть пусть существуют $t_i \in \{1, 2, \dots, l_i\}$ $(i = 1, 2)$, что $e'_1(t_1) = e'_2(t_2)$. Так как $\langle \mathrm{T}(a_1)\rangle \cap \langle \mathrm{T}(a_2)\rangle = \varnothing$, то без ограничения общности можно считать, что $e'_1(t_1+1) \neq e'_2(t_2+1)$. Понятно, что $q_1(t_1+1) = q_2(t_2+1)$, следовательно $\gamma_1(t_1+1) \neq \gamma_2(t_2+1)$, и значит, $\psi(e'_1(t_1)) \neq \psi(e'_2(t_2))$, откуда следует, что $e'_1(t_1) \neq e'_2(t_2)$. Противоречие.

Следовательно, так как $a_1, a_2 \in \mathcal{A}$ — произвольные сверхслова, такие что $\langle \mathrm{T}(a_1)\rangle \cap \langle \mathrm{T}(a_2)\rangle = \varnothing$, то число ребер в диаграмме Мура автомата $V(q_0)$ не менее, чем $|\langle \mathrm{T}(\mathcal{A})\rangle|$. А так как из каждой вершины диаграммы Мура исходит ровно k ребер, то число состояний автомата $V(q_0)$ не менее, чем $|\langle \mathrm{T}(\mathcal{A})\rangle| / k$.

Лемма доказана.

Лемма 2 (Нижняя оценка). *Если $\mathcal{A}$ — некоторое множество периодических сверхслов, то*

$$\omega(\mathcal{A}) \geqslant \|\mathcal{A}\|.$$

Доказательство. Пусть f — такая частичная детерминированная функция, определенная на множестве всех префиксов сверхслов (включая сами сверхслова) из $\mathcal{A}$, что для любого $x \in \mathcal{A}$ найдется такое $N_x \geqslant 0$, что выполнено условие:

$$f(x)(t + N_x) = x(t + N_x + 1), \quad t \geqslant 0. \tag{1}$$

Если существует доопределение f до всюду определенной ограниченно-детерминированной функции, то оно будет задавать инициальный автомат из $\mathbb{G}\,(\mathcal{A})$. При этом, число состояний у такого автомата не меньше числа различимых остаточных функций функции f (см. [2, 3]). Будем обозначать через f_a остаточную функцию для f, соответствующую слову $a \in E_k^*$, то есть $f_a(x) \equiv f(ax)$.

Фиксируем $\gamma \in E_k$. Рассмотрим произвольные слова d_1, d_2 из $\mathrm{T}(\mathcal{A})$ такие, что $d_i \in \langle \mathrm{T}(a_i) \rangle$, $a_i \in \mathcal{A}$, $d_i(1) = \gamma$, $i = 1, 2$, $d_1 \neq d2$.

Пусть $b_1 = d_1^{\tau(a_2)}$, $b_2 = d_2^{\tau(a_1)}$. Тогда $|b_1| = |b_2|$, и b_1, b_2 представляются в виде:

$$b_1 = \gamma b c_1, \quad b_2 = \gamma b c_2, \tag{2}$$

где $|c_1| = |c_2| > 0,\ c_1(1) \neq c_2(1)$. Действительно, если предположить, что $b_1 = b_2$, то получим, что по крайней мере одно из слов d_1, d_2 не является циклическим сдвигом минимальной периодической части соответственно слов a_1, a_2, что противоречит изначальному предположению.

Пусть t'_i — такое число, что $a_i = a_i \underset{t_i}{]}\, b_i^{\infty}$, $f_i = f_{a_i \underset{t_i}{]}}$, где $t_i = \min\left\{t'_i + l \cdot \tau\left(a_i\right) \geqslant N_{a_i} : l \in \mathbb{N}\right\}$, $i = 1, 2$. Тогда из (1) и (2) следует, что $f_1(\gamma b) = bc_1(1) \neq bc_2(1) = f_2(\gamma b)$. То есть, f_1 и f_2 отличимы.

Таким образом, мы каждой паре различных слов из $\mathrm{T}(\mathcal{A})$, начинающихся на символ γ, сопоставили пару различных остаточных функций. Следовательно, число попарно различимых остаточных функций функции f не меньше, чем $\|\mathcal{A}\|_\gamma$. Взяв максимум по γ, получаем утверждение леммы.

Легко видеть, что $\|\mathcal{A}\| = \max\limits_{\gamma \in E_k} \|\mathcal{A}\|_\gamma \geqslant \left|\langle \mathrm{T}(\mathcal{A}) \rangle\right| / k$. Отсюда следует, что лемма 1 является следствием леммы 2.

4. Верхняя оценка сложности прогнозирующего автомата

Лемма 3. *Пусть $V = (E_k, Q, E_k, \psi, \varphi)$ — автомат. Тогда если для любого $q \in Q$ выполнено $V(q) \in \mathbb{G}(\mathcal{A})$, то для любого $q \in Q$ выполнено $V(q) \in \mathbb{G}(\mathcal{A}^*)$.*

Доказательство. Пусть $V(q) \in \mathbb{G}(\mathcal{A})$ для любого $q \in Q$. Возьмем произвольное состояние $q' \in Q$. Возьмем произвольное слово $a \in \mathcal{A}^*$. По определению $a = bc$, где $b \in E_k^*$, $c \in \mathcal{A}$. Пусть $\varphi(b, q') = q''$. Так как автомат $V(q'')$ прогнозирует слово $c \in \mathcal{A}$, то автомат $V(q')$ прогнозирует слово $a = bc \in \mathcal{A}^*$. В силу произвольности слова a, $V(q') \in \mathbb{G}(\mathcal{A}^*)$.

Лемма доказана.

Лемма 4 (Верхняя оценка). *Если $\mathcal{A}$ — некоторое множество периодических сверхслов, то*

$$\omega(\mathcal{A}^*) \leqslant |\langle \mathrm{T}(\mathcal{A}) \rangle| .$$

Доказательство. Пусть $\mathcal{B} = \{(\mathrm{T}(a))^\infty : a \in \mathcal{A}\}$, то есть $\mathcal{B}$ получено из $\mathcal{A}$ отбрасыванием предпериодической части у сверхслов из $\mathcal{A}$. Понятно, что $\mathcal{A}^* = \mathcal{B}^*$ и $|\langle \mathrm{T}(\mathcal{A}) \rangle| = |\langle \mathrm{T}(\mathcal{B}) \rangle|$, поэтому сразу можно считать, что $\mathrm{T}'(\mathcal{A}) = \{\Lambda\}$.

Пусть $\alpha \in E_k$, $u, q \in E_k^*$. Обозначим через $\widehat{u}$ слово u, записанное в обратном порядке, то есть $\widehat{u} = u(|u|)u(|u|-1)\ldots u(1)$. Обозначим

$$r(\alpha, u, q) = \max\left\{l : (\alpha\widehat{u}^\infty)\underset{l}{]} = (\widehat{q}^\infty)\underset{l}{]}, l \in \{0, 1, 2, \ldots\}\right\}.$$

Здесь $r(\alpha, u, q) \leqslant 1$, если $u = \Lambda$; $r(\alpha, u, q) = 0$, если $q = \Lambda$; и $r(\alpha, u, q) = \infty$, если $\alpha\widehat{u}^\infty = \widehat{q}^\infty$, $u \neq \Lambda$, $q \neq \Lambda$. Функция $r(\alpha, u, q)$ говорит на сколько совпадают начала слов $\alpha\widehat{u}^\infty$ и $\widehat{q}^\infty$, или, что то же самое, насколько совпадают концы слов, где первое слово получено приписыванием к букве α слева сколь угодно много раз слова u, а второе слово получено приписыванием к слову q слева сколь угодно много раз слова q. Причем $\alpha\widehat{u}^\infty = \widehat{q}^\infty$ только тогда, когда α есть последняя буква слова q, а слово u или некоторая его степень есть циклический сдвиг вправо на единицу слова q.

Обозначим

$$Q(\mathcal{A}) = \langle \mathrm{T}(\mathcal{A}) \rangle\,,$$
$$R(\alpha, u) = \max_{q \in Q(\mathcal{A})} r(\alpha, u, q),$$
$$A(\alpha, u) = \{q \in Q(\mathcal{A}) : r(\alpha, u, q) = R(\alpha, u)\}\,.$$

Через $s(q) = q(2) \ldots q(|q|)q(1)$ обозначим циклический сдвиг на единицу влево слова q.

Поскольку для любых $q_1, q_2 \in Q(\mathcal{A})$, $q_1 \neq q_2$, выполнено $q_1^\infty \neq q_2^\infty$, и поскольку $q(1)\widehat{q}^\infty = (\widehat{s(q)})^\infty$, то для всех $q \in Q(\mathcal{A})$ выполняется соотношение $r(q(1), q, s(q)) = \infty$, а, учитывая, что $s(q) \in Q(\mathcal{A})$, то выполняется

$$A\,(q(1), q) = \{s(q)\}. \tag{3}$$

Поскольку мы предположили, что $\mathrm{T}'(\mathcal{A}) = \{\Lambda\}$, то для любого $x \in \mathcal{A}$ существует такое слово $q \in Q(\mathcal{A})$, что $q = \mathrm{T}(x)$, то есть $x = q^\infty$. Отсюда следует, что для любого $t \geqslant 1$ существует такое слово $q' \in Q(\mathcal{A})$, что

$$(\widehat{q'})^\infty]_t = \widehat{x}]_t, \tag{4}$$

то есть приписывая к слову q' слева достаточное количество раз слово q', мы получим слово, конец которого полностью совпадает со словом $x]_t$. Причем в качестве q' можно взять слово, полученное из q циклическим сдвигом влево на остаток от деления t на $\tau\,(x) = |q|$.

Рассмотрим некоторое отображение $\Phi : E_k \times E_k^* \to Q(\mathcal{A})$ такое, что $\Phi(\alpha, u) \in A(\alpha, u),\ \alpha \in E_k,\ u \in E_k^*$.

Рассмотрим автомат $V_{\mathcal{A}} = \big(E_k, Q(\mathcal{A}), E_k, \varphi, \psi\big)$ со следующими каноническими уравнениями:

$$\begin{cases} q(t+1) = \varphi\big(q(t), x(t)\big) = \Phi\big(x(t), q(t)\big), \\ y(t) = \psi\big(q(t), x(t)\big) = \Phi\big(x(t), q(t)\big)(1). \end{cases} \tag{5}$$

Понятно, что запись $\Phi\big(x(t), q(t)\big)(1)$ означает первую букву слова $\Phi\big(x(t), q(t)\big)$.

Отметим, что если $q(t)(1) = x(t)$, то согласно (3)

$$q(t+1) = s(q(t)).$$

Покажем по индукции, что для любых $t \geqslant 1$, $x \in \mathcal{A}$ и $q(1) \in Q(\mathcal{A})$ выполнено равенство

$$(\widehat{q(t+1)})^{\infty}\underset{t}{]} = \widehat{x}\underset{t}{]}, \tag{6}$$

что означает, что каждое состояние помнит все поступившие до этого буквы входного слова.

Базис индукции. $t = 1$.

$q(2) = \Phi(x(1), q(1)) \in A(x(1), q(1))$. Так как в $Q(\mathcal{A})$ есть состояние, заканчивающееся на $x(1)$, например, $s(\mathrm{T}(x))$, то $q(2)(|q(2)|) = x(1)$.

Индуктивный переход. По предположению индукции имеем $(\widehat{q(t)})^{\infty}\underset{t-1}{]} = \widehat{x\underset{t-1}{]}}$. Следовательно, $x(t)((\widehat{q(t)})^{\infty}\underset{t-1}{]}) = x(t)\widehat{x\underset{t-1}{]}} = \widehat{x}\underset{t}{]}$. Поскольку согласно (4) существует $q' \in Q(\mathcal{A})$ такое, что $(\widehat{q'})^{\infty}\underset{t}{]} = \widehat{x}\underset{t}{]}$, то и для $q(t+1) = \Psi(x(t), q(t)) \in A(x(t), q(t))$ выполнено $(\widehat{q(t+1)})^{\infty}\underset{t}{]} = \widehat{x}\underset{t}{]}$.

Тем самым равенство (6) доказано.

Обозначим $q' = \mathrm{T}(x)$. Из равенства (6) следует, что при достаточно больших t (например, при $t > |q'| + \max_{q \in Q(\mathcal{A})} |q|$) состояние $q(t+1)$ определяется однозначно и это состояние, полученное из q' циклическим сдвигом влево на остаток от деления t на $|q'|$, то есть $q(t+1) = x(t - \tau(x) + 1)x(t - \tau(x) + 2)\ldots x(t)$. Поэтому $y(t) = q(t+1)(1) = x(t - \tau(x) + 1) = x(t+1)$.

Таким образом, мы показали, что для любого $q(1)$ автомат $V_{\mathcal{A}}(q(1)) \in \mathbb{G}(\mathcal{A})$, откуда в силу леммы 3 имеем $V_{\mathcal{A}}(q(1)) \in \mathbb{G}(\mathcal{A}^*)$. Лемма доказана.

В доказательстве леммы 4 содержится оптимальный по порядку метод построения прогнозирующего автомата для произвольного прогнозируемого множества периодических сверхслов. Суть алгоритма в следующем. В качестве алфавита состояний выбираем $\langle \mathrm{T}(\mathcal{A}) \rangle$ — множество всех циклических сдвигов периодических частей сверхслов из $\mathcal{A}$. Из состояния q по $x(t)$ переходим в такое состояние q', что слова $x(t)(\widehat{q})^{\infty}$ и $(\widehat{q'})^{\infty}$ имеют общий префикс наибольшей длины, или, что то же самое, слово, полученное приписыванием к букве $x(t)$ слева достаточное количество раз слова q, имеет общий суффикс наибольшей длины со словом, полученным приписыванием к слову

q' слева достаточное количество раз слова q'. Множество всех таких q' обозначаем $A(x(t), q)$. Если $x(t) = q(1)$, то q' представляет собой циклический сдвиг влево слова q. Выходное значение $y(t)$ полагаем равным $q'(1)$.

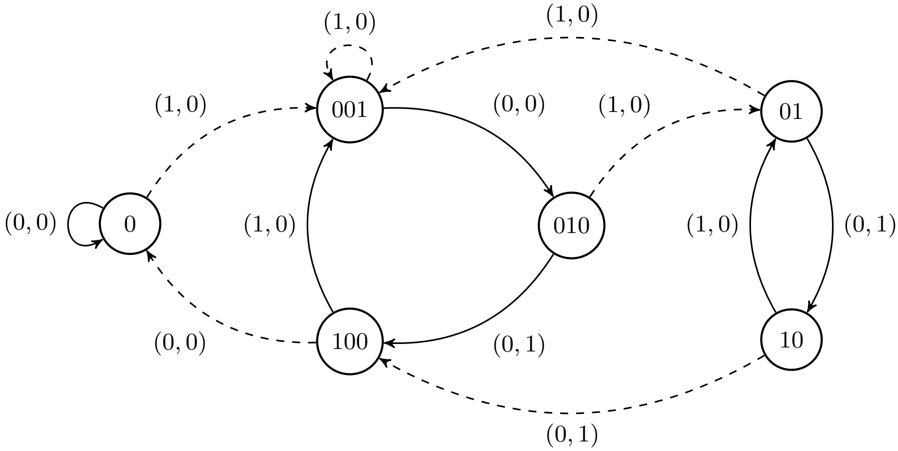

Рис. 1. Прогнозирующий автомат для примера.

Пример. Построить прогнозирующий автомат со вход-выходным алфавитом $\{0, 1\}$ для множества сверхслов $\mathcal{A} = \{0^\infty, (01)^\infty, (001)^\infty\}$.

Таблица 1.

q	0	01	10	001	010	100
$A(\overline{q(1)}, q)$	{001}	{01,001}	{100}	{01,001}	{01}	{0}

Решение. $Q(\mathcal{A}) = \langle \mathrm{T}(\mathcal{A}) \rangle = \{0, 01, 10, 001, 010, 100\}$. Значения для $A\left(q(1), q\right)$ определяются однозначно и равны $\{s(q)\}$. Значения для $A(\overline{q(1)}, q)$ определены в таблице 1. На рисунке 1 приведена диаграмма Мура автомата, построенного по указанному методу. Пунктиром обозначены переходы из q по $\overline{q(1)}$, а сплошными линиями — по $q(1)$. Тем самым переходы по сплошным линиям образуют циклы, по которым будут ходить слова из $\mathcal{A}$, после того как выйдут на свои циклы. Таких циклов 3: $\{0\}$, $\{01, 10\}$ и $\{001, 010, 100\}$. Пунктирные переходы помогают входным словам найти свои циклы. Поскольку, как мы показали в лемме 1, наличие сплошных переходов абсолютно необходимо, то сокращение числа состояний прогнозирующего автомата возможно только за счет сокращения пунктирных переходов. В данном случае оказывается, что автомат можно минимизировать. Полученный после минимизации автомат изображен на рисунке 2. Количество его состояний равно четырем, что совпадает с нижней оценкой из леммы 2 $\|\mathcal{A}\| = \max\{\|\mathcal{A}\|_0, \|\mathcal{A}\|_1\} = \max\{4, 2\} = 4$. Таким образом, построенный автомат прогнозирует $\mathcal{A}^*$ и является оптимальным по числу состояний.

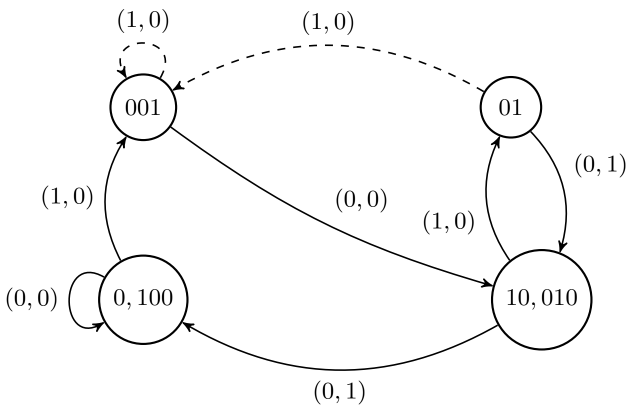

Рис. 2. Прогнозирующий автомат для примера после минимизации.

5. Критерий прогнозируемости и порядок сложности прогнозирующего автомата

Докажем теорему 1.

Необходимость. Пусть $\mathcal{A}$ прогнозируемо и $V(q_0) = (E_k, Q, E_k, \varphi, \psi, q_0)$ — прогнозирующий его автомат. Тогда для любого $x \in \mathcal{A}$ найдется такое натуральное число N, что $x(t+1) = \psi(q_0, x\underset{t}{]})$ при $t \geqslant N$. Тогда сверхслово $x(N+1)x(N+2)\ldots$ можно рассматривать, как генерируемое конечным автономным автоматом. В самом деле, пусть q' — состояние автомата $V(q_0)$ после N тактов. Возьмем автомат $V(q')$ и подадим на его вход через задержку с начальным состоянием $x(N)$

его собственный выход. Получаем автономный автомат, который на выходе будет выдавать сверхслово $x(N+1)x(N+2)\ldots$. Откуда согласно теореме 16 из [4] это слово будет периодическим. Следовательно, x также периодично. Поскольку можно считать, что вход автономного автомата периодическое слово с длиной периода 1, то согласно теореме 16 из [4] длина периода сверхслова x не больше $|Q|$.

Достаточность. Пусть $\mathcal{A}$ состоит из периодических сверхслов с ограниченным периодом. Тогда по лемме 4 существует прогнозирующий его автомат с количеством состояний, не превосходящим $|\langle \mathrm{T}(\mathcal{A})\rangle|$.

Теорема 1 доказана.

Поскольку $|\langle \mathrm{T}(\mathcal{A})\rangle|/k \leqslant \|\mathcal{A}\|$, то теорема 2 является простым следствием лемм 2, 4 и теоремы 1.

Список литературы

[1] Вереникин А. Г., Гасанов Э. Э. Об автоматной детерминизации множеств сверхслов // Дискретная математика. — 2006. — Т. 18, № 2. — С. 84–97.

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

[3] Трахтенброт Б. А., Барздинь Я. М. Конечные автоматы (поведение и синтез). — М.: Наука, 1970.

[4] Яблонский С. В. Введение в дискретную математику. — М.: Высшая школа, 2008.