УДК 519.713.4

И.К. РЫСЦОВ

ТЕОРИЯ ОБОБЩЕННЫХ ЛИНЕЙНЫХ АВТОМАТОВ

Ключевые слова: *линейные автоматы, аффинные автоматы, линейные дискретные системы*.

ВВЕДЕНИЕ

Почти 40 лет назад А.А. Мучник опубликовал пионерскую работу об общих линейных автоматах [1]. Основная цель данной статьи — систематизация результатов в этой области, полученных к настоящему времени. В работе определяются линейные и аффинные автоматы и вводится понятие размерности для конечного автомата. Рассматривается проблема достижимости состояний в автоматах и доказывается, что она алгоритмически неразрешима для двумерных аффинных автоматов. Далее рассматриваются линейные автоматы с выходом и доказывается аналог теоремы Мура об эквивалентных состояниях, а также линейные аналоги теорем для установочных и диагностических слов, рассмотрены приложения линейных автоматов в математической экономике.

1. ЛИНЕЙНЫЕ АВТОМАТЫ

(Полным) линейным n-мерным автоматом (без выхода) A над полем K называется тройка $A=(V, X, f)$, где $V=K^n$ — линейное n-мерное пространство состояний (векторов длины n) над полем K, $X=\{a_1,\ldots,a_k\}$ — конечный входной алфавит, f — отображение вида $f: X \to \mathrm{Mat}(n,K)$, которое каждому символу из $x \in X$ ставит в соответствиие квадратную матрицу n-го порядка $f(x)$ с элементами из K. Число n называется размерностью линейного автомата.

Отображение f определяет функцию переходов $F(v,x)=v\cdot f(x)$, где справа стоит произведение вектора-строки $v=(k_1,\ldots,k_n)$ на $n\times n$ матрицу $f(x)$. В этом случае функция переходов будет линейной только по первому аргументу, поэтому А. Мучник назвал эти автоматы «общими» линейными автоматами [1], чтобы подчеркнуть их отличие от классических линейных автоматов. Подробнее на этом отличии остановимся ниже.

Поскольку множество матриц $\mathrm{Mat}(n,K)$ является мультипликативным моноидом, то отображение f можно продолжить до гомоморфизма, определенного на свободном моноиде X^*, который каждому входному слову (моному), $w=x_1\ldots x_m$, $x_i \in X$, $1\le i\le m$, сопоставляет произведение базовых матриц:

$$f(w)=f(x_1)\cdot f(x_2)\cdot\ldots\cdot f(x_m). \qquad (1)$$

Заметим, что $f(e)=I$, где e — пустое слово, I — единичная матрица. Тогда функция переходов определяется для любых входных слов по формуле $F(v,w)=$ $=v\cdot f(w)$, где правая часть по-прежнему понимается как произведение вектора на матрицу.

В теории линейных систем кроме обычных входных символов принято рассматривать «обобщенные» входы (входные векторы), которые являются линейными комбинациями элементов из X с коэффициентами из поля K, т.е. элементами линейного пространства $K(X)$. Для входного вектора $u=c_1a_1+\ldots+c_ka_k$ матрица переходов $f(u)$ определяется линейным образом $f(u)=c_1 f(a_1)+\ldots+c_k f(a_k)$. Итак, линейный автомат можно рассматривать как тройку $A=(V, X, f)$, где $f: K(X)\to \mathrm{Mat}(n,K)$ — линейное отображение векторных пространств.

Затем функция f продолжается на «обобщенные» слова $p=u_1\ldots u_m$, $u_i \in K(X)$, $1\le i\le m$, по аналогии с формулой (1). Но поскольку множество матриц

$\mathrm{Mat}\,(n, K)$ является алгеброй, то гомоморфизм (1) можно продолжить до гомоморфизма алгебр следующим образом. Свободный моноид X^* естественно вкладывается в свободную ассоциативную алгебру $K(X^*)$, которая состоит из полиномов от не коммутирующих переменных из X с коэффициентами из поля K. Тогда функцию переходов можно доопределить для любого полинома $p = c_1 w_1 + \ldots + c_m w_m$ линейным образом:

$$f(p) = c_1 f(w_1) + \ldots + c_m f(w_m), \; F(v, p) = v \cdot f(p). \tag{2}$$

Хотя автоматный смысл имеют только обобщенные слова, которые можно рассматривать как произведения полиномов первой степени, однако общий подход обладает рядом преимуществ. Функция переходов $F : V \times K(X^*) \to V$ становится линейной по обоим аргументам и, следовательно, линейный автомат можно рассматривать как модуль над кольцом и пользоваться стандартными алгебраическими конструкциями из теории модулей [2]. Это позволяет обобщить подход Калмана к дискретным линейным системам, как модулям над кольцом полиномов [3].

Рассмотрим некоторые свойства линейных автоматов. Обозначим $\mathrm{Kr}\, f(p)$ ядро матрицы $f(p)$, т.е. подпространство состояний, переходящих в ноль под действием полинома p, а $\mathrm{Im}(p)$ — подпространство ее значений (порожденное строками матрицы). Размерность подпространства $U \subseteq V$ обозначим $\dim(U)$. Действие линейного автомата можно распространить на подмножества состояний $S \subseteq V$ по формуле $F(S, p) = \{F(s, p) \mid s \in S\}$. Если U — подпространство, то его образ $F(U, p)$ также будет подпространством, для которого выполняется следующее условие монотонности:

$$\dim(F(U, p)) \le \dim(U). \tag{3}$$

Кроме того, выполняется следующее условие, верное для любого подпространства U и полинома p:

$$\dim(F(U, p)) < \dim(U) \leftrightarrow (U \cap \mathrm{Kr}\, f(p)) \ne \{0\}. \tag{4}$$

Подмножество $S \subseteq V$, инвариантное относительно умножения на все входные символы, т.е. $F(S, x) \subseteq S$ для всех $x \in X$, определяет мультипликативный подавтомат $B = (S, V, X, f)$ линейного автомата $A = (V, X, f)$. Заметим, что подмножество S может быть не замкнуто относительно сложения состояний, однако, взяв его линейную оболочку $\mathrm{lin}\,(S)$, получим подпространство, инвариантное относительно сложения и умножения, т.е. линейный подавтомат или подмодуль. Отметим, что процедура построения фактор-автомата по линейному подавтомату сводится в этом случае к стандартной алгебраической процедуре построения фактор-модуля по подмодулю [2].

2. ОБОБЩЕННЫЕ ЛИНЕЙНЫЕ АВТОМАТЫ

Определение линейного автомата, которое было дано в предыдущем разделе довольно жесткое. Например, не каждый конечный автомат оказывается полным линейным автоматом над конечным полем, поскольку число элементов в конечном поле — степень простого числа. Поэтому А. Мучник дает следующее более общее определение [1].

Определение 1. Обобщенным линейным автоматом $B = (S, V, X, f)$ над полем K называется мультипликативный подавтомат полного линейного автомата $A = (V, X, f)$ над этим полем.

В этом случае подмножество S называется множеством состояний подавтомата. Размерность автомата B равна размерности подпространства $\mathrm{lin}\,(S)$. Чтобы выделить линейные автоматы, у которых $S = V$, мы назвали их «полными» линейными автоматами.

Обобщенные линейные автоматы классифицируются по структуре их множества состояний. Сначала рассмотрим случай, когда множество состояний подавтомата конечно.

Пусть $A = (S, X, \delta)$ — конечный детерминированный автомат (без выхода) с n состояниями $S = \{s_1, \ldots, s_n\}$ и K — произвольное поле. Рассмотрим известное матричное представление конечного автомата. Для этого с каждым $x \in X$ свяжем двоичную $n\times n$ матрицу $f_0(x)$ следующим образом: $e_{ij}(x)=1$, если $\delta(s_i, x)=s_j$, и $e_{ij}(x)=0$ в противном случае. Погрузим множество S в линейное пространство всех формальных линейных комбинаций $K(S)= =\{c_1\cdot s_1+\ldots+c_n\cdot s_n \mid c_i\in K, s_i\in S\}$ и доопределим функцию переходов для всех линейных комбинаций v:

$$F(v,x)=v\cdot f_0(x)=c_1\cdot\delta(s_1,x)+\ldots+c_n\cdot\delta(s_n,x).$$

Отметим, что конечный автомат A будет действовать на базисе S. Значит, A будет конечным мультипликативным подавтоматом линейного n-мерного автомата $K(A)=(K(S), X, f_0)$, который называется стандартным линейным расширением конечного автомата A над полем K. Таким образом, имеем следующее утверждение [1].

Предложение 1. Каждый конечный автомат с n состояниями является обобщенным линейным n-мерным автоматом над любым полем.

Для дальнейшего нам понадобится также подавтомат $K_2(A)$ линейного автомата $K(A)$, который называется линейным автоматом пар.

Следствие 1. Подпространство $K_2(S)$, порожденное состояниями s_i-s_j, $1\le i<j\le n$, в линейном расширении конечного автомата A, определяет линейный подавтомат $K_2(A)=(K_2(S), X, f_0)$ размерности $n-1$.

В связи с предложением 1 возникает важное понятие размерности конечного автомата. Для простоты определим его только для двоичного поля $GF(2)=\{0,1\}$, поскольку для остальных конечных полей оно определяется аналогично.

Определение 2. Размерностью $\dim_2(A)$ конечного автомата A называется наименьшее число d такое, что автомат A изоморфно вкладывается в линейный автомат размерности d над двоичным полем.

Из этого определения и предложения 1 получаем следующие неравенства для размерности конечного автомата A с n состояниями:

$$\log_2(n)\le\dim_2(A)\le n. \qquad (5)$$

Нижняя оценка здесь достигается на любом линейном автомате A размерности d над полем $GF(2)$, поскольку он будет иметь $n=2^d$ состояний. Верхняя оценка в (5) достигается на автомате Черны, который задается двумя подстановками $\delta(x)=(12\ldots(n-1)1)$ и $\delta(y)=(23\ldots n1)$ на множестве состояний $S=[1,n]$.

Теорема 1. Размерность автомата Черны с n состояниями равна n.

Доказательство. Пусть $A=([1,n],\{x,y\},\delta)$ — автомат Черны с n состояниями. Предположим противное, тогда должен существовать взаимно однозначный мономорфизм (вложение) $\varphi\colon A\to(GF(2)^d, X, f)$, где $d<n$. Но тогда состояния из подмножества $\varphi([1,n])$ должны быть линейно зависимы, и, следовательно, существует состояние $\varphi(i)$, которое линейно выражается через предыдущие $\varphi(i)= =\varphi(i_1)+\ldots+\varphi(i_m)$, где $0\le i_j<i$, $1\le j\le m$. Умножая это равенство на слово y^{n-i}, увеличиваем номера всех состояний на $n-i$ и получаем $\varphi(n)=\varphi(l_1)+\ldots+\varphi(l_m)$, где $0\le l_j<n$, $1\le j\le m$. Наконец, умножая это равенство на x, получаем $\varphi(1)=\varphi(n)$, что противоречит взаимной однозначности φ.

Таким образом, теорема доказана.

Следующий важный частный случай обобщенных линейных автоматов связан с вероятностными автоматами, когда множество состояний — ограниченное выпуклое множество (симплекс). Вероятностным автоматом с n состояниями (без выхода) называется тройка $A=(S, X, f_1)$, где S — конечное множество состояний, X — конечное множество входных символов, а f_1 — набор стохастических матриц $f_1\colon X\to\mathrm{Mat}(n,R^+)$. В этом случае из базисных состояний S будут достижимы только распределения $(r_1,\ldots,r_n)$, расположенные в единичном n-мерном симплексе Δ_n, т.е. точки, координаты которых удовлетворяют уравнению $r_1+\ldots+r_n=1$, где $0\le r_i\le 1$, $1\le i\le n$. Значит, симплекс $\Delta_n\subset R^n$ будет в этом случае мультипликативным подавтоматом линейного автомата $R(A)=(R^n, X, f_1)$, который естественно назвать линейным расширением вероятностного автомата. Таким образом, получаем следующее утверждение [1].

Предложение 2. Каждый вероятностный автомат с *n* состояниями является обобщенным линейным *n*-мерным автоматом над полем вещественных чисел.

Для вероятностных автоматов подавтомат пар также оказывается полезным, но эту конструкцию можно рассмотреть и с общих позиций.

Определение 3. Мультипликативный подавтомат $B=(S, V, X, f)$ линейного автомата $A=(V, X, f)$ над полем K назовем аффинным, если множество состояний S является аффинной гиперплоскостью в V.

Напомним, что аффинное подпространство S — это линейное подпространство L, сдвинутое на некоторый вектор v, т.е. $S=L+v$. В этом случае подпространство L называется направляющим (несущим) для S [4]. Для гиперплоскости размерность направляющего подпространства равна $n-1$. Из конструкций линейных расширений для конечных и вероятностных автоматов видно, что множества их состояний естественно вкладываются в инвариантную аффинную гиперплоскость $k_1+\ldots+k_n=1$. Значит, любой конечный или вероятностный автомат можно расширить до аффинного подавтомата. Кроме того, из мультипликативной замкнутости аффинного подпространства S следует мультипликативная замкнутость ее направляющего подпространства L, следовательно, автомат пар (L, X, f) можно определить для любого аффинного подавтомата $B=(S, V, X, f)$.

Следствие 2. Для любого аффинного подавтомата $B=(S, V, X, f)$ подпространство, порожденное состояниями $s-t$, $s, t \in S$, определяет линейный подавтомат пар $B_2=(L, X, f)$ размерности $n-1$.

3. ДОСТИЖИМОСТЬ СОСТОЯНИЙ В ЛИНЕЙНЫХ АВТОМАТАХ

Достижимость состояний в линейных автоматах можно трактовать с двух точек зрения в зависимости от того, допускаются полиномы на входе автомата или достижимость осуществляется только с помощью мономов. Но общие результаты, не зависящие от основного поля, удается получить только для обобщенной достижимости.

Пусть $A=(V, X, f)$ — полный линейный автомат, тогда множеством мономиально достижимых состояний для состояния $v \in V$ в автомате A называется следующее множество: $\mathrm{Rm}_A(v)=\{F(v,w) \mid w \in X^*\}$. При ограничении на длину слов соответствующее множество достижимости обозначим $\mathrm{Rm}_i(v)=\{F(v,w) \mid l(w) \le i, w \in X^*\}$, где $l(w)$ — длина слова, $i \ge 0$. Соответствующие множества полиномиально достижимых состояний обозначим $\mathrm{Rp}_A(v)=\{F(v,p) \mid p \in K(X^*)\}$ и $\mathrm{Rp}_i(v)=$ $=\{F(v,p) \mid d(p) \le i, p \in K(X^*)\}$, где $d(p)$ — степень полинома p.

Нетрудно видеть, что множества $\mathrm{Rp}_i(v)$ будут линейными подпространствами, поскольку из свойства (2) вытекает равенство для любого полинома $p=c_1w_1+\ldots+c_mw_m$:

$$F(v, p)=c_1F(v, w_1)+\ldots+c_mF(v, w_m). \tag{6}$$

Отсюда следует, что $\mathrm{Rp}_i(v)$ — линейная оболочка для $\mathrm{Rm}_i(v)$:

$$\mathrm{Rp}_i(v)=\mathrm{lin}(\mathrm{Rm}_i(v)), \quad i \ge 0. \tag{7}$$

Далее из определений очевидно, что подпространства $\mathrm{Rp}_i(v)$ образуют возрастающую (по i) последовательность

$$\mathrm{Rp}_0(v) \subseteq \mathrm{Rp}_1(v) \subseteq \ldots \subseteq \mathrm{Rp}_n(v) \subseteq \ldots . \tag{8}$$

Этот ряд должен стабилизироваться в силу конечномерности пространства состояний. Действительно, для ненулевого состояния v имеем $\dim(\mathrm{Rp}_0(v))=1$, следовательно, строгие неравенства в ряду (8) могут продолжаться не далее n-го места, и найдется такое число $i, 1 \le i \le n$, что $\mathrm{Rp}_{i-1}(v)=\mathrm{Rp}_i(v)$. После этого ряд (8) должен стабилизироваться, поскольку $F(\mathrm{Rp}_{i-1}(v), x) \subseteq \mathrm{Rp}_{i-1}(v)$ для всех $x \in X$. Значит, в этом случае $\mathrm{Rp}_{i-1}(v)=\mathrm{Rp}_{n-1}(v)=\mathrm{Rp}_A(v)$ и мы приходим к следующему утверждению, которое верно для любого поля, любого n-мерного автомата A над этим полем и любого состояния v:

$$\mathrm{Rp}_A(v) = \mathrm{Rp}_{n-1}(v). \tag{9}$$

Этому равенству можно придать форму, аналогичную конечному случаю, но вместо конечности здесь фигурирует конечномерность.

Теорема 2. Если в n-мерном линейном автомате A над полем K выполняется условие $t \in \mathrm{Rp}_A(s)$, то существует полином p степени, не большей $n-1$, такой, что $F(s, p) = t$.

Отметим, что $\mathrm{Rp}_{n-1}(v)$ будет наименьшим подмодулем, содержащим состояние v. Эту теорему можно доказать и для обобщенных слов $p = u_1 \ldots u_m$, если положить $L_m(v) = \{F(v, p) \mid l(p) \le m, p \in K(X)^*\}$ и строить подпространства $L_m(v)$ индуктивно по формуле:

$$L_m(v) = L_{m-1}(v) + \sum_{j=1}^{k} F(L_{m-1}(v), a_j), \ 1 \le m \le n.$$

Отсюда следует, что стабилизация здесь также наступит не позже $n-1$ шага и будет выполняться равенство $L_{n-1}(v) = \mathrm{Rp}_{n-1}(v)$. Таким образом, эти два вида полиномиальной достижимости по существу совпадают между собой.

Линейный автомат $A = (V, X, f)$ над полем K назовем сильно-связным, если для любых двух ненулевых состояний s и t из S найдется полином p такой, что $F(s, p) = t$. Отметим, что линейный сильно-связный автомат может быть мономиально не связным как абстрактный автомат [1]. Если полный линейный автомат сильно-связный, то с точки зрения алгебры он является простым (неприводимым) модулем, т.е. модулем, не содержащим ненулевых подмодулей [2].

Проблема мономиальной достижимости состояний в линейных автоматах может рассматриваться как алгоритмической проблема, в которой по заданным A, s и t требуется проверить существование слова $w \in X^*$ такого, что $F(s, w) = t$. К сожалению, над бесконечными полями проблема мономиальной достижимости состояний в линейных автоматах алгоритмически неразрешима, начиная с трехмерных автоматов [5]. Кроме того, она разрешима для одномерных линейных автоматов [6] и открыта в двумерном случае.

4. АФФИННЫЕ АВТОМАТЫ

Аффинные автоматы в силу их важности определим явным образом. Аффинным n-мерным автоматом A над полем K называется четверка объектов $A = (V, X, f, g)$, где $V = K^n$ — линейное n-мерное пространство состояний над полем K, $X = \{a_1, \ldots, a_k\}$ — конечный входной алфавит, f — отображение вида $f: X \to \mathrm{Mat}(n, K)$, а g — отображение вида $g: X \to V$. Функция переходов аффинного автомата определяется следующим образом:

$$F(v, x) = v \cdot f(x) + g(x), \ v \in V, \ x \in X. \tag{10}$$

Напомним, что аффинным отображением линейного пространства называется линейное отображение, сдвинутое на некоторый вектор [4], а аффинное отображение переводит аффинные подпространства снова в аффинные подпространства. Кроме того, аффинное отображение на пространстве V будет линейным на аффинных комбинациях состояний $c_1 \cdot v_1 + \ldots + c_m \cdot v_m$, т.е. комбинациях, у которых $c_1 + \ldots + c_m = 1$. Из формулы (10) видно, что функция переходов аффинного автомата при фиксированном $x \in X$ будет аффинным отображением пространства V и, следовательно, для аффинных комбинаций состояний выполняется условие

$$F(c_1 \cdot v_1 + \ldots + c_m \cdot v_m, x) = c_1 \cdot F(v_1, x) + \ldots + c_m \cdot F(v_m, x).$$

Заметим, что линейные автоматы являются частным случаем аффинных, когда $g(x) = 0$ для всех $x \in X$. С другой стороны, следующее предложение показывает, что понятие аффинного автомата согласовано с понятием аффинного подавтомата линейного автомата.

Теорема 3. Каждый аффинный n-мерный автомат изоморфно вкладывается в $(n+1)$-мерный линейный автомат.

Доказательство. Пусть $A=(V, X, f, g)$ — аффинный n-мерный автомат над полем K, тогда в пространстве K^{n+1} выделим аффинную гиперплоскость $S = \{(v,1) \mid v \in K^n\}$ и определим функцию $f_1: X \to \mathrm{Mat}\,(n+1, K)$ следующим образом:

$$f_1(x) = \begin{bmatrix} f(x) & \bar{0}_n \\ g(x) & 1 \end{bmatrix},$$

где $\bar{0}_n$ — нулевой вектор-столбец длины n. Тогда будем иметь $(v,1)\cdot f_1(x) = (F(v,x),1)$ для всех $x \in X$, где $F(v,x)$ — функция переходов аффинного автомата. Следовательно, автомат A изоморфен мультипликативному подавтомату линейного автомата (K^{n+1}, X, f_1).

Теорема доказана.

В некоторых прикладных ситуациях необходима более широкая трактовка аффинного автомата, когда функции f и g смещаются на некоторое постоянное значение. Чтобы не усложнять формальное изложение, будем просто считать, что в нашем распоряжении имеется дополнительный вход a_0, на котором определены функции f и g и который не входит во входной алфавит, а используется только для смещения. Тогда наши аффинные автоматы будут обобщать классические линейные автоматы, в которых функция переходов имеет вид $F(v,u) = v\cdot F_0 + u\cdot G_0$, где $u \in K(X)$ — входной вектор, а F_0 и G_0 — постоянные матрицы размерности $n\times n$ и $k\times n$ соответственно [7, 8]. Действительно, в нашем случае достаточно продолжить по линейности функции f и g для всех векторов a_0+u, где $u \in K(X)$, положив $f(a_0)=F_0$, $g(a_0)=0$, $f(x)=0$ для всех $x \in X$. Тогда из (10) следует, что $F(v, a_0+u) = v\cdot F_0 + g(u) = v\cdot F_0 + u\cdot G_0$, где $g(u) = u\cdot G_0$. Таким образом, получается классический линейный автомат, который правильнее было бы назвать классическим аффинным автоматом.

Далее продолжим функцию переходов аффинного автомата на входные слова. Пусть $w = x_1x_2\ldots x_m$ — входное слово, $x_i \in X$, $1\le i\le m$, и пусть $\sigma_i(w) = x_{i+1}\ldots x_m$ — i-й суффикс слова w, $0\le i\le m$. Предполагается, что $\sigma_m(w)=e$, где e — пустое слово. Определим теперь функцию f на входных словах по формуле (1), положив $f(w) = f(x_1)\cdot\ldots\cdot f(x_m)$. Следующая обобщающая формула, которая легко доказывается по индукции, определяет функцию переходов для аффинных автоматов:

$$F(v,w) = v\cdot f(w) + \sum_{i=1}^{m} g(x_i)\cdot f(\sigma_i(w)). \tag{11}$$

Вектор, равный значению суммы в правой части этой формулы, обозначим $g(w)$, тогда формула (11) примет более лаконичный вид:

$$\mathrm{F}(v, \mathrm{w}) = v\cdot \mathrm{f}(\mathrm{w}) + \mathrm{g}(\mathrm{w}). \tag{12}$$

Продолжим теперь функцию переходов на полиномы. Пусть $p = c_1\cdot w_1 + \ldots + c_m\cdot w_m$ — входной полином, тогда по определению положим:

$$f(p) = c_1 f(w_1) + \ldots + c_m f(w_m),\; g(p) = c_1 g(w_1) + \ldots + c_m g(w_m). \tag{13}$$

Функция переходов аффинного автомата по-прежнему определяется по формуле $F(v,p) = v\cdot f(p) + g(p)$. Таким образом, функция переходов будет аффинной по первому аргументу и линейной по второму.

5. ДОСТИЖИМОСТЬ СОСТОЯНИЙ В АФФИННЫХ АВТОМАТАХ

Пусть $A=(V, X, f, g)$ — n-мерный аффинный автомат, тогда множества мономиально достижимых состояний $\mathrm{Rm}_A(v)$ и $\mathrm{Rm}_i(v)$ определяются так же, как и для линейных автоматов. Далее, обозначим через $\mathrm{Ka}(X^*)$ множество аффинных полиномов, т.е. полиномов, у которых сумма коэффициентов равна единице. Множества полиномиальной достижимости определим также по аналогии с линейным случаем $\mathrm{Rp}_A(v)=\{F(v,p)\mid p\in \mathrm{Ka}(X^*)\}$ и $\mathrm{Rp}_i(v)=\{F(v,p)\mid d(p)\le i,$ $p\in \mathrm{Ka}(X^*)\}$, $i\ge 0$.

Нетрудно видеть, что множества достижимости $\mathrm{Rp}_i(v)$ — аффинные подпространства, поскольку из свойства (13) следует, что для аффинных полиномов выполняется свойство (6). Значит, $\mathrm{Rp}_i(v)$ — аффинная оболочка множества $\mathrm{Rm}_i(v)$, т.е. $\mathrm{Rp}_i(v)=\mathrm{aff}(\mathrm{Rm}_i(v))$. Напомним, что аффинная оболочка — это замыкание подмножества до наименьшего аффинного подпространства, содержащего это подмножество [4]. Тогда будем иметь следующую возрастающую последовательность аффинных подпространств:

$$\mathrm{Rp}_0(v)\subseteq \mathrm{Rp}_1(v)\subseteq \ldots\subseteq \mathrm{Rp}_n(v)\subseteq \ldots. \tag{14}$$

Длина этого ряда до стабилизации может быть на единицу больше, чем длина ряда (8), поскольку $\mathrm{Rp}_0(v)=\{v\}$ — нульмерное аффинное подпространство. Напомним, что размерность аффинного подпространства определяется размерностью соответствующего направляющего линейного подпространства [4]. Отсюда получаем аналог теоремы 2.

Теорема 4. Если в n-мерном аффинном автомате A выполняется условие $t\in \mathrm{Rp}_A(s)$, то существует аффинный полином p степени, не большей n, такой что $F(s,p)=t$.

Как уже оговаривалось, проблема мономиальной достижимости состояний в линейных и аффинных автоматах алгоритмически неразрешима. Покажем, например, алгоритмическую неразрешимость проблемы мономиальной достижимости состояний в двумерных аффинных автоматах над полем рациональных чисел, что будет некоторым усилением результата Патерсона [5]. Изложим сначала идею доказательства. Во-первых, в качестве известной алгоритмически неразрешимой проблемы, к которой будем сводить нашу проблему, возьмем хорошо известную проблему соответствия Поста (ПСП), в которой для двух словарных гомоморфизмов (морфизмов) $h_1\colon X^*\to Y^*$ и $h_2\colon X^*\to Y^*$, где Y — еще один конечный алфавит, требуется определить существование непустого слова w такого, что $h_1(w)=h_2(w)$ [9].

Во-вторых, одномерные аффинные автоматы могут вычислять словарные морфизмы, если вместо конкатенации строк использовать умножение чисел, которое фактически заменяется сдвигами. Действительно, без ограничения общности можно считать, что $Y\subseteq\{1,2,\ldots,9\}$, тогда строки в этом алфавите будем рассматривать как обычные десятичные числа и словарный морфизм $h\colon X^*\to Y^*$ можно вычислить с помощью следующего аффинного автомата: $F(v,x)=v\cdot 10^{l(h(x))}+h(x)$. Под «вычислением» здесь понимается то, что автомат, находясь вначале в нулевом состоянии, после обработки слова w окажется в состоянии $h(w)$. В-третьих, двумерный аффинный автомат, который является прямым произведением двух одномерных автоматов, может моделировать одновременную работу этих автоматов и одновременно вычислять два морфизма. Этого в принципе достаточно для неразрешимости, а оставшиеся технические детали приводятся в следующей теореме.

Теорема 5. Проблема мономиальной достижимости состояний в двумерных аффинных автоматах над полем рациональных чисел алгоритмически неразрешима.

Доказательство. Пусть заданы морфизмы $h_1\colon X^*\to Y^*$, $h_2\colon X^*\to Y^*$, где $Y=\{2,3\}$. Как известно, двух букв достаточно для неразрешимости проблемы Поста, а единица нам понадобится для других целей. Для каждого входного символа $x\in X$ добавим во входной алфавит автомата «зеркальный» символ $\bar{x}$ и, кроме того, еще один «терминальный» символ x_0. Рассмотрим двумерный аффинный автомат $A=(Q^2,\ X\cup\bar{X}\cup\{x_0\},\ f,g)$, где Q — поле рациональных чисел, который определяется по морфизмам h_1, h_2 следующим образом:

$$f(x)=\begin{bmatrix}10^{l(h_1(x))} & 0\\ 0 & 10^{l(h_2(x))}\end{bmatrix},\ f(\bar{x})=\begin{bmatrix}10^{l(h_1(x))} & 0\\ 0 & 10^{l(h_2(x))+1}\end{bmatrix},\ f(x_0)=\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}, \quad (15)$$

$$g(x)=(h_1(x),h_2(x)),\ g(\bar{x})=(h_1(x),1h_2(x)),\ g(x_0)=(0,0).$$

Начальным состоянием автомата будем считать вектор (1, 0), а финальным — нулевое состояние $(0,0)$. Пусть $L(a,b)=\{(ac,bc)\mid c\in Q\}$ — прямая, которая проходит через начало координат и точку (a,b). Из формул (15) ясно, что в нулевое состояние можно попасть только с помощью символа x_0, находясь на прямой $L(1,1)$. В любом другом случае под действием символа x_0 мы попадаем в ненулевую точку на прямой $L(1,-1)$, из которой нулевое состояние недостижимо.

Далее, первым входным символом должен быть один из символов $\bar{x}\in\bar{X}$, поскольку первая компонента состояния всегда начинается с единицы. После этого зеркальных символов больше быть не должно, поскольку тогда во второй компоненте появится вторая единица, в то время как в первой компоненте она останется только одна. Таким образом, для слова $\bar{x}w$, где $w\in X^*$, имеем $F((0,1),\bar{x}w)=$ $=(1h_1(xw),1h_2(xw))$. Отсюда получаем следующее условие:

$$F((0,1),\bar{x}wx_0)=(0,0)\leftrightarrow h_1(xw)=h_2(xw).$$

Значит, ПСП сводится к проблеме достижимости и, следовательно, проблема достижимости также неразрешима.

Теорема доказана.

Таким образом, статус проблемы мономиальной достижимости состояний в аффинных автоматах остается открытым только для одномерных аффинных автоматов. Попутно упомянем о проблеме мортальности, которая состоит в том, чтобы по линейному автомату определить существование нулевого слова, т.е. слова, которое переводит все состояния в нулевое состояние. Статус проблемы мортальности открыт только для двумерных линейных автоматов [6].

6. ЭКВИВАЛЕНТНОСТЬ СОСТОЯНИЙ

Определим сначала автоматы с выходом. (Полным) линейным автоматом с выходом над полем K называется пятерка объектов $A=(V,X,Y,f,h)$, где (V,X,f) — линейный автомат без выхода, $Y=\{b_1,\dots,b_l\}$ — конечное множество выходных символов и h — отображение, которое каждому входному символу x сопоставляет $n\times l$ матрицу с элементами из K.

Отображение h определяет функцию выходов $H(v,x)=v\cdot h(x)$. Отметим, что эта функция принимает значения в линейном пространстве $K(Y)$, т.е. в пространстве формальных линейных комбинаций выходных символов с коэффициентами из K. Функция выходов естественным образом продолжается на входные слова (мономы) $w=x_1\dots x_m$, а именно:

$$H(v,w)=H(v,x_1)H(v_1,x_2)\dots H(v_{m-1},x_m), \quad (16)$$

где $v_1,\dots,v_{m-1}$ — состояния, через которые проходит автомат под действием слова w. Правую часть равенства (16) можно понимать как слово в потенциально бесконечном алфавите $K(Y)$. Функцию выходов при необходимости можно продолжить на входные полиномы, но ее будем рассматривать как функцию вида $H: V\times X^*\to K(Y)^*$.

Определение 4. Два состояния s и t линейного автомата A называются эквивалентными (неотличимыми), если $H(s,w)=H(t,w)$ для всех $w\in X^*$.

Эквивалентные состояния обозначим $s\sim t$. Заметим, что если $s\sim t$, то $H(s-t,w)=0$ для всех $w\in X^*$, другими словами, состояние $s-t$ неотличимо от нулевого состояния. Это очевидное замечание, как ни странно, намного облегчает все последующие доказательства.

Пусть $A=(V,X,Y,f,h)$ — n-мерный линейный автомат, обозначим $\mathrm{Kr}\,h(w)$ подпространство состояний, которые неотличимы от нуля, словом w. Заметим, что в силу линейности функция выходов в нулевом состоянии должна быть тождественно равна нулю. Далее, пусть V_i — подпространство состояний, которые неотличимы от нуля словами (мономами) длины i, тогда получаем следующую убывающую цепочку включений:

$$V = V_0 \supseteq V_1 \supseteq \ldots \supseteq V_n \supseteq \{0\}. \qquad (17)$$

Строгие неравенства в цепочке (17) могут продолжаться не далее n-го места, следовательно, найдется такое число i, $0 \le i \le n$, что $V_i = V_{i+1}$. После этого цепочка (17) должна стабилизироваться, поскольку $F(V_i, x) \subseteq V_i$ для любого $x \in X$.

Таким образом, подпространство V_n будет подавтоматом (подмодулем), который обозначим E, тогда в E попадут состояния, эквивалентные нулю. Линейный автомат будет приведенным, т.е. не будет иметь эквивалентных состояний, если $E = \{0\}$. Если считать E подмодулем, то приведение автомата сводится к стандартной процедуре построения фактор-модуля V / E по подмодулю E. Из (17) получаем следующее предложение [1].

Теорема 6 (Мучник). Если два состояния неэквивалентны в n-мерном линейном автомате, то их можно различить словом длины n.

Известная теорема Мура для конечных автоматов [10] следует из теоремы 6. Действительно, если дополнить стандартное линейное расширение $K(A)$ конечного автомата A функцией выхода, то по следствию 1 и теореме 6 для неэквивалентных состояний s и t состояние $s - t$ будет отличимым от нуля в подавтомате пар словом длины $n-1$. Отсюда получаем следующее утверждение.

Следствие 3 (Мур). Если два состояния неэквивалентны в конечном автомате с n состояниями, то их можно различить словом длины $n-1$.

Аналогичным образом из теоремы 6, пользуясь следствием 2, можно вывести теорему Карлайла для вероятностных автоматов [11].

Следствие 4 (Карлайл). Если два состояния (распределения вероятностей) неэквивалентны в вероятностном автомате с n состояниями, то их можно различить словом длины $n-1$.

Отметим, что до стабилизации размерность подпространств ряда (17) должна убывать, откуда получаем следующие утверждения.

Следствие 5. Для всех подпространств V_i ряда (17) таких, что $V_i \ne E$, выполняется условие $\dim(V_i) \le n - i$.

Следствие 6. Если $U \not\subset E$, то $U \not\subset V_{n-d+1}$, где $d = \dim(U)$.

Действительно, если предположить, что $U \subseteq V_{n-d+1}$, то из следствия 5 получаем противоречие $d = \dim(U) \le \dim(V_{n-d+1}) \le d - 1$.

7. УСТАНОВОЧНЫЕ И ДИАГНОСТИЧЕСКИЕ СЛОВА

Рассмотрим установочные и диагностические эксперименты с линейными автоматами. Пусть задан полный линейный автомат $A = (V, X, Y, f, h)$, тогда слово $w \in X^*$ называется установочным, если оно позволяет с точностью до эквивалентности установить заключительное состояние автомата по реакции на это слово, т.е. выполняется следующее условие для всех состояний s и t:

$$H(s, w) = H(t, w) \to F(s, w) \sim F(t, w).$$

Для линейных автоматов это условие запишем следующим образом:

$$F(\mathrm{Kr}\, h(w), w) \subseteq E.$$

В дальнейшем нам понадобится равенство, которое можно проверить непосредственно для любых слов w_1 и w_2:

$$F(\mathrm{Kr}\, h(w_1 w_2), w_1 w_2) = F(F(\mathrm{Kr}\, h(w_1), w_1) \cap \mathrm{Kr}\, h(w_2), w_2). \qquad (18)$$

Теорема 7 (Мучник). Для всякого линейного n-мерного автомата существует установочное слово длины, не большей $n(n+1)/2$.

Доказательство. Пусть $A=(V, X, Y, f, h)$ — линейный n-мерный автомат, тогда положим $U_1=V$, $w_0=e$, $d_1=\dim(U_1)$, $i=1$, и будем действовать следующим образом. Если $U_i \subseteq E$, то построение заканчивается и его результатом будет слово w_{i-1}. В противном случае согласно следствию 6 находим слово y_i длины, не большей $n-d_i+1$, такое что $U_i \not\subset \operatorname{Kr} h(y_i)$. Положим, $w_i=w_{i-1} y_i$, $U_{i+1}= =F(\operatorname{Kr} h(w_i), w_i)$, $d_{i+1}=\dim(U_{i+1})$ и заметим, что из свойств (4) и (18) будем иметь $d_{i+1} \leq \dim(U_i \cap \operatorname{Kr} h(y_i))<d_i$. После этого повторяем цикл, увеличивая индекс i на единицу, и т.д.

В результате построения получим установочное слово $w_m=y_1 \ldots y_m$, где $l(y_i) \leq n-d_i+1$, $1 \leq i \leq m$, и размерности подпространств d_i образуют убывающую последовательность $n=d_1>d_2>\ldots>d_m \geq 1$. Отсюда получаем следующую оценку на длину слова w_m:

$$l(w_m)=l(y_1 \ldots y_m) \leq \sum_{i=1}^m(n-d_i+1) \leq \sum_{i=1}^n(n-i+1)=\frac{n(n+1)}{2}.$$

Таким образом, теорема доказана.

Из этой теоремы непосредственно вытекает известный результат Хиббарда для конечных автоматов [12], однако при этом надо учитывать, что линейное расширение приведенного конечного автомата может не быть приведенным как линейный автомат.

Слово $w \in X^*$ называется диагностическим для линейного автомата A, если по реакции на это слово можно определить начальное состояние автомата, т.е. выполняется следующее условие для всех состояний s и t:

$$H(s, w)=H(t, w) \rightarrow s=t.$$

Для линейных автоматов это условие эквивалентно условию $\operatorname{Kr} h(w)=0$.

В отличие от установочных слов, в автомате может не быть диагностических слов, поэтому автомат называется диагностируемым, если у него есть диагностическое слово. Возможное отсутствие диагностических слов в приведенном конечном автомате Мур назвал дискретным аналогом принципа неопределенности Гейзенберга [10, 11]. Следующий результат доказан в [13].

Теорема 8. В диагностируемом линейном n-мерном автомате существует диагностическое слово длины, не большей $n(n+1)/2$.

Доказательство этой теоремы проводится стандартным редукционным методом, который описан в предыдущей теореме. Единственное отличие состоит в том, что для редукции в этой теореме нужно использовать допустимые слова, т.е. слова, для которых выполняется условие $\operatorname{Kr} f(w) \cap \operatorname{Kr} h(w)=0$.

Может показаться, что теорема 8 противоречит известным результатам об экспоненциальной длине диагностических слов для конечных автоматов [14]. Но дело в том, что линейное расширение диагностируемого конечного автомата может не быть диагностируемым как линейный автомат, поэтому напрямую теорему 8 нельзя применять к конечным автоматам.

8. ПРИМЕНЕНИЯ В МАТЕМАТИЧЕСКОЙ ЭКОНОМИКЕ

Известно, что в экономике широко применяются линейные математические модели. Однако многие из них носят стационарный характер, что, конечно, ограничивает их практическую ценность. Как правило, не указывается, как соотносятся между собой переменные состояния в различные моменты времени. Попытка же ввести «динамику» в эти модели фактически приводит к появлению дискретных линейных систем и линейных автоматов.

Покажем это на примере известной транспортной задачи [15]. Пусть имеется m источников, производящих некоторую продукцию, которая потребляется n клиентами. Объем производства в i-м источнике в единицу времени равен a_i, $1 \leq i \leq m$, а объем потребления j-м клиентом в единицу времени равен b_j, $1 \leq j \leq n$. Известны стоимости c_{ij} перевозки единицы продукции из i-го пункта в j-й, которые характеризуются матрицей $C=\{c_{ij}\}$. Требуется составить план перевозок $U=\{u_{ij} \geq 0\}$ таким образом, чтобы минимизировать транспортные издержки при следующих условиях:

$$\sum_{j=1}^{n} u_{ij} = a_i, \ 1 \le i \le m, \ \sum_{i=1}^{m} u_{ij} = b_j, \ 1 \le j \le n. \tag{19}$$

Транспортные издержки характеризуются величиной $\langle C, U \rangle = \sum_{i,j} c_{ij} u_{ij}$.

Переведем транспортную задачу на язык аффинных автоматов. Пусть R — поле вещественных чисел, тогда положим $V = R^{m+n}$ и будем считать начальным состоянием автомата единичное состояние $e = (1, \dots, 1)$. Положим $X = \{e_{ij} \mid 1 \le i \le m, 1 \le j \le n\}$, где e_{ij} — элементарная матрица, у которой на месте (i, j) расположена единица, а на остальных местах — нули. Тогда X будет стандартным базисом пространства $m \times n$ матриц и план перевозок $U = \{u_{ij}\}$ может рассматриваться как обобщенный вход автомата, т.е. как элемент линейного пространства $R(X)$. Далее, пусть $d = a \oplus (-b)$ — вектор, состоящий из вектора $a = (a_1, \dots, a_m)$, к которому приписан вектор $-b = (-b_1, \dots, -b_n)$. Наконец, положим $f(U) = (-U^{\mathrm{T}}) \oplus U$, где U^{T} — транспонированная для U матрица, а символ суммы означает прямую сумму матриц. Очевидно, что f будет линейным отображением, которое переводит $m \times n$ матрицы в квадратные $(m+n) \times (m+n)$ матрицы. Тогда условия (19) можно записать так:

$$e \cdot f(U) + d = 0. \tag{20}$$

Таким образом, получен аффинный автомат $A = (V, X, f, g)$, где $g(a_0) = d$ и $g(U) = 0$ для всех $U \in R(X)$. Требуется найти план перевозок $U \ge 0$ (управляющий вектор), который за единицу времени переводит автомат наиболее экономным способом из единичного состояния в нулевое. Экономность перехода определяется скалярным произведением $\langle C, U \rangle$, где $m \times n$ матрицы C и U рассматриваются как векторы. Если равенство (20) умножить справа на единичный вектор-столбец e^{T}, то получим условие $d \cdot e^{\mathrm{T}} = 0$, которое, как известно, является необходимым и достаточным для разрешимости транспортной задачи в замкнутой форме. Конечно, такая формулировка не дает решения исходной оптимизационной задачи, но позволяет легко изменять модель, придавая ей необходимую гибкость и динамику. Во-первых, можно перейти от полного двудольного графа перевозок, который неявно фигурирует в транспортной задаче, к произвольным транспортным сетям (транзитная задача). Пусть граф перевозок имеет m ориентированных дуг, n вершин и матрицу инцидентности G (размерности $m \times n$), в которой в каждой строке находится -1 на месте, откуда исходит дуга, и $+1$ на месте, куда она заходит, а на остальных позициях стоят нули. Далее, пусть производство и потребление в каждой вершине характеризуется вектором $d = (d_1, \dots, d_n)$ (потребление — это отрицательное производство), а вектор $u = (u_1, \dots, u_m)$ характеризует загрузку дуг графа. Тогда уравнение баланса приобретает следующий вид [16]:

$$u \cdot G + d = 0. \tag{21}$$

Другими словами, выходной поток продукта из каждой вершины равен входному потоку в эту вершину плюс производство. Таким образом, получаем другой аффинный автомат, в котором d — начальное состояние, u — входной вектор, $f(a_0) = I$, $f(u) = 0$, $g(u) = u \cdot G$ для всех $u \in R(X)$. Требуется найти план перевозок $u \ge 0$, который переводит автомат наиболее экономным способом из начального состояния в нулевое. Экономность характеризуется скалярным произведением $\langle c, u \rangle$ векторов c и u размерности m. Если равенство (21) умножить справа на единичный вектор-столбец e^{T}, то в силу специфики матрицы G снова получим условие $d \cdot e^{\mathrm{T}} = 0$, которое здесь будет необходимым, но, вообще говоря, недостаточным для существования плана.

Теперь рассмотрим динамическую модель. Введем вектор состояния $v(k)$ длины n, отражающий запасы продукции в k-й период времени у производителей и потребителей. Тогда уровень запасов в период времени $k+1$ определяется алгебраической суммой пополнения и истощения запасов [17]:

$$v(k+1)=v(k)+u(k)\cdot G+d,\ 1\le k\le N. \qquad (22)$$

Снова получаем аффинный автомат $A=(V, X, f, g)$, где $f(a_0)=I$, $f(u)=0$, $g(u)=u\cdot G$, $g(a_0)=d$ для всех $u\in R(X)$. В процессе изменения структуры автомата усложняется и целевая функция, в которой можно учесть затраты на хранения продукции, характеризуемые вектором $y(k)$, отразить влияние среды, сделав переменным вектор стоимостей $c(k)$, и т.д. Кроме того, целевую функцию теперь можно вычислять на всем периоде моделирования:

$$J=\sum_{k=1}^{N}(y(k)\cdot v(k)+c(k)\cdot u(k)).$$

Точкой здесь обозначено скалярное произведение векторов. Таким образом, транспортная задача превратилась в задачу управления запасами на произвольном графе. Нужно найти динамический план перевозок (входное слово $p=u(1)\ldots u(N)$), который минимизирует запасы наиболее экономным способом. Конечно, решение теперь усложняется и его нужно искать в виде многошагового процесса принятия решений в духе динамического программирования Р. Беллмана [18].

Таким же образом можно ввести динамику в известную модель Леонтьева, а также в ряд других экономических моделей [17].

ЗАКЛЮЧЕНИЕ

Конечно, в одной работе трудно рассмотреть все проблемы, относящиеся к этой тематике. Остались не рассмотренными такие вопросы, как исследование управляемости и наблюдаемости, структурная и схемная реализация линейных автоматов и т.д. Но уже из изложенного ясно, что обобщенные линейные автоматы являются «мостом» между теорией автоматов и теорией дискретных линейных систем и поэтому приобретают значение в рамках всей математической теории систем.

СПИСОК ЛИТЕРАТУРЫ

1. Мучник А. Общие линейные автоматы // Проблемы кибернетики. — 1970. — **23.** — С. 171–208.
2. Ленг С. Алгебра. — М.: Мир, 1968. — 564 с.
3. Калман Р., Фалб П., Арбиб М. Очерки по математической теории систем. — М.: Мир, 1971. — 400 с.
4. Кострикин А., Манин Ю. Линейная алгебра и геометрия. — М.: Наука, 1986. — 303 с.
5. Paterson M. Unsolvability in 3×3 matrices. // Studies in Appl. Mathemat. — 1970. — **49.** — P. 105–107.
6. Рысцов И.К. Проблема мортальности и аффинные автоматы // Кибернетика и системный анализ. — 2008. — № 2. — С. 24–29.
7. Гилл А. Линейные последовательностные машины. — М.: Наука, 1974. — 287 с.
8. Заде Л., Дезоер Ч. Теория линейных систем. — М.: Наука, 1970. — 703 с.
9. Саломаа А. Жемчужины теории формальных языков. — М.: Мир, 1986. — 159 с.
10. Мур Э. Умозрительные эксперименты с последовательностными машинами // Автоматы. — М.: Изд-во иностр. лит., 1956. — С. 179–210.
11. Карлайл Е. Приведенные формы для стохастических последовательностных машин // Кибернетический сборник. — 1966. — № 3. — С. 101–111.
12. Хиббард Т. Точные верхние границы длин минимальных экспериментов, определяющих заключительное состояние последовательностных машин // Там же. — 1966. — № 2. — С. 7–23.
13. Рысцов И. Оценка длины диагностического слова для общих линейных автоматов // Исследования по алгебре. — 1974. — № 4. — С. 95–101.
14. Рысцов И. Об асимптотической оценке длины диагностического слова для конечного автомата // Кибернетика. — 1980. — № 2. — С. 31–35.
15. Гольштейн Е., Юдин Д. Задачи линейного программирования транспортного типа. — М.: Наука, 1969. — 382 с.
16. Романовский И. Алгоритмы решения экстремальных задач. — М.: Наука, 1977. — 352 с.
17. Негойцэ К. Применение теории систем к проблемам управления. — М.: Мир, 1981. — 180 с.
18. Беллман Р., Калаба Р. Динамическое программирование и современная теория управления. — М.: Наука, 1969. — 118 с.

Поступила 22.07.2008