Теория вероятностных автоматов
Часть 1

А.М.Миронов

Оглавление

Введение

Понятие **вероятностного автомата (ВА)** впервые было сформулировано в 1963 г. в основополагающей работе М. Рабина [1]. Данное понятие возникло как синтез понятий конечного детерминированного автомата [2] и цепи Маркова [3] и было предназначено для построения математических моделей динамических систем, в которых присутствует неопределённость, описываемая статистическими закономерностями. Эта неопределённость связана:

- с неточностью знаний о состояниях, в которых моделируемые системы находятся в процессе своего функционирования, и

- с недетерминированностью правил изменения этих состояний.

Неопределённость в ВА может быть вызвана различными причинами, которые подразделяются на два класса.

1. Причины из первого класса связаны с природой системы, моделируемой вероятностным автоматом. К ним относятся:

    - влияние случайных факторов на функционирование системы, например: случайные сбои компонентов системы или отказы в их работе, случайное изменение условий функционирования анализируемой системы, случайность потока заявок в системе массового обслуживания и т. п.;

    - несовершенство (или невозможность) точного измерения состояний этой системы.

2. Второй класс причин связан с преднамеренным внесением неточности и неопределённости в математические модели анализируемых систем. Это делается в тех случаях, когда точные модели анализируемых систем имеют неприемлемо высокую сложность и проведение анализа поведения таких систем возможно только с использованием их упрощённых моделей, в которых некоторые компоненты состояний этих систем игнорируются. В частности, анализ поведения сложной программной системы (например, операционной системы компьютера) в большинстве случаев возможен только с использованием таких упрощённых математических моделей этих систем, в которых принимаются во внимание значения лишь некоторых программных переменных, от которых существенно зависит поведение анализируемой программной системы.

Как правило, моделирование систем при помощи ВА производится:

- либо с целью анализа свойств этих систем (к числу которых относятся, например, корректность, безопасность, надёжность, устойчивость функционирования в непредусмотренных ситуациях и т. д.),
- либо с целью вычисления различных количественных характеристик анализируемых систем, среди которых могут быть, например, следующие:
  - частота выполнения тех или иных действий или переходов в анализируемых системах,
  - вероятность отказа компонентов анализируемых систем,
  - вероятность вторжения злоумышленника в компьютерную сеть,
  - математическое ожидание времени отклика веб-сервиса.

Первоначальное понятие ВА, введённое в работе М. Рабина [1], было предназначено главным образом для изучения вопросов представимости регулярных языков вероятностными автоматами. Затем оно было обобщено до такого понятия, которое позволило моделировать вероятностные преобразователи информации. Определение ВА в общей форме было введено независимо в работах Дж. Карлайла [4], Р. Г. Бухараева [5] и П. Штарке [6].

С начала возникновения понятия ВА исследовательская деятельность в этой области отличалась высокой активностью. Результаты первых лет исследований в области ВА были систематизированы в книге [7]. Подробный список (около 500) ссылок на работы с наиболее существенными теоретическими и практическими результатами по ВА, полученными до 1985 г., можно найти в фундаментальной монографии Р. Г. Бухараева [8], которую можно рассматривать как итог первого периода развития теории ВА, продолжавшегося более двух десятилетий.

В последующие годы произошло некоторое снижение активности исследований в этой области, но в настоящее время теория ВА вновь находится в состоянии подъёма. Возрождение исследовательской активности в области ВА в значительной степени связано с тем, что в связи с бурным развитием современных информационных технологий возник широкий круг новых задач, в решении которых ВА могут служить эффективным инструментом. К числу таких задач относятся задачи в следующих областях:

- верификация программ и протоколов передачи данных в компьютерных сетях,
- информационный поиск в Интернете,
- финансово-экономический анализ,
- обработка и извлечение знаний из больших массивов данных (data mining и process mining), в частности, в задачах анализа бизнес-процессов, биоинженерии и биоинформатики,
- извлечение смысла из текстов на естественных языках,
- машинное зрение и обработка изображений и др.

Началом современного этапа развития теории ВА можно считать работу [9], в которой рассмотрены ВА, возникающие при моделировании параллельных вычислительных систем с асинхронным взаимодействием. В качестве вводных текстов в современную теорию ВА можно назвать работы [10] и [11].

Главное отличие нового понятия ВА от того, которое изучалось в предшествующий период, заключается в том, что в новом понимании ВА определяется как **система переходов** (**transition system**), с которой связано некоторое множество переменных. ВА функционирует путём выполнения переходов, после каждого из которых происходит обновление значений переменных этого ВА. Можно доказать, что если множество переменных ВА конечно и множества значений этих переменных тоже конечны, то новое и старое понятия ВА будут эквивалентны.

Наряду с упомянутыми выше понятиями ВА существуют и другие модели динамических систем со случайным поведением, например скрытые марковские модели (hidden Markov models) [12], байесовские сети (Bayesian networks) [13], вероятностные графические модели [14], марковские решающие процессы (Markov decision processes) [15], вероятностные I/O автоматы (probabilistic I/O automata) [16]. Все эти модели являются частными случаями исходного понятия ВА общего вида [8].

Наряду с перечисленными выше моделями в последние годы изучаются модели динамических систем со случайным поведением, переходы в которых могут быть ассоциированы не только с вероятностями их выполнения, но и с модальностями must и may, которые позволяют существенно усилить выразительные возможности этих моделей по сравнению с другими упомянутыми выше моделями. Основные концепции и методы, относящиеся к таким моделям, содержатся в статье [17].

Также изучаются и другие обобщения понятия ВА, в частности вероятностные сети Петри [18], [19], ВА с непрерывным временем [20], вероятностные процессные алгебры [21].

Глава 1

Вспомогательные понятия

1.1 Случайные функции

1.1.1 Понятие случайной функции

Пусть задана пара множеств X, Y.

Случайной функцией (СФ) из X в Y называется произвольная функция f вида

$$f : X \times Y \to [0,1], \tag{1.1}$$

удовлетворяющая условиям:

- $\forall\, x \in X$ множество $\{y \in Y \mid f(x,y) > 0\}$ конечно или счётно,
- $\forall\, x \in X \quad \sum\limits_{y \in Y} f(x,y) = 1$.

Для любых $x \in X$ и $y \in Y$ значение $f(x,y)$ можно интерпретировать как вероятность того, что СФ f отображает x в y.

Если f – СФ из X в Y, то мы будем обозначать этот факт записью $f : X \underset{r}{\to} Y$. Мы будем называть X **областью определения** СФ f, а Y – **областью значений** СФ f.

Если f и g – СФ вида $f : X \underset{r}{\to} Y, \quad g : Y \underset{r}{\to} Z$ то их **композицией** называется СФ $fg : X \underset{r}{\to} Z$, определяемая следующим образом:

$$\forall\, x \in X,\ \forall\, z \in Z \quad (fg)(x,z) \overset{\text{def}}{=} \sum_{y \in Y} f(x,y)g(y,z) \tag{1.2}$$

СФ (1.1) называется **детерминированной**, если для каждого $x \in X$ существует единственный $y \in Y$, такой, что $f(x, y) = 1$. Если f – детерминированная СФ вида (1.1), и x, y – такие элементы X и Y соответственно, что $f(x, y) = 1$, то мы будем говорить, что f **отображает** x **в** y.

1.1.2 Матрицы, соответствующие конечным случайным функциям

СФ называется **конечной (КСФ)**, если её область определения и область значений являются конечными множествами.

Пусть задана КСФ $f : X \underset{r}{\to} Y$, и на X и Y заданы упорядочения их элементов, которые имеют вид $(x_1, \ldots .x_m)$ и $(y_1, \ldots, y_n)$ соответственно. Тогда f можно представить в виде матрицы (обозначаемой тем же символом f)

$$f = \begin{pmatrix} f(x_1, y_1) & \ldots & f(x_1, y_n) \\ \ldots & \ldots & \ldots \\ f(x_m, y_1) & \ldots & f(x_m, y_n) \end{pmatrix}. \tag{1.3}$$

Ниже мы будем отождествлять каждую КСФ f с соответствующей ей матрицей (1.3).

Мы будем предполагать, что для каждого множества X, являющегося областью определения или областью значений какой-либо из рассматриваемых КСФ, на X задано фиксированное упорядочение его элементов. Таким образом, для каждой рассматриваемой КСФ соответствующая ей матрица определена однозначно.

Согласно определению произведения матриц, из (1.2) следует, что матрица fg является произведением матриц f и g.

1.1.3 Вероятностные распределения

Вероятностным распределением (или просто **распределением**) на множестве X называется произвольная СФ ξ вида

$$\xi : \mathbf{1} \underset{r}{\to} X$$

где **1** – множество, состоящее из одного элемента, который мы будем обозначать символом e. Совокупность всех распределений на X мы будем обозначать записью $X^{\triangle}$. Для каждого $x \in X$ и каждого $\xi \in X^{\triangle}$ значение $\xi(e,x)$ мы будем обозначать более коротко записью x^{ξ}. Для каждого $x \in X$ мы будем обозначать записью ξ_x распределение из $X^{\triangle}$, определяемое следующим образом: $\forall\, y \in X \;\; y^{\xi_x} \overset{\text{def}}{=} 1$, если $y = x$, и $y^{\xi_x} \overset{\text{def}}{=} 0$, если $y \neq x$.

1.2 Строки и функции на строках

1.2.1 Строки и связанные с ними понятия

Для каждого множества X мы будем обозначать записью X^* совокупность всех конечных строк, компонентами которых являются элементы X. Множество X^* содержит **пустую строку**, она обозначается символом ε.

Для каждого $x \in X$ строка, состоящая из одного этого элемента, обозначается той же записью x.

Для каждой строки $u \in X^*$ её **длиной** называется количество компонентов этой строки. Длина пустой строки равна нулю. Длина строки u обозначается записью $|u|$.

Для каждого целого числа $k \geq 0$ записи X^k, $X^{\leq k}$, $X^{<k}$, $X^{\geq k}$, $X^{>k}$, обозначают совокупности всех строк из X^*, длина которых равна k, меньше или равна k, и т.д., соответственно.

Для каждой пары строк $u, v \in X^*$ их **конкатенацией** называется строка, обозначаемая записью uv, и определяемая следующим образом:

- $u\varepsilon \overset{\text{def}}{=} \varepsilon u \overset{\text{def}}{=} u$, и
- если $u = x_1 \ldots x_n$ и $v = x'_1 \ldots x'_m$, то $uv \overset{\text{def}}{=} x_1 \ldots x_n x'_1 \ldots x'_m$.

Для каждой строки $u \in X^*$ запись $\tilde{u}$ обозначает строку u, записанную в обратном порядке, т.е. $\tilde{\varepsilon} = \varepsilon$, и если $u = x_1 \ldots x_n$, то $\tilde{u} = x_n \ldots x_1$.

1.2.2 Функции на строках

Пусть задано конечное множество X.

Функцией на строках из X^* мы будем называть произвольную функцию вида $f : X^* \to \mathbf{R}$ (где символ $\mathbf{R}$ обозначает множество действительных чисел). Совокупность всех функций на строках из X^* мы будем обозначать записью $\mathbf{R}^{X^*}$.

На множестве $\mathbf{R}^{X^*}$ определены следующие операции.

- Для функций $f_1, f_2 \in \mathbf{R}^{X^*}$ их **сумма** $f_1 + f_2$ и **разность** $f_1 - f_2$ определяются следующим образом:

$$\forall\, u \in X^* \begin{cases} (f_1 + f_2)(u) \overset{\text{def}}{=} f_1(u) + f_2(u), \\ (f_1 - f_2)(u) \overset{\text{def}}{=} f_1(u) - f_2(u). \end{cases} \tag{1.4}$$

- Для каждого $a \in \mathbf{R}$ и каждой функции $f \in \mathbf{R}^{X^*}$ **произведение** af определяется следующим образом:

$$\forall\, u \in X^* \quad (af)(u) \overset{\text{def}}{=} af(u). \tag{1.5}$$

Множество $\mathbf{R}^{X^*}$ можно рассматривать как векторное пространство над $\mathbf{R}$ относительно определённых выше операций сложения и умножения на числа из $\mathbf{R}$.

1.3 Автоматы Мура

1.3.1 Понятие автомата Мура

Автомат Мура – это совокупность объектов

$$M = (X, Y, S, \delta, \lambda, s^0) \tag{1.6}$$

(называемая в этом параграфе просто **автоматом**), компоненты которой имеют следующий смысл:

- X, Y, S – множества, элементы которых называются соответственно **входными сигналами**, **выходными сигналами**, и **состояниями** автомата M,
- $\delta : S \times X \to S$ и $\lambda : S \to Y$ – отображения, называемые соответственно **отображением перехода** и **отображением выхода** автомата M,

- s^0 – элемент S, называемый **начальным состоянием** автомата M.

Автомат является моделью динамической системы, работа которой происходит в дискретном времени и заключается в

- изменении состояний под воздействием входных сигналов, поступающих на её вход, и
- выдаче в каждый момент времени $t = 0, 1, \dots$ некоторого выходного сигнала.

Функционирование автомата M вида (1.6) происходит следующим образом. В каждый момент времени $t = 0, 1, \dots$ автомат M находится в некотором состоянии $s(t)$, причем $s(0) \stackrel{\text{def}}{=} s^0$. В каждый момент времени t автомат M

- получает входной сигнал $x(t) \in X$,
- переходит в состояние $s(t+1) \stackrel{\text{def}}{=} \delta(s(t), x(t))$, и
- выдаёт выходной сигнал $y(t) \stackrel{\text{def}}{=} \lambda(s(t))$.

1.3.2 Достижимые состояния и реакция автомата

Пусть M – автомат вида (1.6). Для каждого $s \in S$ и каждой строки $u \in X^*$ запись su обозначает состояние, определяемое индуктивно следующим образом: $s\varepsilon \stackrel{\text{def}}{=} s$, и если $u = vx$, где $v \in X^*$ и $x \in X$, то $su \stackrel{\text{def}}{=} \delta(sv, x)$. Нетрудно видеть, что если строка u имеет вид $x_0 \dots x_n$ то su – это состояние, в которое перейдёт M через $n + 1$ тактов времени, при условии, что

- в текущий момент времени t он находился в состоянии s, и
- в моменты $t, t + 1, \dots, t + n$ на вход M подавались сигналы $x_0, x_1, \dots, x_n$ соответственно.

Состояние $s \in S$ называется **достижимым**, если оно имеет вид $s^0 u$ для некоторого $u \in X^*$.

Реакция автомата M – это отображение $f_M : X^* \to Y$, определяемое следующим образом:

$$\forall\, u \in X^* \quad f_M(u) \stackrel{\text{def}}{=} \lambda(s^0 u).$$

Нетрудно видеть, что если строка $u \in X^*$ имеет вид $x_0 \dots x_n$, то $f_M(u)$ – это выходной сигнал, который выдает M в момент $n+1$, если в моменты $0, 1, \dots, n$ на его вход подавались сигналы $x_0, x_1, \dots, x_n$ соответственно.

Автоматы называются **эквивалентными**, если их реакции совпадают.

1.3.3 Достижимая часть автомата

Пусть M – автомат вида (1.6). Обозначим

- символом S' множество всех достижимых состояний M, и
- символом M' автомат, получаемый из M заменой S на S', и отображений δ и λ на ограничения этих отображений на подмножества $S' \times X$ и S' соответственно.
  (нетрудно видеть, что $\forall\, s \in S'$, $\forall\, x \in X \;\; \delta(s, x) \in S'$)

Автомат M' называется **достижимой частью** автомата M. Очевидно, что M и M' эквивалентны.

Если X и S конечны, то S' может быть найдено следующим образом: определим последовательность $S_0 \subseteq S_1 \subseteq S_2 \subseteq \dots$, где

- $S_0 \stackrel{\text{def}}{=} \{s^0\}$,
- $\forall\, i \geq 0 \quad S_{i+1} \stackrel{\text{def}}{=} S_i \cup \{sx \mid s \in S_i,\ x \in X\}$.

Т.к. все члены последовательности $S_0, S_1, \dots$ – подмножества конечного множества S, то $\exists\, k < |S| : S_k = S_{k+1}$. Нетрудно видеть, что $S_k = S'$.

1.4 Линейные автоматы

Пусть заданы конечное множество X и натуральное число n.

Линейным автоматом (ЛА) размерности n над X мы будем называть тройку L вида

$$L = (\xi^0, \{L^x \mid x \in X\}, \lambda), \tag{1.7}$$

где

- ξ^0 – вектор-строка размерности n над $\mathbf{R}$,
- $\forall\, x \in X \;\; L^x$ – квадратная матрица размерности n над $\mathbf{R}$, и
- λ – вектор-столбец размерности n над $\mathbf{R}$.

Для каждого ЛА L мы будем обозначать записью $\dim L$ размерность этого ЛА.

ЛА (1.7) определяет автомат Мура, обозначаемый тем же символом L,

- множествами входных и выходных сигналов которого являются X и $\mathbf{R}$ соответственно,
- множеством состояний которого является совокупность $\mathbf{R}^n$ всех вектор–строк размерности n над $\mathbf{R}$,
- начальным состоянием – вектор-строка ξ^0,
- отображение перехода сопоставляет паре $(\xi, x) \in \mathbf{R}^n \times X$ вектор-строку ξL^x, и
- отображение выхода сопоставляет состоянию $\xi \in \mathbf{R}^n$ число $\xi\lambda \in \mathbf{R}$.

Нетрудно видеть, что реакция f_L данного автомата сопоставляет каждой строке $u \in X^*$ число $\xi^0 L^u \lambda$, где $L^\varepsilon \stackrel{\text{def}}{=} E$ (единичная матрица размерности n), и если строка u имеет вид $x_1 \ldots x_k$, то $L^u \stackrel{\text{def}}{=} L^{x_1} \ldots L^{x_k}$.

Пусть f – функция из $\mathbf{R}^{X^*}$. Мы будем называть её **линейно-автоматной функцией (ЛАФ)**, если для некоторого ЛА L над X верно равенство $f = f_L$.

Глава 2

Вероятностные автоматы и вероятностные реакции

2.1 Вероятностные автоматы

2.1.1 Понятие вероятностного автомата

Вероятностный автомат (ВА) – это пятерка A вида

$$A = (X, Y, S, P, \xi^0) \tag{2.1}$$

компоненты которой имеют следующий смысл.

1. X, Y и S – конечные множества, элементы которых называются соответственно **входными сигналами**, **выходными сигналами** и **состояниями** ВА A.

2. P – СФ вида $P : S \times X \underset{r}{\to} S \times Y$, называемая **поведением** ВА A. $\forall (s, x, s', y) \in S \times X \times S \times Y$ значение $P(s, x, s', y)$ понимается как вероятность того, что

   - если в текущий момент времени (t) A находится в состоянии s, и в этот момент времени на его вход поступил сигнал x,
   - то в следующий момент времени $(t+1)$ A будет находиться в состоянии s', и в момент времени t выходной сигнал A равен y.

3. ξ^0 – распределение на S, называемое **начальным распределением** ВА A. $\forall\, s \in S$ значение s^{ξ^0} понимается как вероятность того, что в начальный момент времени ($t = 0$) ВА A находится в состоянии s.

ВА (2.1) называется **детерминированным**, если $\xi^0 = \xi_s$ для некоторого $s \in S$, и СФ P является детерминированной.

2.1.2 Матрицы, связанные с вероятностными автоматами

Пусть A – ВА вида (2.1), и упорядочение множества S его состояний имеет вид $(s_1, \ldots, s_n)$. Для любых $x \in X$ и $y \in Y$ мы будем обозначать записью A^{xy} матрицу порядка n

$$\begin{pmatrix} P(s_1, x, s_1, y) & \ldots & P(s_1, x, s_n, y) \\ \ldots & \ldots & \ldots \\ P(s_n, x, s_1, y) & \ldots & P(s_n, x, s_n, y) \end{pmatrix} \tag{2.2}$$

и для любой пары строк $u \in X^*, v \in Y^*$ мы будем обозначать записью $A^{u,v}$ (запятая в этой записи может опускаться) матрицу порядка n, определяемую следующим образом:

- $A^{\varepsilon,\varepsilon} = E$ (единичная матрица),
- если $|u| \neq |v|$, то $A^{u,v} = 0$ (нулевая матрица), и
- если $u = x_1 \ldots x_k$ и $v = y_1 \ldots y_k$, то $A^{u,v} = A^{x_1 y_1} \ldots A^{x_k y_k}$.

Пусть s – произвольное состояние из S, и в упорядочении элементов S данное состояние имеет номер i (т.е. $s = s_i$). Мы будем называть

- строку номер i матрицы $A^{u,v}$ – **строкой** s, и обозначать её записью $\vec{A}_s^{u,v}$
- столбец номер i матрицы $A^{u,v}$ – **столбцом** s, и обозначать его записью $A_s^{u,v\downarrow}$

Для любых $s, s' \in S$ мы будем обозначать записью $A^{u,v}_{s,s'}$ элемент матрицы $A^{u,v}$, находящийся в строке s столбце s'.

Если строки $u \in X^*$ и $v \in Y^*$ имеют вид $x_0 \dots x_k$ и $y_0 \dots y_k$ соответственно, то $A^{u,v}_{s,s'}$ можно понимать как вероятность того, что

- если в текущий момент (t) A находится в состоянии s, и, начиная с этого момента, на вход A последовательно поступали элементы строки u (т.е. в момент t поступил сигнал x_0, в момент $t+1$ поступил сигнал x_1, и т.д.)

- то в моменты $t, t+1, \dots, t+k$ выходные сигналы A равны $y_0, \dots, y_k$ соответственно, и в момент $t+k+1$ A будет находиться в состоянии s'.

2.1.3 Реакция вероятностного автомата

Пусть заданы ВА A вида (2.1) и распределение $\xi \in S^{\triangle}$.

Мы будем говорить, что **ВА A в момент времени t имеет распределение** ξ, если для каждого состояния $s \in S$ вероятность того, что A в момент времени t находится в состоянии s, равна s^{ξ}.

Реакцией ВА A в распределении ξ называется функция

$$A^{\xi} : X^* \times Y^* \to \mathbf{R}$$

определяемая следующим образом:

$$\forall\, u \in X^*, \forall\, v \in Y^* \quad A^{\xi}(u, v) \stackrel{\text{def}}{=} \xi A^{u,v} I$$

где запись I обозначает вектор-столбец порядка $|S|$, все компоненты которого равны 1.

Реакцией ВА A мы будем называть реакцию этого ВА в его начальном распределении. Мы будем обозначать реакцию ВА A записью f_A. Нетрудно доказать, что если ВА A детерминированный, то СФ f_A – детерминированная.

Если строки $u \in X^*$ и $v \in Y^*$ имеют вид $x_0 \dots x_k$ и $y_0 \dots y_k$ соответственно, то $f_A(u, v)$ можно понимать как вероятность того, что если, начиная с момента 0, на вход A последовательно поступали элементы строки u (т.е. в момент 0 поступил сигнал x_0, в момент 1 поступил сигнал x_1, и т.д.), то в моменты $0, 1, \ldots, k$ выходные сигналы A равны $y_0, \ldots, y_k$ соответственно.

Теорема 1. Если A – ВА вида (2.1) и $\xi \in S^{\triangle}$, то A^{ξ} – СФ.

Доказательство.

Поскольку $\forall\, u \in X^*, \forall\, v \in X^*$ значение $A^{\xi}(u, v)$ неотрицательно, то для доказательства теоремы достаточно доказать, что $\forall\, u \in X^* \quad \sum\limits_{v \in Y^*} A^{\xi}(u, v) = 1$, т.е.

$$\forall\, u \in X^* \qquad \sum_{v \in Y^*} \xi A^{u,v} I = 1. \tag{2.3}$$

Поскольку $A^{u,v} = 0$ при $|u| \neq |v|$, то (2.3) эквивалентно условию: $\forall\, k \geq 0$

$$\forall\, u \in X^k \qquad \sum_{v \in Y^k} \xi A^{u,v} I = 1. \tag{2.4}$$

Докажем (2.4) индукцией по k. Если $k = 0$, то (2.4) следует из того, что $A^{\varepsilon,\varepsilon} = E$ и $\xi E I = \xi I = 1$ (т.к. $\xi \in S^{\triangle}$).

Пусть (2.4) верно для некоторого k. Докажем, что

$$\forall\, u \in X^{k+1} \qquad \sum_{v \in Y^{k+1}} \xi A^{u,v} I = 1. \tag{2.5}$$

(2.5) эквивалентно соотношению

$$\forall\, u \in X^k,\ \forall\, x \in X \qquad \sum_{v \in Y^k,\ y \in Y} \xi A^{ux,vy} I = 1. \tag{2.6}$$

Т.к. $A^{ux,vy} = A^{u,v} A^{xy}$, то (2.6) можно переписать в виде

$$\forall\, u \in X^k,\ \forall\, x \in X \qquad \sum_{v \in Y^k} \xi A^{u,v} \Big(\sum_{y \in Y} A^{xy} I\Big) = 1. \tag{2.7}$$

(2.7) следует из (2.4) и равенства

$$\sum_{y \in Y} A^{xy} I = I \tag{2.8}$$

которое верно потому, что если A^{xy} имеет вид (2.2), то $\forall\, i \in \{1, \ldots, n\}$ элемент с индексом i столбца $\sum\limits_{y \in Y} A^{xy} I$ равен сумме

$$\sum_{y \in Y,\ j = 1, \ldots, n} P(s_i, x, s_j, y)$$

которая равна 1, т.к. P – СФ вида $P : S \times X \underset{r}{\to} S \times Y$. ■

Распределения $\xi_1, \xi_2 \in S^{\triangle}$ называются **эквивалентными относительно** A, если реакции A^{ξ_1} и A^{ξ_2} совпадают, т.е.

$$\forall\, u \in X^*, \forall\, v \in Y^* \qquad \xi_1 A^{u,v} I = \xi_2 A^{u,v} I.$$

Если распределения ξ_1 и ξ_2 эквивалентны относительно A, то мы будем обозначать этот факт записью $\xi_1 \underset{A}{\sim} \xi_2$.

2.1.4 Базисные матрицы вероятностных автоматов

Ниже мы будем использовать следующее обозначение: для каждого множества W элементов какого-либо линейного пространства мы будем обозначать записью $\langle W \rangle$ подпространство этого линейного пространства, порожденное векторами из W.

Пусть A – ВА вида (2.1). Обозначим записью AI совокупность всех вектор-столбцов вида $A^{u,v}I$, где $u \in X^*, v \in Y^*$.

Базисной матрицей ВА A называется матрица, обозначаемая записью $[A]$, и удовлетворяющая условиям:

- каждый столбец матрицы $[A]$ является элементом AI,
- столбцы матрицы $[A]$ образуют базис пространства $\langle AI \rangle$.

Нетрудно видеть, что для любых $\xi_1, \xi_2 \in S^{\triangle}$

$$\xi_1 \underset{A}{\sim} \xi_2 \quad \Leftrightarrow \quad \xi_1[A] = \xi_2[A].$$

Для каждого $s \in S$ мы будем называть **строкой** s матрицы $[A]$ ту её строку, которая содержит значения вида $\vec{A}_s^{u,v} I$. Мы будем обозначать эту строку записью $[A]_s$.

Матрица $[A]$ м.б. построена при помощи излагаемого ниже алгоритма.

Пусть $k \geq 0$. Обозначим записью AI_k совокупность вектор-столбцов вида $A^{u,v}I$, где $u \in X^*$, $v \in Y^*$, $|u| = |v| \leq k$. Нетрудно видеть, что

$$\langle AI_0 \rangle \subseteq \langle AI_1 \rangle \subseteq \langle AI_2 \rangle \subseteq \ldots \quad \text{и} \quad \bigcup_{k \geq 0} \langle AI_k \rangle = \langle AI \rangle. \qquad (2.9)$$

Поскольку все пространства $\langle AI_k \rangle$ являются подпространствами конечномерного линейного пространства (размерности $|S|$), то, следовательно, последовательность включений в (2.9) не может неограниченно возрастать, т.е. для некоторого k верны равенства

$$\langle AI_k \rangle = \langle AI_{k+1} \rangle = \langle AI_{k+2} \rangle = \ldots = \langle AI \rangle. \qquad (2.10)$$

Алгоритм построения матрицы $[A]$ основан на следующей теореме.

Теорема 2. Если для некоторого k верно равенство

$$\langle AI_{k+1} \rangle = \langle AI_k \rangle \qquad (2.11)$$

то k обладает свойством (2.10).

Доказательство.
Достаточно доказать равенство

$$\langle AI_{k+2} \rangle = \langle AI_k \rangle. \qquad (2.12)$$

Пусть $V \in AI_{k+2} \setminus AI_{k+1}$, тогда V имеет вид $A^{xy}A^{u,v}I$, где $x \in X, y \in Y$ и $|u| = |v| = k+1$. Поскольку $A^{u,v}I \in AI_{k+1} \subseteq \langle AI_k \rangle$, то, следовательно, $A^{u,v}I$ является линейной комбинацией вида

$$A^{u,v}I = \sum_{i=1}^{m} \lambda_i V_i \quad (\forall\, i = 1, \ldots, m \quad \lambda_i \in \mathbf{R},\ V_i \in AI_k).$$

Следовательно,

$$V = A^{xy}\Big(\sum_{i=1}^{m} \lambda_i V_i \Big) = \sum_{i=1}^{m} \lambda_i (A^{xy} V_i) = \sum_{i=1}^{m} \lambda_i W_i \qquad (2.13)$$

где $W_i \in AI_{k+1} \subseteq \langle AI_k \rangle$. откуда на основании (2.13) заключаем, что V является линейной комбинацией элементов $\langle AI_k \rangle$, поэтому $V \in \langle AI_k \rangle$. Таким образом, $AI_{k+2} \subseteq \langle AI_k \rangle$, откуда следует (2.12).
■

Из теоремы 2 непосредственно следует, что если k – наименьший номер, для которого верно (2.11), то $k \leq |S| - 1$.

Используя теорему 2, можно определить следующий алгоритм построения матрицы $[A]$. Мы будем обозначать записью CA переменную, значениями которой являются множества вектор-столбцов порядка $|S|$. Алгоритм состоит из перечисленных ниже трёх шагов. Шаг 2 может выполняться несколько раз.

1. Значение CA полагается равным $\{I\}$ $(= AI_0)$.

2. Пусть $V_1, \dots, V_m$ – список всех столбцов вида $A^{xy}V$, где $x \in X, y \in Y$ и $V \in CA$. Выполняется цикл:

   `for i=1 to m do {`
   `if` $V_i \notin \langle CA \rangle$ `then` V_i добавляется к CA
   `}`

3. Если во время выполнения шага 2 множество CA изменилось, то шаг 2 выполняется ещё раз, иначе алгоритм заканчивает работу.

Обоснуем корректность данного алгоритма. Нетрудно видеть, что если перед выполнением шага 2 было верно равенство $\langle CA \rangle = \langle AI_k \rangle$ для некоторого $k \geq 0$, то после выполнения этого шага будет верно равенство $\langle CA \rangle = \langle AI_{k+1} \rangle$. Следовательно, через не более чем $|S| - 1$ выполнений шага 2 будет верно равенство $\langle CA \rangle = \langle AI \rangle$, и шаг 2 выполнится не более $|S|$ раз. Поскольку каждый добавляемый к CA вектор V_i не принадлежит пространству $\langle CA \rangle$, то, следовательно, в каждый момент времени CA состоит из линейно независимых векторов, т.е. после завершения работы алгоритма CA является базисом пространства $\langle AI \rangle$. ■

2.1.5 Матричные обозначения

Мы будем использовать следующие обозначения, связанные с матрицами.

- Если ξ_1 и ξ_2 – вектор-строки размерностей n_1 и n_2 соответственно, то запись (ξ_1, ξ_2) обозначает вектор-строку размерности $n_1 + n_2$, первые n_1 компонентов которой совпадают с соответствющими компонентами ξ_1, а остальные компоненты – с соответствующими компонентами ξ_2.

- Если λ_1 и λ_2 – вектор-столбцы размерностей n_1 и n_2 соответственно, то запись $\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$ обозначает вектор-столбец размерности $n_1 + n_2$, первые n_1 компонентов которого совпадают с соответствующими компонентами λ_1, а остальные компоненты – с соответствующими компонентами λ_2.

- Если A и B – матрицы размерностей (m, n) и (k, l) соответственно, то запись $\begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix}$ обозначает матрицу размерности $(m + k, n + l)$, определяемую естественным образом.

- Для каждой матрицы A запись $\tilde{A}$ (или $A^{\sim}$) обозначает матрицу, транспонированную к матрице A.

2.1.6 Эквивалентность вероятностных автоматов

Пусть задана пара ВА A_1, A_2, у которых одинаковы множества входных сигналов и множества выходных сигналов, т.е. A_1 и A_2 имеют вид

$$A_i = (X, Y, S_i, P_i, \xi_i^0) \quad (i = 1, 2).$$

A_1 и A_2 называются **эквивалентными**, если их реакции совпадают, т.е. верно равенство

$$f_{A_1} = f_{A_2}. \tag{2.14}$$

Нетрудно доказать, что равенство (2.14) равносильно соотношению $\xi_1 \underset{A}{\sim} \xi_2$, где A имеет вид $(X, Y, S_1 \sqcup S_2, P, \xi^0)$, и

$$\begin{aligned} &\forall\, x \in X,\ y \in Y \quad A^{xy} = \begin{pmatrix} A_1^{xy} & \mathbf{0} \\ \mathbf{0} & A_2^{xy} \end{pmatrix}, \\ &\xi_1 = (\xi_1^0, \mathbf{0}),\ \ \xi_2 = (\mathbf{0}, \xi_2^0) \end{aligned} \tag{2.15}$$

(символы $\mathbf{0}$ в (2.15) изображают нулевые матрицы или вектор-строки соответствующих размеров).

Если ВА A_1 и A_2 эквивалентны, то мы будем обозначать этот факт записью $A_1 \sim A_2$.

2.2 Редукция вероятностных автоматов

Редукция ВА заключается в построении по заданному ВА A такого ВА, который был бы эквивалентен A, и содержал меньше состояний, чем A (если это возможно). Мы будем рассматривать два метода редукции: выделение достижимой части и удаление выпуклых комбинаций.

2.2.1 Выделение достижимой части

Пусть A – ВА вида (2.1). Понятие **достижимого состояния** ВА A определяется рекурсивно: состояние $s \in S$ достижимо, если

- либо $s^{\xi^0} \neq 0$,
- либо существует достижимое состояние $s' \in S$, такое, что

$$\exists\, x \in X, y \in Y: \quad P(s', x, s, y) > 0. \tag{2.16}$$

Нетрудно доказать, что $A \sim A_r \stackrel{\text{def}}{=} (X, Y, S_r, P_r, \xi^0_r)$, где

- S_r состоит из всех достижимых состояний ВА A, и
- P_r и ξ^0_r являются соответствующими ограничениями P и ξ^0.

ВА A_r называется **достижимой частью** ВА A. Алгоритм построения по заданному ВА его достижимой части аналогичен соответствующему алгоритму для детерминированных автоматов (см. конец пункта 1.3.3).

2.2.2 Удаление выпуклых комбинаций

Пусть A – ВА вида (X, Y, S, P, ξ^0). Мы будем говорить, что состояние $s \in S$ является **выпуклой комбинацией** других состояний ВА A, если строка s матрицы $[A]$ является выпуклой комбинацией других строк этой матрицы, т.е. существует распределение $\xi \in (S \setminus \{s\})^{\triangle}$, удовлетворяющее условию

$$[A]_s = \sum_{s' \in S \setminus \{s\}} (s')^{\xi} [A]_{s'}. \tag{2.17}$$

Если в множестве S состояний ВА A есть состояние s, являющееся выпуклой комбинацией других состояний этого ВА, то можно определить ВА B, который эквивалентен A, и множество состояний которого имеет вид $S \setminus \{s\}$. Мы будем говорить, что B получается из A путем удаления выпуклой комбинации s.

Автомат B определяется следующим образом. Пусть упорядочение множества S имеет вид $(s_1, \ldots, s_n)$, и вышепомянутое состояние s является последним в этом упорядочении (т.е. $s = s_n$). Обозначим символом M матрицу

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ & \ldots & & & \\ 0 & 0 & \ldots & 1 & 0 \\ s_1^\xi & s_2^\xi & \ldots & s_{n-1}^\xi & 0 \end{pmatrix}$$

и обозначим символом C ВА $(X, Y, S, Q, \xi^0 M)$, где

$$\forall\, x \in X,\ y \in Y \quad C^{xy} = A^{xy} M. \tag{2.18}$$

Докажем, что $\forall\, u \in X^*, v \in Y^*$ верно равенство

$$C^{u,v} I = A^{u,v} I. \tag{2.19}$$

(2.19) верно, когда u и v имеют разную длину. Для u и v одинаковой длины будем доказывать (2.19) индукцией по длине u.

1. (2.19) верно, когда $u = v = \varepsilon$.

2. Пусть (2.19) верно для некоторых u, v. Тогда $\forall\, x \in X, y \in Y$

$$C^{xu,yv} I = C^{xy} C^{u,v} I = A^{xy} M A^{u,v} I \tag{2.20}$$

(второе равенство в (2.20) следует из (2.18) и (2.19)).

Докажем, что верно равенство

$$M A^{u,v} I = A^{u,v} I. \tag{2.21}$$

Из (2.17) следует. что

$$(s_1^\xi\ \ldots\ s_{n-1}^\xi\ 0)[A] = [A]_{s_n} = (0\ \ldots\ 0\ 1)[A]$$

откуда следует

$$M[A] = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} [A] = [A] \qquad (2.22)$$

Из (2.22) следует, что для каждого столбца V, матрицы $[A]$ верно равенство

$$MV = V. \qquad (2.23)$$

Поскольку столбцы $[A]$ образуют базис $\langle AI \rangle$, то, следовательно, (2.23) верно в том случае, когда V является произвольным элементом $\langle AI \rangle$. В частности, (2.23) верно для всех векторов из AI. Таким образом, равенство (2.21) доказано.

Из (2.21) и из (2.20) следует, что

$$C^{xu,yv}I = A^{xy}MA^{u,v}I = A^{xy}A^{u,v}I = A^{xu,yv}I. \qquad (2.24)$$

Таким образом, если (2.19) верно, то будет верно равенство, получаемое из (2.19) заменой u на xu, а v – на yv.

Следовательно, (2.19) верно для всех $u \in X^*$, $v \in Y^*$.

Докажем, что ВА A и C эквивалентны, т.е. $f_A = f_C$. Данное равенство равносильно утверждению

$$\forall\, u \in X^*, \forall\, v \in Y^* \qquad \xi^0 A^{u,v} I = \xi^0 M C^{u,v} I. \qquad (2.25)$$

(2.25) следует из (2.19) и из (2.21).

Заметим, что состояние s_n ВА C не является достижимым. Действительно, т.к. последний столбец матрицы M является нулевым, то

- значение $s_n^{\xi^0 M}$, которое является последним элементом вектор-строки $\xi^0 M$, равно 0, и

- для каждого $x \in X$ и каждого $y \in Y$ последний столбец матрицы $C^{xy} = A^{xy}M$ является нулевым, поэтому неравенство (2.16), в котором P заменено на Q, и s – на s_n, неверно для каждого $s' \in S$.

Искомый ВА B определяется как ВА, получаемый из ВА C удалением недостижимого состояния s_n и соответствующим ограничением поведения и начального распределения ВА C. Из утверждения в пункте 2.2.1 следует, что $C \sim B$. Поскольку свойство эквивалентности ВА является транзитивным, то из $A \sim C$ и $C \sim B$ следует, что $A \sim B$. ■

2.2.3 Метод распознавания выпуклых комбинаций состояний

Для реализации изложенного в предыдущем пункте метода редукции ВА путем удаления выпуклых комбинаций состояний необходимо иметь алгоритм решения следующей задачи: пусть задан ВА A, и s – одно из состояний этого ВА, требуется

- определить, является ли состояние s выпуклой комбинацией других состояний ВА A, т.е. является ли строка $[A]_s$ выпуклой комбинацией других строк матрицы $[A]$, и
- если ответ на этот вопрос положителен, то найти коэффициенты этой выпуклой комбинации.

Данную задачу можно свести к задаче линейного программирования (ЗЛП), на основе нижеследующей теоремы.

Теорема 3.

Пусть задан ВА A, и s – одно из состояний этого ВА. Обозначим записью $\{W_1, \ldots, W_m\}$ совокупность строк матрицы $[A]$, за исключением строки $[A]_s$.

Тогда следующие утверждения эквивалентны.

1. $[A]_s$ является выпуклой комбинацией строк $\{W_1, \ldots, W_m\}$, т.е.

$$\exists\, (\xi_1, \ldots, \xi_m) \in \{1, \ldots, m\}^{\triangle} : [A]_s = \sum_{i=1}^{m} \xi_i W_i. \qquad (2.26)$$

2. Существует решение ЗЛП, в которой

- множество переменных имеет вид

$$\{x_1, \ldots, x_m, y_1, \ldots, y_n\}$$

  где n – число состояний ВА A,

- ограничения в форме неравенств имеют вид $x_i \geq 0$ и $y_j \geq 0$, где $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$,
- ограничения в форме равенств выражаются в виде матричного равенства $(X,\ Y) \begin{pmatrix} W \\ E_n \end{pmatrix} = [A]_s$, где
  - X и Y – вектор-строки переменных:

$$X = (x_1, \ldots, x_m), \quad Y = (y_1, \ldots, y_n).$$

  - W – матрица, получаемая из $[A]$ путем удаления строки $[A]_s$,
  - E_n – единичная матрица порядка n,

  а также равенства $\sum\limits_{i=1}^{m} x_i + \sum\limits_{i=1}^{n} y_i = 1$,

- целевая функция имеет вид $\sum\limits_{i=1}^{n} y_i \to \min$,

и значение целевой функции на этом решении равно 0.

Доказательство.
Пусть верно утверждение 1. Тогда решение ЗЛП имеет вид $x_1 = \xi_1, \ldots, x_m = \xi_m$, $y_1 = 0, \ldots, y_n = 0$.

Обратно, пусть верно утверждение 2, т.е. существует решение ЗЛП, значение целевой функции на котором равно 0. Тогда из ограничения $y_j \geq 0$ $(j = 1, \ldots, n)$ следует, что значения переменных $y_1, \ldots, y_n$ на этом решении равны 0. Нетрудно видеть, что совокупность $(\xi_1, \ldots, \xi_m)$ значений переменных $x_1, \ldots, x_m$ на этом решении удовлетворяет условиям в соотношении (2.26). ■

Отметим, что одно из опорных решений ЗЛП, сформулированной в теореме 3, имеет вид $X = \mathbf{0}$, $Y = [A]_s$.

2.3 Вероятностные реакции

2.3.1 Понятие вероятностной реакции

Пусть X и Y – конечные множества.

Вероятностной реакцией (ВР) из X в Y называется СФ $f : X^* \underset{r}{\to} Y^*$, удовлетворяющая условию: $\forall\, u \in X^*, \forall\, v \in Y^*$

$$\begin{array}{l} \text{если } |u| \neq |v|, \text{то } f(u,v) = 0 \\ \forall\, x \in X \quad f(u,v) = \sum\limits_{y \in Y} f(ux, vy). \end{array} \tag{2.27}$$

Запись $R(X,Y)$ обозначает совокупность всех ВР из X в Y.

Теорема 4.
Для каждого ВА $A = (X, Y, S, P, \xi^0)$ и каждого $\xi \in S^\triangle$

$$A^\xi \in R(X,Y).$$

Доказательство.
Докажем, что $\forall\, u \in X^*, \forall\, v \in Y^*$ СФ $f \overset{\text{def}}{=} A^\xi$ удовлетворяет условию (2.27).

- Если $|u| \neq |v|$, то $A^{u,v} = 0$, поэтому $A^\xi(u,v) = \xi A^{u,v} I = 0$,
- $\forall\, x \in X$

$$\begin{array}{l} \sum\limits_{y \in Y} A^\xi(ux, vy) = \sum\limits_{y \in Y} \xi A^{ux,vy} I = \\ = \sum\limits_{y \in Y} \xi A^{u,v} A^{xy} I = \xi A^{u,v} (\sum\limits_{y \in Y} A^{xy} I) = \\ = \xi A^{u,v} I = A^\xi(u,v) \end{array} \tag{2.28}$$

(в (2.28) используется равенство (2.8)). ■

2.3.2 Остаточные вероятностные реакции

Пусть заданы

- конечные множества X и Y,
- ВР $f \in R(X,Y)$, и

- строки $u \in X^*, v \in Y^*$, такие, что $f(u,v) \neq 0$.

Обозначим записью $f_{u,v}$ функцию вида $f_{u,v} : X^* \times Y^* \to \mathbf{R}$, определяемую следующим образом:

$$\forall\, u' \in X^*, \forall\, v' \in Y^* \quad f_{u,v}(u',v') \stackrel{\text{def}}{=} \frac{f(uu',vv')}{f(u,v)}. \tag{2.29}$$

Теорема 5.
Если $f \in R(X,Y)$ и $f(u,v) \neq 0$, то функция $f_{u,v}$, определяемая соотношением (2.29), является СФ.

Доказательство.
Поскольку все значения функции $f_{u,v}$ неотрицательны, то достаточно доказать, что

$$\forall\, u' \in X^* \quad \sum_{v' \in Y^*} f_{u,v}(u',v') = 1. \tag{2.30}$$

(2.30) эквивалентно соотношению

$$\forall\, u' \in X^* \quad \sum_{v' \in Y^*} f(uu',vv') = f(u,v). \tag{2.31}$$

Из предположения $f(u,v) \neq 0$ следует, что $|u| = |v|$. Поэтому $f(uu',vv') = 0$ при $|u'| \neq |v'|$, и, следовательно, (2.31) эквивалентно условию: $\forall\, k \geq 0$

$$\forall\, u' \in X^k \quad \sum_{v' \in Y^k} f(uu',vv') = f(u,v). \tag{2.32}$$

Докажем (2.32) индукцией по k. Если $k = 0$, то (2.32), очевидно, верно. Пусть (2.32) верно для некоторого k. Докажем, что

$$\forall\, u' \in X^{k+1} \quad \sum_{v' \in Y^{k+1}} f(uu',vv') = f(u,v). \tag{2.33}$$

(2.33) эквивалентно утверждению: $\forall\, u' \in X^k$, $\forall\, x \in X$

$$\sum_{v' \in Y^k,\; y \in Y} f(uu'x,vv'y) = f(u,v). \tag{2.34}$$

Поскольку $\forall\, x \in X,\ \forall\, u' \in X^k,\ \forall\, v' \in Y^k$

$$f(uu', vv') = \sum_{y \in Y} f(uu'x, vv'y)$$

то (2.34) можно переписать следующим образом:

$$\begin{aligned} &\sum_{v' \in Y^k,\, y \in Y} f(uu'x, vv'y) = \sum_{v' \in Y^k} \Big(\sum_{y \in Y} f(uu'x, vv'y) \Big) = \\ &= \sum_{v' \in Y^k} f(uu', vv') = f(u, v). \end{aligned} \tag{2.35}$$

Последнее равенство в (2.35) совпадает с равенством в (2.32), и оно верно по индуктивному предположению. ■

Теорема 6.
Если $f \in R(X, Y)$ и $f(u, v) \neq 0$, то $f_{u,v} \in R(X, Y)$.

Доказательство.
Требуется доказать, что $\forall\, u' \in X^*, \forall\, v' \in Y^*$

$$\begin{array}{l} \text{если } |u'| \neq |v'|, \text{то } f_{u,v}(u', v') = 0 \\ \forall\, x \in X \quad f_{u,v}(u', v') = \sum_{y \in Y} f_{u,v}(u'x, v'y). \end{array} \tag{2.36}$$

Из условия $f(u, v) \neq 0$ следует, что $|u| = |v|$, поэтому, согласно определению функции f_{uv}, можно переписать (2.36) следующим образом:

$$\begin{array}{l} \text{если } |uu'| \neq |vv'|, \text{то } f(uu', vv') = 0 \\ \forall\, x \in X \quad f(uu', vv') = \sum_{y \in Y} f(uu'x, vv'y). \end{array} \tag{2.37}$$

Первое утверждение в (2.37) верно потому, что f – ВР, а второе утверждение следует из доказанного выше соотношения (2.32) (в данном случае $k = 1$). ■

Теорема 7.
Если $f \in R(X, Y)$, $f(u, v) \neq 0$ и $f(uu', vv') \neq 0$, то

$$f_{uu', vv'} = (f_{u,v})_{u', v'}.$$

Доказательство.
$\forall\, u'' \in X^*, \forall\, v'' \in Y^*$

(a) $f_{uu',vv'}(u'',v'') = \frac{f(uu'u'',vv'v'')}{f(uu',vv')}$

(b) $(f_{u,v})_{u',v'}(u'',v'') = \frac{f_{u,v}(u'u'',v'v'')}{f_{u,v}(u',v')} = \frac{f(uu'u'',vv'v'')/f(u,v)}{f(uu',vv')/f(u,v)}$

Нетрудно видеть, что правые части в (a) и (b) совпадают. ■

Если $f \in R(X,Y)$ и u, v – строки из X^* и Y^* соответственно, такие, что $f(u,v) \neq 0$, то $f_{u,v}$ называется **остаточной ВР** для f. Мы будем обозначать записью S_f совокупность всех остаточных ВР для f. Отметим, что $f \in S_f$, т.к. $f(\varepsilon,\varepsilon) = 1$, поэтому $f_{\varepsilon,\varepsilon} = f$.

Обозначим записью A_f пятерку (X, Y, S_f, P_f, ξ_f), где P_f – СФ вида

$$P_f : S_f \times X \underset{r}{\rightarrow} S_f \times Y,$$

определяемая следующим образом:

$$\begin{array}{l} \forall\, g, g' \in S_f,\ \forall\, x \in X,\ \forall\, y \in Y \\ P_f(g,x,g',y) \stackrel{\text{def}}{=} \begin{cases} g(x,y), & \text{если } g' = g_{x,y}, \\ 0, & \text{иначе.} \end{cases} \end{array} \tag{2.38}$$

Теорема 8.
Если $f \in R(X,Y)$ и $|S_f| < \infty$, то A_f – ВА, и $f_{A_f} = f$.

Доказательство.
Из определения СФ P_f следует, что для любых $g \in S_f$, $x \in X$, $y \in Y$, таких, что $g(x,y) \neq 0$, верно равенство

$$\xi_g A_f^{xy} = g(x,y)\xi_{g_{x,y}}. \tag{2.39}$$

(напомним, что ξ_g – распределение, такое, что $\forall\, h \in S_f \quad h^{\xi_g} = 1$, если $h = g$, и $h^{\xi_g} = 0$, если $h \neq g$).

Докажем, что $\forall\, u \in X^*$, $\forall\, v \in Y^* : f(u,v) \neq 0$

$$\xi_f A_f^{u,v} = f(u,v)\xi_{f_{u,v}}. \tag{2.40}$$

Доказательство будем вести индукцией по длине u.

Если $|u| = 0$, т.е. $u = v = \varepsilon$, то обе части (2.40) равны ξ_f.

Иначе u и v имеют вид $u'x$ и $v'y$ соответственно, причём $f(u',v') \neq 0$, (т.к. если $f(u',v') = 0 = \sum\limits_{y' \in Y} f(u'x, v'y')$, то $f(u,v) =$

$f(u'x, v'y) = 0$), и, по индуктивному предположению, верно равенство

$$\xi_f A_f^{u',v'} = f(u',v')\xi_{f_{u',v'}}. \tag{2.41}$$

Используя (2.39), (2.41) и теорему 7, получаем цепочку равенств

$$\begin{array}{l} \xi_f A_f^{u,v} = \\ = \xi_f A_f^{u'x,v'y} = \xi_f A_f^{u',v'} A_f^{x,y} = \\ = f(u',v')\xi_{f_{u',v'}} A_f^{x,y} = \\ = f(u',v') f_{u',v'}(x,y)\xi_{(f_{u',v'})_{x,y}} = \\ = f(u',v')\frac{f(u'x,v'y)}{f(u',v')}\xi_{f_{u'x,v'y}} = \\ = f(u'x,v'y)\xi_{f_{u'x,v'y}} = f(u,v)\xi_{f_{u,v}}. \end{array} \tag{2.42}$$

Таким образом, для любых $u \in X^*$, $v \in Y^*$, таких, что $f(u,v) \neq 0$, верно равенство (2.40), из которого следует цепочка равенств

$$\begin{array}{l} A_f^{\xi_f}(u,v) = \xi_f A_f^{u,v} I = \\ = f(u,v)\xi_{f_{u,v}} I = f(u,v) \cdot 1 = f(u,v). \end{array}$$

Следовательно, в случае $f(u,v) \neq 0$ верно равенство

$$f_{A_f}(u,v) \stackrel{\text{def}}{=} A_f^{\xi_f}(u,v) = f(u,v). \tag{2.43}$$

Докажем, что (2.43) верно и в случае $f(u,v) = 0$.

- Если $|u| \neq |v|$, то левая часть (2.43) равна 0 по определению матриц вида $A^{u,v}$.
- Пусть $|u| = |v| > 0$, и u и v имеют вид $x_1 \ldots x_n$ и $y_1 \ldots y_n$ соответственно. Существуют номер $k \in \{1, \ldots, n\}$ и строки u', v', такие, что

  $$\begin{array}{l} u = u'x_k \ldots x_n, \; v = v'y_k \ldots y_n, \\ f(u',v') \neq 0, \; f(u'x_k, v'y_k) = 0. \end{array} \tag{2.44}$$

  Имеем цепочку равенств

  $$\begin{array}{l} A_f^{\xi_f}(u,v) = \xi_f A_f^{u',v'} A_f^{x_k,y_k} \ldots A_f^{x_n,y_n} I = \\ = f(u',v')\xi_{f_{u',v'}} A_f^{x_k,y_k} \ldots A_f^{x_n,y_n} I \end{array} \tag{2.45}$$

Строка $\xi_{f_{u',v'}} A_f^{x_k,y_k}$ является нулевой, поскольку её элементы имеют вид

$$P_f(g, x_k, g', y_k) \tag{2.46}$$

где $g = f_{u',v'}$, и, согласно определению (2.38) СФ P_f, элемент (2.46) отличен от 0 если и только если $f_{u',v'}(x_k, y_k) \neq 0$, т.е. $f(u'x_k, v'y_k) \neq 0$. Учитывая (2.44), получаем, что все элементы (2.46) равны 0, т.е. $\xi_{f_{u',v'}} A_f^{x_k,y_k}$ является нулевой строкой. Таким образом, правая часть в (2.45) равна 0, откуда следует, что равенство (2.43) в рассматриваемом случае также верно. ■

2.3.3 Реализуемость вероятностных реакций

ВР f называется **реализуемой**, если $\exists$ ВА $A: \ f_A = f$.

Согласно теореме 8, если $|S_f| < \infty$, то f реализуема. Обращение этого утверждения неверно: согласно нижеследующей теореме, существует реализуемая ВР f, такая, что $|S_f| = \infty$.

Теорема 9.

Пусть $f = f_A$, где A – ВА вида (X, Y, S, P, ξ^0), компоненты которого удовлетворяют условиям:

$$\begin{array}{l} |S| = 2, \ \xi^0 = (1, 0), \\ \exists\, x \in X, \exists\, y \in Y : A^{x,y} = \alpha \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \quad (\alpha, \beta \in (0,1)). \end{array}$$

Тогда $|S_f| = \infty$.

Доказательство.

Если $u_k = \underbrace{x \ldots x}_{k}$, $v_k = \underbrace{y \ldots y}_{k}$, то $A^{u_k,v_k} = \alpha^k \begin{pmatrix} 1 & k\beta \\ 0 & 1 \end{pmatrix}$, поэтому $f(u_k, v_k) = \alpha^k(1 + k\beta)$. $\forall\, k \geq 1$ определена остаточная ВР f_{u_k,v_k}, и нетрудно видеть, что $\forall\, s \geq 1$

$$f_{u_k,v_k}(u_s, v_s) = \tfrac{f(u_k u_s, v_k v_s)}{f(u_k, v_k)} = \tfrac{\alpha^{k+s}(1+(k+s)\beta)}{\alpha^k(1+k\beta)} = \tfrac{\alpha^s(1+(k+s)\beta)}{1+k\beta}.$$

Если для некоторых $k_1, k_2 \geq 1$ функции $f_{u_{k_1},v_{k_1}}$ и $f_{u_{k_2},v_{k_2}}$ совпадают, то

$$\forall\, s \geq 1 \quad \tfrac{\alpha^s(1+(k_1+s)\beta)}{1+k_1\beta} = \tfrac{\alpha^s(1+(k_2+s)\beta)}{1+k_2\beta},$$

откуда следует, что $k_1 = k_2$. Таким образом, при различных k функции f_{u_k,v_k} различны, т.е. $|S_f| = \infty$. ■

Пусть X и Y – конечные множества. Мы будем использовать следующие определения и обозначения.

- Запись $[0,1]^{X^* \times Y^*}$ обозначает множество функций вида
$$f : X^* \times Y^* \to [0,1].$$

- Для каждого $\Gamma \subseteq [0,1]^{X^* \times Y^*}$ **конусом** над Γ называется подмножество $C_0(\Gamma) \subseteq [0,1]^{X^* \times Y^*}$, состоящее из функций вида $\sum\limits_{i=1}^{n} a_i f_i$, где
  - $\forall\, i = 1, \ldots, n \quad a_i \in [0,1],\ f_i \in \Gamma, \quad \sum\limits_{i=1}^{n} a_i \leq 1$, и
  - $\forall\, (u,v) \in X^* \times Y^* \quad \Big(\sum\limits_{i=1}^{n} a_i f_i\Big)(u,v) \stackrel{\text{def}}{=} \sum\limits_{i=1}^{n} a_i f_i(u,v).$

- $\forall\, x \in X, \forall\, y \in Y$ запись D^{xy} обозначает отображение вида
$$D^{xy} : [0,1]^{X^* \times Y^*} \to [0,1]^{X^* \times Y^*},$$
называемое **сдвигом**, сопоставляющее каждой функции f из $[0,1]^{X^* \times Y^*}$ функцию, обозначаемую записью fD^{xy}, где
$$\forall\, u \in X^*, \forall\, v \in Y^* \quad (fD^{xy})(u,v) \stackrel{\text{def}}{=} f(xu, yv). \qquad (2.47)$$

- Подмножество $\Gamma \subseteq [0,1]^{X^* \times Y^*}$ называется **устойчивым относительно сдвигов**, если
$$\forall\, f \in \Gamma,\ \forall\, x \in X, \forall\, y \in Y \quad fD^{xy} \in C_0(\Gamma).$$

Теорема 10.
Пусть X и Y – конечные множества, и $f \in R(X,Y)$. Следующие условия эквивалентны:

- f реализуема,

- $\exists$ конечное $\Gamma \subseteq R(X,Y)$, устойчивое относительно сдвигов, и такое, что $f \in C_0(\Gamma)$.

Доказательство.
Пусть f реализуема, т.е. $\exists$ ВА $A = (X, Y, S, P, \xi^0)$:

$$\forall\, u \in X^*, \forall\, v \in Y^* \quad f(u,v) = \xi^0 A^{u,v} I.$$

$\forall\, s \in S$ обозначим записью A_s ВА (X, Y, S, P, ξ_s). В качестве искомого Γ можно взять множество $\{f_{A_s} \mid s \in S\}$.

$f \in C_0(\Gamma)$, т.к. $f = \sum\limits_{s \in S} s^{\xi^0} f_{A_s}$, и $\Gamma \subseteq R(X,Y)$ (по теореме 4).

Докажем, что Γ устойчиво относительно сдвигов, т.е. $\forall\, s \in S$, $\forall\, x \in X$, $\forall\, y \in Y$ $f_{A_s} D^{xy} \in C_0(\Gamma)$. Согласно (2.47),

$$\begin{array}{l} \forall\, u \in X^*, \forall\, v \in Y^* \\ (f_{A_s} D^{xy})(u,v) = f_{A_s}(xu, yv) = \xi_s A^{xu,yv} I = \xi_s A^{xy} A^{u,v} I. \end{array} \tag{2.48}$$

Нетрудно видеть, что

$$\xi_s A^{xy} A^{u,v} I = \sum_{s' \in S} a_{s'} f_{A_{s'}}(u,v)$$

где $\forall\, s' \in S$ $\;a_{s'}$ – компонента вектор-строки $\xi_s A^{xy}$, соответствующая состоянию s' (т.е. элемент матрицы A^{xy}, находящийся в строке s столбце s'). Свойства $\forall\, s' \in S$ $\;a_{s'} \in [0,1]$ и $\sum\limits_{s' \in S} a_{s'} \leq 1$ являются следствием соответствующих свойств матрицы A^{xy}.

Обратно, пусть $f \in C_0(\Gamma)$, где $\Gamma = \{f_1, \ldots, f_n\} \subseteq R(X,Y)$, и Γ устойчиво относительно сдвигов. Определим A как ВА

$$A \stackrel{\text{def}}{=} (X, Y, S, P, \xi^0), \tag{2.49}$$

компоненты которого имеют следующий вид.

- $S \stackrel{\text{def}}{=} \{1, \ldots, n\}$.
- $\xi^0 = (a_1, \ldots, a_n)$, где $a_1, \ldots, a_n$ – коэффициенты представления f в виде суммы

$$f = \sum_{i=1}^{n} a_i f_i, \quad \text{где } \forall\, i = 1, \ldots, n \;\; a_i \geq 0,\; \sum_{i=1}^{n} a_i \leq 1. \tag{2.50}$$

  По предположению, $f \in R(X,Y)$, в частности, $f(\varepsilon, \varepsilon) = 1$, откуда следует равенство $\sum\limits_{i=1}^{n} a_i = 1$, поэтому $\xi^0 \in S^{\triangle}$.

- Поведение $P : S \times X \times S \times Y \to [0,1]$ ВА (2.49) определяется матрицами A^{xy} порядка n $(x \in X, y \in Y)$:

$$P(i,x,j,y) \stackrel{\text{def}}{=} A^{xy}_{ij},$$

где $\forall\, x \in X$, $\forall\, y \in Y$, $\forall\, i = 1,\dots,n$ строка i матрицы A^{xy} состоит из элементов $a_{i1},\dots,a_{in}$ представления функции $f_i D^{xy}$ в виде суммы $\sum\limits_{j=1}^{n} a_{ij} f_j$ $(\forall\, i,j \;\; a_{ij} \geq 0,\ \sum\limits_{j=1}^{n} a_{ij} \leq 1)$.

Докажем, что P является СФ вида $S \times X \underset{r}{\to} S \times Y$. Данное утверждение эквивалентно соотношению $\Big(\sum\limits_{y \in Y} A^{xy}\Big) I = I$.

$\forall\, i = 1,\dots,n$ из

$$f_i D^{xy} = \sum_{j=1}^{n} A^{xy}_{ij} f_j \tag{2.51}$$

следует, что

$$(f_i D^{xy})(\varepsilon,\varepsilon) = \sum_{j=1}^{n} A^{xy}_{ij} f_j(\varepsilon,\varepsilon). \tag{2.52}$$

Т.к. $\forall\, i = 1,\dots,n$ $\; f_j \in R(X,Y)$, то $f_j(\varepsilon,\varepsilon) = 1$. Кроме того, левая часть (2.52) равна $f_i(x,y)$. Поэтому (2.52) можно переписать в виде $f_i(x,y) = \sum\limits_{j=1}^{n} A^{xy}_{ij}$, откуда следует соотношение

$$\sum_{y \in Y} f_i(x,y) = \sum_{y \in Y} \sum_{j=1}^{n} A^{xy}_{ij}. \tag{2.53}$$

Т.к. $f_i \in R(X,Y)$, то, согласно второму соотношению в (2.27), левая часть (2.53) равна $f_i(\varepsilon,\varepsilon)$, т.е. равна 1. Учитывая это, и поменяв порядок суммирования в правой части (2.53), получаем соотношение

$$\sum_{j=1}^{n} \sum_{y \in Y} A^{xy}_{ij} = 1. \tag{2.54}$$

Нетрудно видеть, что истинность (2.54) $\forall\, i = 1,\dots,n$ эквивалентна доказываемому равенству $\Big(\sum\limits_{y \in Y} A^{xy}\Big) I = I$.

Докажем, что реакция ВА (2.49) совпадает с f, т.е.

$$\forall\, u \in X^*, \forall\, v \in Y^* \quad \xi^0 A^{u,v} I = f(u,v). \tag{2.55}$$

Если $|u| \neq |v|$, то левая часть равенства в (2.55) равна 0 по определению матриц вида $A^{u,v}$, и правая часть равенства в (2.55) равна 0 согласно предположению $f \in R(X,Y)$ и первому соотношению в (2.27).

Пусть $|u| = |v|$. Докажем (индукцией по $|u|$), что

$$A^{u,v} I = \begin{pmatrix} f_1(u,v) \\ \dots \\ f_n(u,v) \end{pmatrix}. \tag{2.56}$$

Если $u = v = \varepsilon$, то обе части (2.56) равны I.

Если $u = xu'$ и $v = yv'$, то, предполагая верным равенство (2.56), в котором u и v заменены на u' и v', имеем:

$$\begin{gathered} A^{u,v} I = A^{xu',yv'} I = A^{xy} A^{u',v'} I = A^{xy} \begin{pmatrix} f_1(u',v') \\ \dots \\ f_n(u',v') \end{pmatrix} = \\ = \begin{pmatrix} \sum\limits_{i=1}^{n} A^{xy}_{1j} f_j(u',v') \\ \dots \\ \sum\limits_{i=1}^{n} A^{xy}_{nj} f_j(u',v') \end{pmatrix}. \end{gathered} \tag{2.57}$$

Из (2.51) следует, что правую часть в (2.57) можно переписать в виде

$$\begin{pmatrix} (f_1 D^{xy})(u',v') \\ \dots \\ (f_n D^{xy})(u',v') \end{pmatrix}. \tag{2.58}$$

Согласно определению (2.47) функций вида fD^{xy}, столбец (2.58) совпадает с правой частью доказываемого равенства (2.56).

Таким образом, равенство (2.56) доказано. Согласно этому равенству, левая часть доказываемого равенства (2.55) равна

$$\xi^0 \begin{pmatrix} f_1(u,v) \\ \dots \\ f_n(u,v) \end{pmatrix} = \sum_{i=1}^{n} \xi^0_i f_i(u,v). \tag{2.59}$$

По определению ξ^0 (см. (2.50)), правая часть (2.59) равна $f(u,v)$, т.е. правой части доказываемого равенства (2.55). ■

2.4 Случайные последовательности

2.4.1 Понятие случайной последовательности

Пусть X – конечное множество.

Случайной последовательностью (СП) над X называется функция $\zeta \in [0,1]^{X^*}$, удовлетворяющая условиям

$$\begin{array}{l} \zeta(\varepsilon) = 1, \\ \forall\, u \in X^* \quad \zeta(u) = \sum\limits_{x \in X} \zeta(ux). \end{array} \tag{2.60}$$

Множество всех СП над X будет обозначаться записью $R(\mathbf{1}, X)$.

Если $\zeta \in R(\mathbf{1}, X)$, то запись D_ζ обозначает множество

$$\{u \in X^* \mid \zeta(u) \neq 0\}.$$

2.4.2 Остаточные случайные последовательности

Если $\zeta \in R(\mathbf{1}, X)$, и u – строка из X^*, такая, что $\zeta(u) \neq 0$, то запись ζ_u обозначает СП из $R(\mathbf{1}, X)$, называемую **остаточной СП** для ζ, и определяемую следующим образом:

$$\forall\, u' \in X^* \quad \zeta_u(u') \stackrel{\text{def}}{=} \tfrac{\zeta(uu')}{\zeta(u)}.$$

$\forall\, \zeta \in R(\mathbf{1}, X)$ запись S_ζ обозначает множество всех остаточных СП для ζ. Отметим, что $\zeta \in S_\zeta$, т.к. $\zeta = \zeta_\varepsilon$.

Обозначим записью A_ζ пятерку $(\mathbf{1}, X, S_\zeta, P_\zeta, \xi_\zeta)$, где $\mathbf{1}$ – одноэлементное множество, единственный элемент которого будет обозначаться символом e, и P_ζ – СФ вида

$$P_\zeta : S_\zeta \times \mathbf{1} \underset{r}{\to} S_\zeta \times X,$$

определяемая следующим образом: $\forall\, \zeta', \zeta'' \in S_\zeta$, $\forall\, x \in X$

$$P_\zeta(\zeta', e, \zeta'', x) \stackrel{\text{def}}{=} \begin{cases} \zeta'(x), & \text{если } \zeta'' = \zeta'_x, \\ 0, & \text{иначе.} \end{cases} \tag{2.61}$$

Теорема 11.

Если $\zeta \in R(\mathbf{1}, X)$ и $|S_\zeta| < \infty$, то A_ζ – ВА, и

$$\forall\, u \in X^* \;\; f_{A_\zeta}(e^{|u|}, u) = \zeta(u)$$

(запись $e^{|u|}$ обозначает строку из $|u|$ символов e, где $\{e\} = \mathbf{1}$).

Доказательство.

Утверждение теоремы следует из того, что если строка u имеет вид $x_1 \ldots x_k$, где $x_1, \ldots, x_k \in X$, то

$$f_{A_\zeta}(u) = \zeta(x_1)\zeta_{x_1}(x_2) \ldots \zeta_{x_1 \ldots x_{k-1}}(x_k). \quad \blacksquare$$

Теорема 12.

Если $\zeta \in R(\mathbf{1}, X)$ и $|S_\zeta| < \infty$, то ζ полностью определяется своими значениями на строках из $X^{\leq 2(|S_\zeta| - |X| + 1)}$. $\blacksquare$

2.4.3 Парные случайные последовательности

Пусть X, Y – пара конечных множеств.

Парной случайной последовательностью (ПСП) над парой (X, Y) называется функция $\eta \in [0,1]^{X^* \times Y^*}$, удовлетворяющая условиям:

$$\begin{array}{ll} \eta(\varepsilon, \varepsilon) = 1, & \\ \forall\, u \in X^*, \forall\, v \in Y^* & \left\{ \begin{array}{l} \text{если } |u| \neq |v|, \text{ то } \eta(u,v) = 0, \\ \eta(u,v) = \sum\limits_{x \in X, y \in Y} \eta(ux, vy). \end{array} \right. \end{array} \quad (2.62)$$

Запись $R(\mathbf{1}, X \times Y)$ обозначает множество всех ПСП над (X, Y).

Если $\eta \in R(\mathbf{1}, X \times Y)$, то записи η^X и η^Y обозначают СП из $R(\mathbf{1}, X)$ и $R(\mathbf{1}, Y)$ соответственно, определяемые следующим образом:

$$\begin{array}{ll} \forall\, u \in X^* & \eta^X(u) = \sum\limits_{v \in Y^*} \eta(u, v), \\ \forall\, v \in Y^* & \eta^Y(v) = \sum\limits_{u \in X^*} \eta(u, v). \end{array} \quad (2.63)$$

Если $\eta \in R(\mathbf{1}, X \times Y)$, и u, v – строки из X^* и Y^* соответственно, такие, что $\eta(u,v) \neq 0$, то запись $\eta_{u,v}$ обозначает ПСП из $R(\mathbf{1}, X \times Y)$, называемую **остаточной ПСП** для η и определяемую следующим образом:

$$\forall\, u' \in X^*, \; \forall\, v' \in Y^* \quad \eta_{u,v}(u', v') \stackrel{\text{def}}{=} \tfrac{\eta(uu', vv')}{\eta(u,v)}.$$

$\forall\, \eta \in R(\mathbf{1}, X \times Y)$ запись S_η обозначает множество всех остаточных ПСП для η. Отметим, что $\eta \in S_\eta$, т.к. $\eta = \eta_{\varepsilon,\varepsilon}$.

Теорема 13.
Пусть X, Y – пара конечных множеств, и $\eta \in R(\mathbf{1}, X \times Y)$.
Тогда следующие условия эквивалентны:

1. $|S_\eta| < \infty$,

2. $|S_{\eta^X}| < \infty$, и существуют

   - ВА A вида (X, Y, S, P, ξ_{s^0}),
   - функция $\delta : S \times (X \times Y) \to S$, и
   - совокупность $\{f_s \mid s \in S\}$ функций вида $X \times Y \to [0,1]$,

   удовлетворяющие условиям:

$$\begin{array}{l} \forall\, s, s' \in S,\ \forall\, x \in X,\ \forall\, y \in Y \\ P(s, x, s', y) \stackrel{\text{def}}{=} \begin{cases} f_s(x, y), & \text{если } s' = s(x, y), \\ 0, & \text{иначе}, \end{cases} \end{array} \tag{2.64}$$

и

$$\begin{array}{l} \forall\, x \in X,\ \forall\, y \in Y,\ \forall\, u \in X^*,\ \forall\, v \in Y^* \\ \eta(ux, vy) = \eta(u, v) \cdot f_{s^0(u,v)}(x, y), \end{array} \tag{2.65}$$

где $\forall\, s \in S$, $\forall\, x \in X$, $\forall\, y \in Y$, $\forall\, u \in X^*$, $\forall\, v \in Y^*$

$$s(\varepsilon, \varepsilon) \stackrel{\text{def}}{=} s, \quad s(ux, vy) \stackrel{\text{def}}{=} \delta(s(u, v), (x, y)). \quad \blacksquare$$

(отметим, что из (2.64) и (2.65) следует равенство $f_A = \eta$).

2.4.4 Автоматные преобразования случайных последовательностей

Для любой СП $\zeta \in R(\mathbf{1}, X)$ и любого ВА A вида $(X, Y, \ldots)$ запись $\eta_{\zeta,A}$ обозначает функцию из $[0,1]^{X^* \times Y^*}$, определяемую следующим образом: $\forall\, u \in X^*$, $\forall\, v \in Y^*$

$$\eta_{\zeta,A}(u, v) \stackrel{\text{def}}{=} \zeta(u) f_A(u, v). \tag{2.66}$$

Теорема 14.
$\forall\,\zeta \in R(\mathbf{1}, X)$, $\forall$ ВА A вида $(X, Y, \ldots)$

$$\eta_{\zeta,A} \in R(\mathbf{1}, X \times Y).$$

Доказательство.
Докажем, что выполнены условия (2.62) для $\eta \stackrel{\text{def}}{=} \eta_{\zeta,A}$.

- $\eta_{\zeta,A}(\varepsilon, \varepsilon) = \zeta(\varepsilon) f_A(\varepsilon, \varepsilon) = 1 \cdot 1 = 1$.
- Если $|u| \neq |v|$, то $\eta_{\zeta,A}(u, v) = \zeta(u) f_A(u, v) = \zeta(u) \cdot 0 = 0$.
- Равенство $\eta_{\zeta,A}(u, v) = \sum\limits_{x \in X, y \in Y} \eta_{\zeta,A}(ux, vy)$, т.е.

  $$\zeta(u) f_A(u, v) = \sum_{x \in X, y \in Y} \zeta(ux) f_A(ux, vy) \tag{2.67}$$

  верно потому, что, согласно (2.60), левая часть (2.67) равна

  $$\zeta(u) f_A(u, v) = \Big(\sum_{x \in X} \zeta(ux) \Big) f_A(u, v) = \sum_{x \in X} \Big(\zeta(ux) f_A(u, v) \Big) \tag{2.68}$$

  и, поскольку $f_A \in R(X, Y)$, то, согласно (2.27),

  $$\forall\, x \in X \quad f_A(u, v) = \sum_{y \in Y} f_A(ux, vy),$$

  поэтому правая часть (2.68) равна

  $$\sum_{x \in X} \Big(\zeta(ux) f_A(u, v) \Big) = \sum_{x \in X} \Big((\sum_{y \in Y} \zeta(ux) f_A(ux, vy)) \Big). \tag{2.69}$$

  Нетрудно видеть, что правая часть (2.69) совпадает с правой частью (2.67). $\blacksquare$

С $\eta_{\zeta,A}$ связаны две СП, определяемые соотношениями (2.63):

- $\eta^X_{\zeta,A}$, нетрудно видеть, что эта СП совпадает с ζ, и
- $\eta^Y_{\zeta,A}$, мы будем обозначать эту СП записью ζ_A, и называть её **результатом преобразования СП ζ вероятностным автоматом** A.

Будем говорить, что СП ζ и ζ' **эквивалентны** (и обозначать это записью $\zeta \sim \zeta'$), если $\exists$ ВА A и B, такие, что $\zeta' = \zeta_A$ и $\zeta = \zeta'_B$.

Теорема 15.

1. Если СП $\zeta \in R(\mathbf{1}, X)$ такова, что множество S_ζ конечно, то $\forall$ ВА $A = (X, \ldots)$ множество S_{ζ_A} конечно.

2. Если СП ζ и ζ' таковы, что множества S_ζ и $S_{\zeta'}$ конечны, то $\exists$ ВА $A : \zeta' = \zeta_A$.

3. Для любых СП $\zeta \in R(\mathbf{1}, X)$ и $\zeta' \in R(\mathbf{1}, Y)$ следующие условия эквивалентны:

   (a) $\zeta \sim \zeta'$

   (b) $\exists$ детерминированный ВА A: $\zeta' = \zeta_A$ и его реакция f_A биективно отображает D_ζ на $D_{\zeta'}$.

Доказательство.

Обоснуем лишь импликацию 3b $\Rightarrow$ 3a. Из 3b следует существование детерминированного ВА B, такого, что ограничения f_A и f_B на D_ζ и $D_{\zeta'}$ соответственно являются взаимно обратными отображениями, откуда нетрудно вывести равенство $\zeta'_B = \zeta$:

$$\forall\, u \in X^* \;\; \zeta'_B(u) = \sum_{v \in Y^*} \zeta'(v) f_B(v, u) = \sum_{v \in D_{\zeta'}} \zeta'(v) f_B(v, u) =$$
$$\sum_{v \in f_B^{-1}(u)} \zeta'(v) = \sum_{v \in f_B^{-1}(u)} \zeta_A(v) = \sum_{v \in f_B^{-1}(u)} \sum_{r \in X^*} \eta_{\zeta, A}(r, v) =$$
$$= \sum_{v \in f_B^{-1}(u)} \sum_{r \in X^*} \zeta(r) \cdot f_A(r, v) = \sum_{v \in f_B^{-1}(u)} \sum_{r \in D_\zeta} \zeta(r) \cdot f_A(r, v) =$$
$$= \sum_{v \in f_B^{-1}(u)} \sum_{r \in f_A^{-1}(v)} \zeta(r) = \sum_{r \in f_A^{-1} f_B^{-1}(u)} \zeta(r) = \zeta(u). \quad \blacksquare$$

2.4.5 Цепи Маркова

Один из способов задания СП связан с понятием цепи Маркова.

Цепь Маркова (ЦМ) – это пятёрка M вида

$$M = (X, S, P, \lambda, \xi^0),$$

компоненты которой имеют следующий смысл.

1. X и S – конечные множества, элементы которых называются соответственно **сигналами** и **состояниями** ЦМ M.

2. P – СФ вида $S \underset{r}{\to} S$, называемая **функцией перехода**. $\forall\,(s, s') \in S \times S$ значение $P(s, s')$ понимается как вероятность того, что если в текущий момент времени (t) M находится в состоянии s, то в следующий момент времени $(t+1)$ M будет находиться в состоянии s'.

3. λ – функция вида $S \to X$, $\forall\, s \in S$ $\lambda(s)$ понимается как сигнал, который ЦМ M выдаёт в текущий момент времени, если в этот момент времени M находится в состоянии s.

4. ξ^0 – распределение на S, называемое **начальным распределением**. $\forall\, s \in S$ s^{ξ^0} понимается как вероятность того, что в начальный момент времени $(t = 0)$ ЦМ M находится в состоянии s.

Функция ЦМ $M = (X, \ldots)$ – это функция $f_M \in [0,1]^{X^*}$, значение которой

- на пустой строке равно 1, и
- на непустой строке $u = x_0 \ldots x_k \in X^*$ равно вероятности того, что в моменты $0, \ldots, k$ сигналы, выдаваемые M, совпадают с $x_0, \ldots, x_k$, соответственно.

Пусть $M = (X, S, P, \lambda, \xi^0)$ – ЦМ, и упорядочение множества S её состояний имеет вид $(s_1, \ldots, s_n)$. Мы будем использовать следующие обозначения.

- Будем обозначать тем же символом M матрицу порядка n

$$\begin{pmatrix} P(s_1, s_1) & \ldots & P(s_1, s_n) \\ \ldots & \ldots & \ldots \\ P(s_n, s_1) & \ldots & P(s_n, s_n) \end{pmatrix}. \tag{2.70}$$

- Для любого $x \in X$ будем обозначать записью E^x квадратную матрицу порядка n, каждый элемент которой равен 0 или 1, и элемент E^x в строке i и столбце j равен 1 тогда и только тогда, когда $i = j$ и $\lambda(s_i) = x$.

Теорема 16.

Пусть M – ЦМ вида $(X, S, P, \lambda, \xi^0)$. Если строка $u \in X^*$ непуста и имеет вид $x_0 \dots x_k$, то

$$f_M(u) = \xi^0 E^{x_0} M E^{x_1} \dots M E^{x_k} I, \tag{2.71}$$

где I – столбец порядка n, все элементы которого равны 1.

Доказательство.

Равенство (2.71) следует из правил вычисления вероятностей несовместных и независимых событий. ■

Теорема 17.

Если M – ЦМ, то $f_M \in R(\mathbf{1}, X)$.

Доказательство.

Первое соотношение в (2.60) выполняется по определению, а второе соотношение в (2.60) для $\zeta \stackrel{\text{def}}{=} f_M$ следует из (2.71), равенства $MI = I$ и того, что $\sum\limits_{x \in X} E^x$ является единичной матрицей. ■

Теорема 18.

Пусть X – конечное множество, и $f \in [0,1]^{X^*}$. Тогда следующие утверждения эквивалентны:

1. $\exists$ ЦМ $M : f = f_M$,
2. $\exists$ ВА A вида $(\mathbf{1}, X, \dots) : f = f_A \circ in$,
   где $in : X^* \to \mathbf{1}^* \times X^*, \forall\, u \in X^* \;\; in(u) = (e^{|u|}, u)$,
3. $\exists \zeta \in R(\mathbf{1}, Y) : \;\forall\, y_1, \dots, y_n \in Y \quad \zeta(y_1 \dots y_n) = \zeta(y_1) \dots \zeta(y_n)$,
   и $\exists$ детерминированный ВА A: $f = \zeta_A$. ■

Глава 3

Вероятностные автоматы Мура с числовым выходом

3.1 Вероятностные автоматы Мили и Мура

С каждым ВА $A = (X, Y, S, P, \xi^0)$ связаны СФ

$$\delta : S \times X \underset{r}{\rightarrow} S, \quad \lambda : S \times X \underset{r}{\rightarrow} Y,$$

называемые соответственно **функцией перехода** и **функцией выхода**, и определяемые следующим образом:

- $\forall\, s, s' \in S, x \in X \quad \delta(s, x, s') \stackrel{\text{def}}{=} \sum\limits_{y \in Y} P(s, x, s', y),$
- $\forall\, s \in S, x \in X, y \in Y \quad \lambda(s, x, y) \stackrel{\text{def}}{=} \sum\limits_{s' \in S} P(s, x, s', y).$

ВА $A = (X, Y, S, P, \xi^0)$ называется **ВА Мили**, если

$$\forall\, s, s' \in S, x \in X, y \in Y \quad P(s, x, s', y) = \delta(s, x, s') \cdot \lambda(s, x, y).$$

ВА $A = (X, Y, S, P, \xi^0)$ называется **ВА Мура**, если он является ВА Мили, и его функция выхода λ не зависит от x (т.е. имеет вид $\lambda : S \underset{r}{\rightarrow} Y$).

Пусть $A = (X, Y, S, P, \xi^0)$ – ВА Мура, и упорядочение множества S его состояний имеет вид $(s_1, \ldots, s_n)$. $\forall\, x \in X$ мы будем обозначать записью A^x матрицу порядка n, называемую **матрицей перехода**, соответствующей входному сигналу x и имеющую вид

$$\begin{pmatrix} \delta(s_1, x, s_1) & \dots & \delta(s_1, x, s_n) \\ \dots & \dots & \dots \\ \delta(s_n, x, s_1) & \dots & \delta(s_n, x, s_n) \end{pmatrix}, \qquad (3.1)$$

где δ – функция перехода ВА A.

$\forall\, x \in X$, $\forall\, s, s' \in S$ мы будем обозначать записью $A^x_{s,s'}$ элемент матрицы A^x, находящийся в строке s столбце s' (т.е. $A^x_{s,s'} = \delta(s, x, s')$). Из того, что δ – СФ, следует, что элементы матрицы A^x обладают свойствами

$$\begin{array}{ll} \forall\, s, s' \in S & A^x_{s,s'} \geq 0, \\ \forall\, s \in S & \sum\limits_{s' \in S} A^x_{s,s'} = 1 \ \ (\text{т.е. } A^x I = I). \end{array} \qquad (3.2)$$

(Матрица, обладающая такими свойствами, называется **стохастической**.)

$\forall\, u \in X^*$ мы будем обозначать записью A^u матрицу порядка n, определяемую следующим образом: $A^\varepsilon \stackrel{\text{def}}{=} E$, и если $u = x_1 \dots x_k$, то $A^u \stackrel{\text{def}}{=} A^{x_1} \dots A^{x_k}$.

Нетрудно доказать, что матрица A^u – стохастическая.

Пусть s – произвольное состояние из S, и в упорядочении элементов S данное состояние имеет номер i (т.е. $s = s_i$). Будем называть

- строку номер i матрицы A^u – **строкой** s, и обозначать её записью $\vec{A}^u_s$
- столбец номер i матрицы A^u – **столбцом** s, и обозначать его записью $A^{u\downarrow}_s$.

$\forall\, u \in X^*$, $\forall\, s, s' \in S$ мы будем обозначать записью $A^u_{s,s'}$ элемент матрицы A^u, находящийся в строке s столбце s'.

Если строка $u \in X^*$ имеет вид $x_0 \dots x_k$, то $A^u_{s,s'}$ можно понимать как вероятность того, что

- если в текущий момент (t) ВА A находится в состоянии s, и, начиная с этого момента, на вход A последовательно поступают элементы строки u (т.е. в момент t поступил сигнал x_0, в момент $t+1$ поступил сигнал x_1, и т.д.)

- то в момент $t+k+1$ A будет находиться в состоянии s'.

ВА Мура с детерминированным выходом – это ВА Мура (X, Y, S, P, ξ^0), функция выходов λ которого является детерминированной (т.е. можно считать, что λ имеет вид $S \to Y$).

Если ВА $A = (X, Y, S, P, \xi^0)$ является ВА Мура с детерминированным выходом, то для обозначения такого ВА мы будем использовать запись

$$(\xi^0, \{A^x \mid x \in X\}, \lambda), \tag{3.3}$$

компоненты которой определяются следующим образом.

- Компонента ξ^0 в (3.3) является вектор-строкой, соответствующей начальному распределению ξ^0 ВА A.
- Компонента $\{A^x \mid x \in X\}$ в (3.3) является совокупностью матриц перехода ВА A.
- Компонента λ в (3.3) является вектор-столбцом $\begin{pmatrix} \lambda(s_1) \\ \dots \\ \lambda(s_n) \end{pmatrix}$ значений функции выхода $\lambda : S \to Y$ ВА A (где $(s_1, \dots, s_n)$ – фиксированное упорядочение множества S.)

Если A – ВА Мура с детерминированным выходом, то запись S_A обозначает множество состояний этого ВА.

3.2 Вероятностные автоматы Мура с числовым выходом

3.2.1 Понятие вероятностного автомата Мура с числовым выходом

ВА Мура с числовым выходом – это ВА Мура с детерминированным выходом, множество выходных сигналов которого является подмножеством множества $\mathbf{R}$ действительных чисел.

Для ВА Мура с числовым выходом можно определить понятие реакции, отличное от того понятия реакции ВА, которое было определено в пункте 2.1.3. Мы будем называть это понятие **усреднённой реакцией**.

3.2.2 Усреднённые реакции

Пусть A – ВА Мура с числовым выходом вида (3.3), и $\xi \in S_A^{\triangle}$.

Усреднённая реакция ВА A в распределении ξ – это функция $A^\xi : X^* \to \mathbf{R}$, определяемая следующим образом:

$$\forall\, u \in X^* \;\; A^\xi(u) \stackrel{\text{def}}{=} \xi A^u \lambda.$$

Усреднённую реакцию ВА A в его начальном распределении мы будем называть просто **усреднённой реакцией** ВА A, и будем обозначать её записью f_A.

Если строка $u \in X^*$ имеет вид $x_0 \ldots x_k$, то значение $A^\xi(u)$ можно понимать следующим образом:

- если A в некоторый момент времени t имеет распределение ξ, и, начиная с этого момента, на его вход последовательно поступают элементы строки u (т.е. в момент t поступает сигнал x_0, в момент $t+1$ поступает сигнал x_1, и т.д.),
- то $A^\xi(u)$ – это среднее значение (т.е. математическое ожидание) выходного сигнала A в момент времени $t+k+1$.

Мы будем говорить, что распределения $\xi_1, \xi_2 \in S_A^{\triangle}$ **эквивалентны по усреднению относительно** A (и обозначать это записью $\xi_1 \underset{A}{\approx} \xi_2$), если усреднённые реакции A^{ξ_1} и A^{ξ_2} совпадают, т.е.

$$\forall\, u \in X^* \qquad \xi_1 A^u \lambda = \xi_2 A^u \lambda.$$

Пусть задана пара A_1, A_2 ВА Мура с числовым выходом, у которых одинаковы множества входных сигналов, т.е. A_1 и A_2 имеют вид

$$A_i = (\xi_i^0, \{A_i^x \mid x \in X\}, \lambda_i) \quad (i = 1, 2).$$

Мы будем говорить, что A_1 и A_2 **эквивалентны по усреднению** (и обозначать это записью $A_1 \approx A_2$), если их усреднённые реакции совпадают, т.е.

$$\forall\, u \in X^* \qquad \xi_1^0 A_1^u \lambda_1 = \xi_2^0 A_2^u \lambda_2.$$

3.2.3 Усреднённые базисные матрицы

Для ВА Мура с числовым выходом можно ввести понятие усреднённой базисной матрицы, аналогично тому, как было введено понятие базисной матрицы в пункте 2.1.4.

Пусть $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ – ВА Мура с числовым выходом. Обозначим записью $\hat{A}$ совокупность всех вектор-столбцов вида $A^u\lambda$, где $u \in X^*$.

Усреднённой базисной матрицей ВА A называется матрица, обозначаемая записью $[\![A]\!]$, и удовлетворяющая условиям:

- каждый столбец матрицы $[\![A]\!]$ является элементом $\hat{A}$,
- столбцы матрицы $[\![A]\!]$ образуют базис линейного пространства $\langle\hat{A}\rangle$.

Нетрудно видеть, что

$$\forall\, \xi_1, \xi_2 \in S_A^{\triangle} \quad \xi_1 \underset{A}{\approx} \xi_2 \quad \Leftrightarrow \quad \xi_1[\![A]\!] = \xi_2[\![A]\!].$$

Матрицу $[\![A]\!]$ можно построить при помощи алгоритма, аналогичного соответствующему алгоритму из пункта 2.1.4.

Для каждого $s \in S_A$ мы будем называть **строкой** s матрицы $[\![A]\!]$ ту её строку, которая содержит значения вида $\vec{A}_s^u\lambda$. Мы будем обозначать эту строку записью $[\![A]\!]_s$.

Мы будем говорить, что состояние $s \in S_A$ является **выпуклой комбинацией** других состояний ВА A, если строка s матрицы $[\![A]\!]$ является выпуклой комбинацией других строк этой матрицы, т.е. существует распределение $\xi \in (S_A \setminus \{s\})^{\triangle}$, удовлетворяющее условию

$$[\![A]\!]_s = \sum_{s' \in S_A \setminus \{s\}} (s')^{\xi} [\![A]\!]_{s'}. \tag{3.4}$$

3.2.4 Редукция вероятностных автоматов Мура с числовым выходом

Пусть A – ВА Мура с числовым выходом.

Редукция ВА A заключается в построении такого ВА Мура с числовым выходом B, который

- был бы эквивалентен A по усреднению, и
- содержал бы меньше состояний, чем A (если это возможно).

К вероятностным автоматам Мура с числовым выходом можно применять те же методы редукции, которые были изложены в пункте 2.2. Мы рассмотрим лишь метод редукции путем удаления выпуклых комбинаций. Данный метод основан на нижеследующей теореме.

Теорема 19.

Пусть $A = (\xi_A^0, \{A^x \mid x \in X\}, \lambda_A)$ – ВА Мура с числовым выходом, и состояние $s \in S_A$ является выпуклой комбинацией других состояний, т.е. $\exists \xi \in (S_A \setminus \{s\})^\triangle$: верно (3.4). Будем считать, что упорядочение S_A имеет вид $(s_1, \ldots, s_n)$ и $s = s_n$.

Обозначим символом B ВА Мура с числовым выходом, который имеет вид $(\xi_B^0, \{B^x \mid x \in X\}, \lambda_B)$, где $S_B = S_A \setminus \{s_n\}$, и

$$\begin{aligned} &\xi_B^0 = \xi_A^0 \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix}, \\ &B^x = (E_{n-1} \ \ \mathbf{0}) A^x \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix}, \\ &\lambda_B = (E_{n-1} \ \ \mathbf{0}) \lambda_A, \end{aligned} \tag{3.5}$$

где E_{n-1} – единичная матрица порядка $n-1$, $\xi = (s_1^\xi, \ldots, s_{n-1}^\xi)$, $\mathbf{0}$ – столбец порядка $n-1$ с нулевыми компонентами.

Тогда $A \approx B$.

Доказательство.

С учетом предположений, изложенных в формулировке теоремы, можно переписать (3.4) в виде

$$[\![A]\!]_{s_n} = \sum_{i=1}^{n-1} s_i^\xi [\![A]\!]_{s_i}. \tag{3.6}$$

Согласно определению матрицы $[\![A]\!]$, равенство (3.6) равносильно соотношению

$$\forall\, u \in X^* \quad A_n^u \lambda_A = \sum_{i=1}^{n-1} s_i^\xi A_i^u \lambda_A, \tag{3.7}$$

где $\forall\, i = 1, \ldots, n$ A^u_i обозначает i–ю строку матрицы A^u.

Можно переписать (3.7) в матричном виде:

$$\forall\, u \in X^* \quad \vec{e}_n A^u \lambda_A = \xi(E_{n-1}\ \ \mathbf{0}) A^u \lambda_A \tag{3.8}$$

(где $\vec{e}_n$ – вектор-строка длины n, у которой n–я компонента равна 1, а остальные компоненты равны 0).

В частности, (3.8) верно при $u = \varepsilon$, т.е.

$$\vec{e}_n \lambda_A = \xi(E_{n-1}\ \ \mathbf{0})\lambda_A. \tag{3.9}$$

Докажем, что $A \approx B$, т.е. $\forall\, u \in X^*$ $\xi^0_A A^u \lambda_A = \xi^0_B B^u \lambda_B$. Согласно (3.5), для этого достаточно доказать, что $\forall\, u \in X^*$

$$A^u \lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} B^u (E_{n-1}\ \ \mathbf{0})\lambda_A. \tag{3.10}$$

Докажем (3.10) индукцией по $|u|$.

Если $u = \varepsilon$, то (3.10) имеет вид

$$\lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} (E_{n-1}\ \ \mathbf{0})\lambda_A. \tag{3.11}$$

Согласно правилам матричного умножения, и учитывая (3.9), можно переписать правую часть (3.11) следующим образом:

$$\begin{pmatrix} E_{n-1}(E_{n-1}\ \ \mathbf{0})\lambda_A \\ \xi(E_{n-1}\ \ \mathbf{0})\lambda_A \end{pmatrix} = \begin{pmatrix} (E_{n-1}\ \ \mathbf{0})\lambda_A \\ \vec{e}_n \lambda_A \end{pmatrix} = \lambda_A.$$

Таким образом, в случае $u = \varepsilon$ равенство (3.10) верно.

Пусть (3.10) верно для некоторого u. Докажем, что $\forall\, x \in X$

$$A^{xu} \lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} B^{xu} (E_{n-1}\ \ \mathbf{0})\lambda_A. \tag{3.12}$$

Используя определение B^x из (3.5), перепишем (3.12) в виде

$$A^{xu} \lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} (E_{n-1}\ \ \mathbf{0}) A^x \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} B^u (E_{n-1}\ \ \mathbf{0})\lambda_A. \tag{3.13}$$

Учитывая индуктивное предположение (3.10), и используя правила матричного умножения перепишем (3.13) в виде

$$
\begin{aligned}
A^{xu}\lambda_A = \begin{pmatrix} E_{n-1}(E_{n-1}\ \ \mathbf{0}) \\ \xi(E_{n-1}\ \ \mathbf{0}) \end{pmatrix} A^x A^u \lambda_A = \\
= \begin{pmatrix} (E_{n-1}\ \ \mathbf{0}) \\ \xi(E_{n-1}\ \ \mathbf{0}) \end{pmatrix} A^{xu}\lambda_A = \begin{pmatrix} (E_{n-1}\ \ \mathbf{0})A^{xu}\lambda_A \\ \xi(E_{n-1}\ \ \mathbf{0})A^{xu}\lambda_A \end{pmatrix}.
\end{aligned} \tag{3.14}
$$

Используя (3.8), перепишем правую часть (3.14) в виде

$$
\begin{pmatrix} (E_{n-1}\ \ \mathbf{0})A^{xu}\lambda_A \\ \vec{e}_n A^{xu}\lambda_A \end{pmatrix} = \begin{pmatrix} (E_{n-1}\ \ \mathbf{0}) \\ \vec{e}_n \end{pmatrix} A^{xu}\lambda_A = A^{xu}\lambda_A. \tag{3.15}
$$

Таким образом, (3.12) верно.

Следовательно, (3.10) верно для любого $u \in X^*$. ■

Отметим, что задачу распознавания того, является ли какое-либо из состояний ВА Мура с числовым выходом выпуклой комбинацией других состояний этого ВА, можно решать методом, аналогичным методу, изложенному в пункте 2.2.3.

3.2.5 Соглашение

Начиная со следующего пункта и до конца книги, все рассматриваемые ВА по умолчанию (т.е. если их вид не указан особо) предполагаются ВА Мура с числовым выходом. Мы будем обозначать эти ВА записями вида (3.3), и для каждого такого ВА A запись f_A обозначает его усреднённую реакцию (которую мы будем называть просто **реакцией**). Те ВА, которые были введены в главе 2, мы будем называть **ВА общего вида**. Для каждого рассматриваемого ВА A запись S_A обозначает множество состояний этого ВА.

3.3 Вероятностная реализуемость функций на строках

Пусть X – конечное множество.

Функция на строках $f \in \mathbf{R}^{X^*}$ называется **вероятностно реализуемой**, если $\exists$ ВА, реакция которого совпадает с f.

Будем использовать следующие определения и обозначения.

- Для каждого $\Gamma \subseteq \mathbf{R}^{X^*}$ **выпуклой оболочкой** множества Γ называется подмножество $C(\Gamma) \subseteq \mathbf{R}^{X^*}$, состоящее из функций вида $\sum\limits_{i=1}^{n} a_i f_i$ (называемых **выпуклыми комбинациями** функций $f_1, \ldots, f_n$), где
  - $\forall\, i = 1, \ldots, n \quad a_i \in [0,1],\ f_i \in \Gamma, \quad \sum\limits_{i=1}^{n} a_i = 1$, и
  - $\forall\, u \in X^* \quad \Big(\sum\limits_{i=1}^{n} a_i f_i\Big) u \stackrel{\text{def}}{=} \sum\limits_{i=1}^{n} a_i f_i(u)$.
- $\forall\, x \in X$ запись D^x обозначает отображение вида
$$D^x : \mathbf{R}^{X^*} \to \mathbf{R}^{X^*},$$
называемое **сдвигом**, сопоставляющее каждой функции f из $\mathbf{R}^{X^*}$ функцию, обозначаемую записью fD^x, где
$$\forall\, u \in X^* \quad (fD^x)(u) \stackrel{\text{def}}{=} f(xu). \tag{3.16}$$
- Подмножество $\Gamma \subseteq \mathbf{R}^{X^*}$ называется **устойчивым относительно сдвигов**, если
$$\forall\, f \in \Gamma,\ \forall\, x \in X \quad fD^x \in C(\Gamma).$$

Теорема 20.

Пусть X – конечное множество, и $f \in \mathbf{R}^{X^*}$. Следущие условия эквивалентны:

- f вероятностно реализуема,
- $\exists$ конечное $\Gamma \subseteq \mathbf{R}^{X^*}$, устойчивое относительно сдвигов, и такое, что $f \in C(\Gamma)$.

Доказательство.

Пусть f вероятностно реализуема, т.е. $\exists$ ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$:
$$\forall\, u \in X^* \quad f(u) = \xi^0 A^u \lambda.$$

Будем считать, что множество состояний S_A этого ВА имеет вид $\{1, \ldots, n\}$, и $\forall\, i \in S_A$ запись ξ_i обозначает распределение из $S_A^{\triangle}$, представляемое вектор-строкой порядка n, i–я компонента которой равна 1, а остальные компоненты равны 0.

$\forall\, i = 1, \ldots, n$ определим $A_i \stackrel{\text{def}}{=} (\xi_i, \{A^x \mid x \in X\}, \lambda)$.

Полагаем $\Gamma \stackrel{\text{def}}{=} \{f_{A_i} \mid i = 1, \ldots, n\}$. $f \in C(\Gamma)$, т.к. $f = \sum\limits_{i=1}^{n} i^{\xi^0} f_{A_i}$.

Докажем, что $\forall\, i = 1, \ldots, n$, $\forall\, x \in X \quad f_{A_i} D^x \in C(\Gamma)$.

Согласно (3.16),

$$\forall\, u \in X^* \quad (f_{A_i} D^x)(u) = f_{A_i}(xu) = \xi_i A^{xu} \lambda = \xi_i A^x A^u \lambda. \tag{3.17}$$

Нетрудно видеть, что

$$\xi_i A^x A^u \lambda = \sum_{j=1}^{n} A^x_{ij} f_{A_j}(u).$$

Таким образом, $f_{A_i} D^x = \sum\limits_{j=1}^{n} A^x_{ij} f_{A_j} \in C(\Gamma)$.

Обратно, пусть $f \in C(\Gamma)$, где $\Gamma = \{f_1, \ldots, f_n\}$, и Γ устойчиво относительно сдвигов. Определим ВА

$$A \stackrel{\text{def}}{=} (\xi^0, \{A^x \mid x \in X\}, \lambda), \tag{3.18}$$

где

- ξ^0 – вектор-строка коэффициентов представления f в виде выпуклой комбинации функций из Γ, т.е.

$$f = \sum_{i=1}^{n} i^{\xi^0} f_i, \tag{3.19}$$

- $\forall\, x \in X$, $\forall\, i = 1, \ldots, n$ строка i матрицы A^x состоит из коэффициентов представления функции $f_i D^x$ в виде выпуклой комбинации функций из Γ, т.е.

$$f_i D^x = \sum_{j=1}^{n} A^x_{ij} f_j, \tag{3.20}$$

- $\lambda \stackrel{\text{def}}{=} \begin{pmatrix} f_1(\varepsilon) \\ \ldots \\ f_n(\varepsilon) \end{pmatrix}$.

Докажем, что реакция ВА (3.18) совпадает с f, т.е.

$$\forall\, u \in X^* \quad \xi^0 A^u \lambda = f(u). \tag{3.21}$$

Для этого сначала докажем (индукцией по $|u|$), что

$$A^u\lambda = \begin{pmatrix} f_1(u) \\ \dots \\ f_n(u) \end{pmatrix}. \tag{3.22}$$

Если $u = \varepsilon$, то обе части (3.22) совпадают по определению λ.

Если $u = xu'$, то, предполагая верным равенство (3.22), в котором u заменено на u', имеем:

$$\begin{gathered} A^u\lambda = A^{xu'}\lambda = A^x A^{u'}\lambda = A^x \begin{pmatrix} f_1(u') \\ \dots \\ f_n(u') \end{pmatrix} = \\ = \begin{pmatrix} \sum\limits_{i=1}^{n} A^x_{1j} f_j(u') \\ \dots \\ \sum\limits_{i=1}^{n} A^x_{nj} f_j(u') \end{pmatrix}. \end{gathered} \tag{3.23}$$

Из (3.20) следует, что правую часть в (3.23) можно переписать в виде

$$\begin{pmatrix} (f_1 D^x)(u') \\ \dots \\ (f_n D^x)(u') \end{pmatrix}. \tag{3.24}$$

Согласно определению (3.16) функций вида fD^x, столбец (3.24) совпадает с правой частью доказываемого равенства (3.22).

Таким образом, равенство (3.22) доказано. Согласно этому равенству, левая часть доказываемого равенства (3.21) равна

$$\xi^0 \begin{pmatrix} f_1(u) \\ \dots \\ f_n(u) \end{pmatrix} = \sum_{i=1}^{n} i^{\xi^0} f_i(u). \tag{3.25}$$

По определению ξ^0 (см. (3.19)), правая часть (3.25) равна $f(u)$, т.е. правой части доказываемого равенства (3.21). ■

3.4 Связь между линейно-автоматными функциями и реакциями вероятностных автоматов

Теорема 21.

Пусть $L = (\xi_L^0, \{L^x \mid x \in X\}, \lambda_L)$ – ЛА размерности n.

Тогда существуют ВА A и число $a > 0$, такие, что $\forall\, u \in X^*$

$$f_A(u) = a^{|u|+1} f_L(u) + \tfrac{1}{n+2}. \tag{3.26}$$

Доказательство.

Если все компоненты λ_L равны нулю, то все значения функции f_L равны нулю, в этом случае искомый ВА может иметь вид

$$(\vec{e}_1, \{E \mid x \in X\}, \tfrac{1}{n+2} e_1^{\downarrow}),$$

где $\vec{e}_1$ и $e_1^{\downarrow}$ – вектор-строка и вектор-столбец соответственно, у которых первая компонента равна 1, а остальные равны 0.

Пусть не все компоненты λ_L равны нулю. Можно считать, что $\lambda_L = e_1^{\downarrow}$ (а если $\lambda_L \neq e_1^{\downarrow}$, то заменим L на эквивалентный ему ЛА

$$\left(\xi_L^0 P, \{P^{-1} L^x P \mid x \in X\}, e_1^{\downarrow}\right),$$

где P – обратимая матрица, первый столбец которой равен λ_L).

$\forall\, x \in X$ определим A_1^x как матрицу порядка $n+2$ вида

$$\begin{pmatrix} & & & a_1 & 0 \\ & L^x & & \dots & \dots \\ & & & a_n & 0 \\ 0 & \dots & 0 & 0 & 0 \\ b_1 & \dots & b_n & b_0 & 0 \end{pmatrix}$$

в которой

- левая верхняя подматрица порядка n совпадает с L^x, и
- компоненты $a_1, \dots, a_n$ и $b_1, \dots, b_n, b_0$ выбраны так, что сумма компонентов в каждой строке и в каждом столбце матрицы A_1^x равна нулю.

Второе из этих свойств можно выразить в виде равенств

$$A_1^x I = 0^{\downarrow}, \quad \tilde{I} A_1^x = \vec{0}, \tag{3.27}$$

где I и $\tilde{I}$ – вектор-столбец и вектор-строка порядка $n+2$, каждый элемент которых равен 1, и $0^{\downarrow}$ и $\vec{0}$ – вектор-столбец и вектор-строка порядка $n+2$, каждый элемент которых равен 0.

Нетрудно видеть, что $\forall\, u \neq \varepsilon$ левая верхняя подматрица порядка n матрицы A_1^u совпадает с L^u, и все компоненты строки $n-1$ и столбца n матрицы A_1^u равны нулю. Кроме того, будут верны равенства (3.27), в которых x заменено на u.

Обозначим записью ξ_1^0 вектор-строку $(\xi_L^0, c, 0)$ порядка $n+2$, в которой левая подстрока порядка n совпадает с ξ_L^0, и число c выбрано так, что сумма компонентов ξ_1^0 равна нулю (т.е. $\xi_1^0 I = 0$).

Выберем число $a > 0$ так, чтобы модуль каждого элемента строки $a\xi_1^0$ и матрицы aA_1^x был меньше $\frac{1}{n+2}$.

Определим матрицу B порядка $n+2$ и вектор-строку ξ порядка $n+2$ следующим образом:

$$B \stackrel{\text{def}}{=} \frac{1}{n+2}\begin{pmatrix} 1 & \dots & 1 \\ \dots & \dots & \dots \\ 1 & \dots & 1 \end{pmatrix}, \quad \xi \stackrel{\text{def}}{=} \frac{1}{n+2}(1 \dots 1).$$

Нетрудно видеть, что

$$\begin{array}{l} B^2 = B,\ BI = I,\ \xi I = 1,\ \xi B = \xi, \\ A_1^x B = 0,\ BA_1^x = 0,\ \xi_1^0 B = 0,\ \xi A_1^x = 0, \end{array} \tag{3.28}$$

где символ 0 в (3.28) обозначает нулевую матрицу или вектор-строку порядка $n+2$.

Искомый ВА имеет вид $A = (\xi_A^0, \{A^x \mid x \in X\}, \lambda_A)$, где

$$\xi_A^0 := a\xi_1^0 + \xi, \quad \forall\, x \in X \quad A^x := aA_1^x + B, \quad \lambda_A := \begin{pmatrix} \lambda_L \\ 0 \\ 0 \end{pmatrix}.$$

Нетрудно видеть, что все элементы вектор-строки ξ_A^0 и матриц A^x положительны, $\xi_A I = 1$, и $\forall\, x \in X \quad A^x I = I$ (т.е. ξ_A – распределение, и $\forall\, x \in X$ матрица A^x определяет СФ).

Докажем, что верно равенство (3.26).

Если $u = \varepsilon$, то левая часть (3.26) равна

$$a\xi_L^0\lambda_L + \xi\lambda_A = a\xi_L^0\lambda_L + \tfrac{1}{n+2},$$

что равно правой части (3.26).

Если $u \neq \varepsilon$, то из (3.28) следует, что $A^u = a^{|u|}A_1^u + B$, и, используя (3.28), получаем:

$$\begin{aligned} &f_A(u) = \xi_A^0 A^u\lambda_A = (a\xi_1^0 + \xi)(a^{|u|}A_1^u + B)\lambda_A = \\ &= (a^{|u|+1}\xi_1^0 A_1^u + \xi)\lambda_A = a^{|u|+1}\xi_1^0 A_1^u\lambda_A + \tfrac{1}{n+2} = \\ &= a^{|u|+1} f_L(u) + \tfrac{1}{n+2}. \blacksquare \end{aligned}$$

3.5 Эргодичные автоматы

3.5.1 Вспомогательные понятия и результаты

Мы будем использовать следующие обозначения:

- если $v = (v_1, \ldots, v_n) \in \mathbf{R}^n$, то записи $|v|$ и $\|v\|$ обозначают соответственно числа

$$\max_{i=1\ldots n} |v_i| \quad \text{и} \quad \max_{i=1\ldots n} v_i - \min_{i=1\ldots n} v_i,$$

- если A – квадратная матрица порядка n, то записи $|A|$ и $\|A\|$ обозначают соответственно числа

$$\max_{i,j=1\ldots n} |a_{ij}| \quad \text{и} \quad \max_{i=1\ldots n} \|A_i^{\downarrow}\|,$$

  где $A_i^{\downarrow}$ – i–й столбец матрицы A,

- если A – матрица (в частности, вектор-строка или вектор-столбец), то запись $A > 0$ означает, что каждый элемент этой матрицы положителен,

- если A – квадратная матрица, то запись $Q(A)$ обозначает **булев шаблон** матрицы A, т.е. матрицу, получаемую из A заменой каждого ненулевого элемента на 1,

- если $\{A^x \mid x \in X\}$ совокупность матриц порядка n, индексированных элементами множества X, то

– $\forall\, u \in X^*$ запись A^u обозначает матрицу, определяемую следующим образом: A^ε – единичная матрица порядка n, и если $u = x_1 \dots x_k$, то $A^u \stackrel{\text{def}}{=} A^{x_1} \dots A^{x_k}$, и

– $\forall\, u \in X^*, \forall\, i, j \in \{1, \dots, n\}$ запись A^u_{ij} обозначает элемент матрицы A^u в строке i и столбце j.

Нетрудно доказать, что если $\{A^x \mid x \in X\}$ – совокупность стохастических матриц одинакового порядка, то

$$\forall\, u, v \in X^* \;\; Q(A^{uv}) = Q(A^u) Q(A^v),$$

где произведение булевых шаблонов определяется аналогично обычному произведению матриц, с единственным отличием: сумма $1 + 1$ считается равной 1 (мы будем называть такое произведение **булевым произведением**).

Теорема 22.

Пусть заданы конечное множество X, натуральное число n, и совокупность $\{A^x \mid x \in X\}$ стохастических матриц порядка n.

Следующие условия эквивалентны:

(a) $\lim\limits_{|u| \to \infty} \|A^u\| = 0$,

(b) $\exists\, k > 0 : \forall\, u \in X^k \;\exists\, i \in \{1, \dots, n\} : A^{u\downarrow}_i > 0$,

(c) $\exists\, k > 0 : \forall\, u \in X^k$

$$\forall\, i, i' \in \{1, \dots, n\} \;\exists\, j : A^u_{ij} > 0, A^u_{i'j} > 0. \tag{3.29}$$

Доказательство.

Схема доказательства: (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).

- (a) $\Rightarrow$ (b): выберем k так, что $\forall\, u \in X^k \; \|A^u\| < \frac{1}{n}$.

  Поскольку $\forall\, u \in X^k$ матрица A^u стохастическая, то в любой её строке существует элемент $\geq \frac{1}{n}$. Столбец $A^{u\downarrow}_i$, в котором содержится этот элемент, обладает свойством $\|A^{u\downarrow}_i\| < \frac{1}{n}$, поэтому $A^{u\downarrow}_i > 0$.

- (b) $\Rightarrow$ (c): очевидно.

- (c) $\Rightarrow$ (a): пусть верно (c), т.е. $\exists\, k > 0 : \forall\, u \in X^k$ верно (3.29). Обозначим символом c минимальный положительный элемент матриц вида A^u, где $u \in X^k$.

Лемма.

Для любой матрицы B порядка n верно неравенство

$$\|A^u B\| \leq (1-c) \cdot \|B\|. \tag{3.30}$$

Доказательство.

Обозначим матрицу $A^u B$ символом D, и элементы матриц B и D – записями b_{ij}, d_{ij} $(i, j = 1, \ldots, n)$ соответственно.

Для доказательства неравенства (3.30) достаточно доказать, что $\forall\, i, i', j \in \{1, \ldots, n\}$ верно неравенство

$$|d_{ij} - d_{i'j}| \leq (1-c)(M_j - m_j), \tag{3.31}$$

где $M_j := \max\limits_t b_{tj}$, $m_j = \min\limits_t b_{tj}$.

Пусть $i, i' \in \{1, \ldots, n\}$. Обозначим записью j_0 индекс, удовлетворяющий условию $A^u_{ij_0} \geq c, A^u_{i'j_0} \geq c$.

Верны равенства

$$\begin{aligned} d_{ij} &= \sum_{t \neq j_0} A^u_{it} b_{tj} + \Big(A^u_{ij_0} M_j - A^u_{ij_0}(M_j - b_{j_0 j})\Big) \\ d_{i'j} &= \sum_{t \neq j_0} A^u_{i't} b_{tj} + \Big(A^u_{i'j_0} m_j + A^u_{i'j_0}(b_{j_0 j} - m_j)\Big) \end{aligned}$$

из которых следуют соотношения

$$\begin{aligned} d_{ij} &\leq \Big(\sum_t A^u_{it}\Big) M_j - A^u_{ij_0}(M_j - b_{j_0 j}) = M_j - A^u_{ij_0}(M_j - b_{j_0 j}) \\ d_{i'j} &\geq \sum_t A^u_{i't} m_j + A^u_{i'j_0}(b_{j_0 j} - m_j) = m_j + A^u_{i'j_0}(b_{j_0 j} - m_j) \end{aligned}$$

из которых следует неравенство

$$d_{ij} - d_{i'j} \leq M_j - A^u_{ij_0}(M_j - b_{j_0 j}) - m_j - A^u_{i'j_0}(b_{j_0 j} - m_j). \tag{3.32}$$

Поскольку $M_j - b_{j_0 j} \geq 0$ и $b_{j_0 j} - m_j \geq 0$, то, используя определение c, можно оценить сверху правую часть (3.32) значением

$$M_j - c(M_j - b_{j_0 j}) - m_j - c(b_{j_0 j} - m_j) = (1-c)(M_j - m_j).$$

Таким образом, доказано неравенство

$$d_{ij} - d_{i'j} \leq (1-c)(M_j - m_j). \tag{3.33}$$

В силу произвольности выбора индексов i, i' верно неравенство

$$d_{i'j} - d_{ij} \leq (1-c)(M_j - m_j). \tag{3.34}$$

Из (3.33) и (3.34) следует (3.31). ■

Из леммы следует, что

$$\forall\, u \in X^* \; \|A^u\| \leq (1-c)^{[|u|/k]} \cdot \|A^l\| \quad (|l| < k). \tag{3.35}$$

Правая часть (3.35) стремится к нулю при $|u| \to \infty$, поэтому условие (a) верно. ■

3.5.2 Понятие эргодичного вероятностного автомата и критерий эргодичности

ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ называется **эргодичным**, если

$$\forall\, s, s' \in S_A \quad |\vec{A}^u_s - \vec{A}^u_{s'}| \to 0 \quad (|u| \to \infty). \tag{3.36}$$

Теорема 23.

ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ эргодичен тогда и только тогда, когда $\forall\, u \in X^* \setminus \{\varepsilon\}$ матрица A^u **регулярна** (т.е. $\exists\, n: \; (A^u)^n > 0$).

Доказательство.

Если A эргодичен, но для некоторого $u \in X^* \setminus \{\varepsilon\}$ матрица A^u нерегулярна, то $\forall\, k \geq 1$ $(A^u)^k$ нерегулярна. Из эргодичности A следует, что выполнено условие (a) в теореме 22 ($\lim\limits_{|u| \to \infty} \|A^u\| = 0$). Поэтому выполнено условие (b) в этой теореме, т.е.

$$\exists\, l > 0 : \forall\, u \in X^l \; \exists\, i : \; A_i^{u\downarrow} > 0.$$

Можно доказать, что это противоречит нерегулярности матриц вида $(A^u)^k$, где k – произвольное натуральное число.

Обратно, пусть $\forall\, u \in X^* \setminus \{\varepsilon\} \; \exists\, n : \; (A^u)^n > 0$.

Мы будем использовать следующие обозначения.

- Запись $\mathcal{Q}_A$ обозначает множество $\{Q(A^u) \mid u \in X^*\}$ булевых шаблонов матриц вида A^u.

  Нетрудно видеть, что множество $\mathcal{Q}_A$ конечно.

- Символ $\mathcal{Q}_A^I$ обозначает множество всех матриц из $\mathcal{Q}_A$, содержащие столбец I (все его элементы равны 1).

- Символ k обозначает число различных матриц из $\mathcal{Q}_A \setminus \mathcal{Q}_A^I$.

Отметим, что

$$P \in \mathcal{Q}_A^I \;\Rightarrow\; \forall\, Q \in \mathcal{Q}_A \;\; QP \in \mathcal{Q}_A^I. \tag{3.37}$$

Докажем, что

$$\forall\, u \in X^{k+1} \;\; Q(A^u) \in \mathcal{Q}_A^I. \tag{3.38}$$

Пусть $u = (x_1 \dots x_{k+1})$ и $Q(A^u) \notin \mathcal{Q}_A^I$. Из (3.37) следует, что

$$\forall\, i = 1, \dots, k+1 \quad Q(A^{x_1}) \cdot \ldots \cdot Q(A^{x_{k+1}}) \in \mathcal{Q}_A \setminus \mathcal{Q}_A^I. \tag{3.39}$$

Поскольку $|\mathcal{Q}_A \setminus \mathcal{Q}_A^I| = k$, то из (3.39) следует, что

$$\begin{array}{l} \exists\, i, j \in \{1, \dots, k+1\} : i < j, \\ Q(A^{x_i}) \cdot \ldots \cdot Q(A^{x_{k+1}}) = Q(A^{x_j}) \cdot \ldots \cdot Q(A^{x_{k+1}}) \notin \mathcal{Q}_A^I. \end{array} \tag{3.40}$$

Определим $v \stackrel{\text{def}}{=} (x_i \dots x_{j-1})$, $w \stackrel{\text{def}}{=} (x_j \dots x_{k+1})$. Из (3.40) следует, что

$$\forall\, t \geq 1 \quad Q((A^v)^t A^w) = Q(A^w) \notin \mathcal{Q}_A^I. \tag{3.41}$$

Поскольку

- из регулярности A^v следует, что $\exists\, t \geq 1 : (A^v)^t > 0$, и

- из регулярности A^w следует, что $\exists\, s : A_s^{w\downarrow} \neq \mathbf{0}^{\downarrow}$,

то $(A^v)^t A_s^{w\downarrow} > 0$, т.е. $Q((A^v)^t A^w) \in \mathcal{Q}_A^I$, что противоречит (3.41).

Таким образом, свойство (3.38) верно. Это свойство совпадает с условием (b) теоремы 22 для ВА A. Следовательно, верно условие (a) этой теоремы, откуда следует эргодичность A. ■

3.6 Устойчивость вероятностных автоматов

3.6.1 Вспомогательные утверждения

Мы будем использовать следующие обозначения: если A – матрица, и $c \in \mathbf{R}$, то записи $A > c$ и $A \geq c$ означают, что каждый элемент этой матрицы $> c$ или $\geq c$ соответственно.

Теорема 24.
Если A – стохастическая матрица порядка n, $A \geq c \geq 0$, и $\lambda \in \mathbf{R}^n$ – вектор-столбец, то

$$\|A\lambda\| \leq (1-2c)\|\lambda\|.$$

Доказательство.
Обозначим элементы матрицы A и вектор-столбца λ записями a_{ij} и λ_i $(i, j = 1, \ldots, n)$ соответственно. Кроме того, обозначим записью $b = (b_1 \ldots b_n)^\sim$ вектор-столбец $A\lambda$. Будем считать, что

$$b_1 = \max b_i, \quad b_2 = \min b_i, \quad \lambda_1 = \max \lambda_i, \quad \lambda_2 = \min b_i$$

(если это не так, то соответствующим образом переставим в матрице A столбцы и строки).

Надо доказать, что $b_1 - b_2 \leq (1-2c)(\lambda_1 - \lambda_2)$.

Мы докажем более сильное неравенство:

$$b_1 - b_2 \leq (1 - a_{12} - a_{21})(\lambda_1 - \lambda_2).$$

Нетрудно видеть, что верны следующие соотношения:

- $b_1 = \sum\limits_j a_{1j}\lambda_j = (1 - \sum\limits_{j\geq 2} a_{1j})\lambda_1 + \sum\limits_{j\geq 2} a_{1j}\lambda_j =$
  $= \lambda_1 - \sum\limits_{j\geq 2} a_{1j}(\lambda_1 - \lambda_j) \leq \lambda_1 - a_{12}(\lambda_1 - \lambda_2)$,

- $b_2 = \sum\limits_j a_{2j}\lambda_j = a_{21}\lambda_1 + (1 - a_{21} - \sum\limits_{j\geq 3} a_{2j})\lambda_2 + \sum\limits_{j\geq 3} a_{2j}\lambda_j =$
  $= \lambda_2 + a_{21}(\lambda_1 - \lambda_2) + \sum\limits_{j\geq 3} a_{2j}(\lambda_j - \lambda_2) \geq \lambda_2 + a_{21}(\lambda_1 - \lambda_2)$.

Таким образом,

$$b_1 - b_2 \le \lambda_1 - a_{12}(\lambda_1 - \lambda_2) - \lambda_2 - a_{21}(\lambda_1 - \lambda_2) =$$
$$= (\lambda_1 - \lambda_2) - (a_{12} + a_{21})(\lambda_1 - \lambda_2) = (\lambda_1 - \lambda_2)(1 - a_{12} - a_{21}). \ \blacksquare$$

Теорема 25.

Если $\exists\, c > 0$: $\forall\, x \in X$ матрица A^x удовлетворяет неравенству $A^x \ge c$ и является стохастической, то $\forall\, u \in X^*$

$$\|A^u\| \le (1 - 2c)^{|u|-1}. \tag{3.42}$$

Доказательство.

Обозначим символом n порядок матриц A^x.

Сначала докажем, что $\forall\, u \in X^*$ верно соотношение

$$\forall\, \lambda \in \mathbf{R}^n \quad \|A^u\lambda\| \le (1 - 2c)^{|u|}\|\lambda\|. \tag{3.43}$$

(3.43) верно для $u = \varepsilon$.

Если (3.43) верно для некоторого $u \in X^*$, то $\forall\, x \in X$, используя предположение (3.43) и теорему 24, получаем:

$$\|A^{ux}\lambda\| = \|A^u A^x \lambda\| \le (1 - 2c)^{|u|}\|A^x\lambda\| \le$$
$$\le (1 - 2c)^{|u|}(1 - 2c)\|\lambda\| = (1 - 2c)^{|ux|}\|\lambda\|.$$

Таким образом, (3.43) верно для всех $u \in X^*$.

Теперь докажем (3.42) индукцией по $|u|$.

- Если $|u| = 0$ или 1, то (3.42) верно.
- Пусть (3.42) верно для некоторой строки $u \in X^*$, и пусть $x \in X$.

  Обозначим совокупность столбцов A^x записью $(\lambda_1 \ldots \lambda_n)$.

  Используя (3.43) и свойство $\forall\, i = 1, \ldots, n$ $\|\lambda_i\| \le 1$, получаем:

$$\|A^{ux}\| = \|A^u A^x\| = \|A^u(\lambda_1 \ldots \lambda_n)\| = \|(A^u\lambda_1 \ldots A^u\lambda_n)\| =$$
$$= \max_i \|A^u\lambda_i\| \le (1 - 2c)^{|u|} \max_i \|\lambda_i\| \le (1 - 2c)^{|u|} =$$
$$= (1 - 2c)^{|ux|-1}.$$

Таким образом, (3.42) верно для любой строки $u \in X^*$. ■

Теорема 26.

Если A – стохастическая матрица порядка n, то для любой матрицы B порядка n верно неравенство

$$|AB - B| \leq \|B\|.$$

Доказательство.

Пусть представление матрицы B в виде последовательности столбцов имеет вид $(B_1^{\downarrow} \dots B_n^{\downarrow})$, тогда

$$\begin{aligned}
&|AB - B| = |A(B_1^{\downarrow} \dots B_n^{\downarrow}) - (B_1^{\downarrow} \dots B_n^{\downarrow})| = \\
&= |(AB_1^{\downarrow} - B_1^{\downarrow} \dots AB_n^{\downarrow} - B_n^{\downarrow})| = \\
&= \max_i |AB_i^{\downarrow} - B_i^{\downarrow}| \leq \max_i \|B_i^{\downarrow}\| = \|B\|,
\end{aligned}$$

где последнее неравенство следует из нижеслежующей леммы.

Лемма.

Если A – стохастическая матрица порядка n, и $\lambda \in \mathbf{R}^n$ – вектор-столбец, то $|A\lambda - \lambda| \leq \|\lambda\|$.

Доказательство.

Пусть элементы A и λ имеют вид a_{ij} и λ_i $(i, j = 1, \dots, n)$. Поскольку $\forall\, i = 1, \dots, n \;\; \sum\limits_j a_{ij} = 1$, то

$$\begin{aligned}
&|A\lambda - \lambda| = \max_i |\sum_j (a_{ij}\lambda_j) - \lambda_i| = \\
&= \max_i |\sum_j (a_{ij}\lambda_j) - \sum_j (a_{ij}\lambda_i)| = \max_i |\sum_j a_{ij}(\lambda_j - \lambda_i)| \leq \\
&\leq \max_i \sum_j a_{ij}|\lambda_j - \lambda_i| \leq \max_i \sum_j a_{ij}\|\lambda\| = \|\lambda\|. \blacksquare
\end{aligned}$$

Теорема 27.

Пусть P, Q – стохастические матрицы порядка n, тогда для любой матрицы B порядка n верно неравенство

$$|PB - QB| \leq \|B\|. \tag{3.44}$$

Доказательство.

Обозначим матрицу $P - Q$ символом A, и элементы матриц A, P, Q – записями a_{ij}, p_{ij}, q_{ij} $(i, j = 1, \ldots, n)$ соответственно.

Матрица A удовлетворяет условию:

$$\forall\, i = 1, \ldots, n \quad \sum_{j=1}^{n} a_{ij} = 0, \quad \sum_{j=1}^{n} a_{ij}^{(+)} \leq 1, \tag{3.45}$$

где $a_{ij}^{(+)} \overset{\text{def}}{=} \begin{cases} a_{ij}, & \text{если } a_{ij} \geq 0 \\ 0, & \text{иначе} \end{cases}$, т.к. $\forall\, i = 1, \ldots, n$

- $\sum\limits_j a_{ij} = \sum\limits_j (p_{ij} - q_{ij}) = \sum\limits_j p_{ij} - \sum\limits_j q_{ij} = 1 - 1 = 0$,
- если $a_{ij_1}, \ldots, a_{ij_k}$ – список всех неотрицательных элементов i–й строки матрицы A, то

$$\sum_s a_{ij_s} = \sum_s (p_{ij_s} - q_{ij_s}) = \sum_s p_{ij_s} - \sum_s q_{ij_s} \leq 1 - \sum_s q_{ij_s} \leq 1.$$

Перепишем доказываемое неравенство (3.44) в виде

$$|AB| \leq \|B\|. \tag{3.46}$$

Обозначим матрицу AB символом C, и элементы матриц A, B, C – записями a_{ij}, b_{ij}, c_{ij} $(i, j = 1, \ldots, n)$ соответственно.

Пусть элемент c_{kr} матрицы C таков, что

$$|c_{kr}| = \max_{i,j} |c_{ij}| \quad (= |AB|\,). \tag{3.47}$$

Определим $M \overset{\text{def}}{=} \max\limits_j b_{jr}$, $m \overset{\text{def}}{=} \min\limits_j b_{jr}$, и $B_r^{\downarrow} \overset{\text{def}}{=}$ r–й столбец матрицы B. Из этого определения следует, что

$$M - m = \|B_r^{\downarrow}\| \leq \|B\|. \tag{3.48}$$

Из (3.47) и (3.48) следует, что для доказательства неравенства (3.46) достаточно доказать неравенство

$$|c_{kr}| \leq M - m. \tag{3.49}$$

Обозначим записью $a^{(+)}$ $(a^{(-)})$ сумму всех положительных (отрицательных) чисел вида a_{kj} $(j = 1, \ldots, n)$. Заметим, что

- из $\sum_{j=1}^{n} a_{kj} = 0$ следует $a^{(+)} + a^{(-)} = 0$, или $a^{(-)} = -a^{(+)}$,
- из $\sum_{j=1}^{n} a_{kj}^{(+)} \leq 1$ следует $a^{(+)} \leq 1$.

Для доказательства неравенства (3.49) рассмотрим отдельно случаи $c_{kr} \geq 0$ и $c_{kr} < 0$.

- Если $c_{kr} \geq 0$, то
$$|c_{kr}| = c_{kr} = \sum_j a_{kj} b_{jr} \leq a^{(+)} M + a^{(-)} m = \\ = a^{(+)}(M - m) \leq M - m,$$
т.е. в данном случае неравенство (3.49) верно.
- Если $c_{kr} < 0$, то
$$c_{kr} = \sum_j a_{kj} b_{jr} \geq a^{(+)} m + a^{(-)} M = a^{(+)}(m - M),$$
поэтому $|c_{kr}| = -c_{kr} = a^{(+)}(M - m) \leq M - m$, т.е. в данном случае неравенство (3.49) также верно. ■

3.6.2 Понятие устойчивости вероятностных автоматов

Пусть заданы ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ и число $\varepsilon > 0$.

Запись $O_\varepsilon(A)$ обозначает множество ВА, каждый из которых получается из A путем прибавления к элементам строки ξ^0 и матриц A^x чисел из интервала $(-\varepsilon, \varepsilon)$.

ВА A называется **устойчивым** относительно ИТС $a \in I(A)$, если $\exists\, \varepsilon > 0 : \forall\, B \in O_\varepsilon(A) \quad a \in I(B)$ и $A_a = B_a$.

Теорема 28.
Пусть задан ВА $A = (\xi^0_A, \{A^x \mid x \in X\}, \lambda)$, и $\forall\, x \in X \;\; A^x > 0$.
Тогда $\forall\, a \in I(A)$ $\;A$ устойчив относительно a.

Доказательство.

По предположению, $\exists\,\delta>0$: $\forall\,u\in X^*$

$$|f_A(u)-a|>\delta. \tag{3.50}$$

Докажем, что $\exists\,\varepsilon>0$: $\forall\,B\in O_\varepsilon(A)$, $\forall\,u\in X^*$

$$|f_A(u)-f_B(u)|\le\tfrac{\delta}{2}. \tag{3.51}$$

Отметим, что из (3.51) следуют требуемые соотношения $a\in I(B)$ и $A_a=B_a$, т.к.

- $a\in I(B)$, верно потому, что $\forall\,u\in X^*$ из неравенства

  $$|f_A(u)-a|\le|f_A(u)-f_B(u)|+|f_B(u)-a|$$

  согласно (3.51) и (3.50) следуют соотношения

  $$\begin{aligned}&|f_B(u)-a|\ge|f_A(u)-a|-|f_A(u)-f_B(u)|>\\&>\delta-\tfrac{\delta}{2}=\tfrac{\delta}{2},\end{aligned} \tag{3.52}$$

- $A_a=B_a$ верно потому, что если $\exists\,u\in X^*$: соотношение

  $$f_A(u)>a\ \Leftrightarrow\ f_B(u)>a$$

  неверно, то, согласно (3.50) и (3.52), будет верно неравенство

  $$|f_A(u)-f_B(u)|>\delta,$$

  которое противоречит (3.51).

Пусть B имеет вид $(\xi^0_B,\{B^x\mid x\in X\},\lambda)$, тогда можно переписать (3.51) в виде

$$|\xi^0_A A^u\lambda-\xi^0_B B^u\lambda|\le\tfrac{\delta}{2}. \tag{3.53}$$

Поскольку

$$\begin{aligned}&|\xi^0_A A^u\lambda-\xi^0_B B^u\lambda|\le\\&\le|\xi^0_A A^u\lambda-\xi^0_A B^u\lambda|+|\xi^0_A B^u\lambda-\xi^0_B B^u\lambda|=\\&=|\xi^0_A(A^u-B^u)\lambda|+|(\xi^0_A-\xi^0_B)B^u\lambda|\le\\&\le|A^u-B^u|\cdot n\cdot|\lambda|+|\xi^0_A-\xi^0_B|\cdot n\cdot|\lambda|\end{aligned}$$

(где $n = |S_A|$), то, следовательно, неравенство (3.53) будет верно, если будут верны неравенства

$$|A^u - B^u| \leq \delta_1, \quad |\xi_A^0 - \xi_B^0| \leq \delta_1, \quad \text{где } \delta_1 = \tfrac{\delta}{4n|\lambda|}. \tag{3.54}$$

Таким образом, для доказательства теоремы (28) достаточно доказать, что $\exists\, \varepsilon \in (0, \delta_1)$: $\forall\, B \in O_\varepsilon(A)$, $\forall\, u \in X^*$

$$|A^u - B^u| < \delta_1. \tag{3.55}$$

По предположению, $\exists\, c > 0$: $\forall\, x \in X \quad A^x \geq c$.

Выберем k так, чтобы было верно неравенство

$$(1-c)^{k-1} < \tfrac{\delta_1}{3}. \tag{3.56}$$

Лемма.

Пусть $\varepsilon \in (0, \frac{c}{2})$, $B \in O_\varepsilon(A)$, и

$$\forall\, u \in X^{\leq k} \quad |A^u - B^u| < \tfrac{\delta_1}{3}. \tag{3.57}$$

Тогда $\forall\, u \in X^*$ верно неравенство (3.55).

Доказательство.

Если $|u| \leq k$, то (3.55) следует из (3.57).

Пусть $|u| > k$, тогда $u = u_1 u_2$, где $|u_2| = k$, и

$$\begin{aligned} &|A^u - B^u| = |A^{u_1 u_2} - B^{u_1 u_2}| \leq \\ &\leq |A^{u_1} A^{u_2} - A^{u_2}| + |B^{u_1} B^{u_2} - B^{u_2}| + |A^{u_2} - B^{u_2}|. \end{aligned} \tag{3.58}$$

Согласно теореме (26), верны неравенства

$$|A^{u_1} A^{u_2} - A^{u_2}| \leq \|A^{u_2}\|, \quad |B^{u_1} B^{u_2} - B^{u_2}| \leq \|B^{u_2}\|,$$

поэтому из (3.58) и из истинности $\forall\, u \in X^k$ неравенства в (3.57) следует неравенство

$$|A^u - B^u| < \|A^{u_2}\| + \|B^{u_2}\| + \tfrac{\delta_1}{3}. \tag{3.59}$$

Согласно теореме 25 и условию (3.56),

- из условия $\forall\, x \in X \ \ A^x \geq c$ следует соотношение

$$\begin{array}{l} \forall\, u \in X^k \\ \|A^u\| \leq (1-2c)^{k-1} < (1-c)^{k-1} < \frac{\delta_1}{3}, \end{array} \tag{3.60}$$

- из условий $B \in O_\varepsilon(A)$, $\varepsilon \in (0, \frac{c}{2})$ и $\forall\, x \in X \quad A^x \geq c$ следует условие $\forall\, x \in X \ \ B^x \geq \frac{c}{2}$, из которого следует соотношение

$$\begin{array}{l} \forall\, u \in X^k \\ \|B^u\| \leq (1-2\cdot\frac{c}{2})^{k-1} = (1-c)^{k-1} < \frac{\delta_1}{3}. \end{array} \tag{3.61}$$

Из (3.59), (3.60), и (3.61) следует, что

$$|A^u - B^u| < \frac{\delta_1}{3} + \frac{\delta_1}{3} + \frac{\delta_1}{3} = \delta_1. \ \blacksquare$$

Таким образом, для доказательства теоремы 28 осталось доказать, что $\exists\, \varepsilon \in (0, \min\{\delta_1, \frac{c}{2}\}) : \forall\, B \in O_\varepsilon(A)$ верно (3.57).

Нетрудно доказать, что для каждой строки $u \in X^{\leq k}$

$$\exists\, \varepsilon_u > 0 : \forall\, \varepsilon \in (0, \varepsilon_u), \forall\, B \in O_\varepsilon(A) \quad |A^u - B^u| < \frac{\delta_1}{3}.$$

(для доказательства этого факта можно использовать следующее соображение: $\forall\, x \in X$ обозначим символом M^x матрицу порядка n, элемент в строке i и столбце j которой имеет вид $A^x_{ij} + t^x_{ij}$, где t^x_{ij} – различные переменные, и если $u = x_1 \ldots x_s$, то $M^u \stackrel{\mathrm{def}}{=} M^{x_1} \ldots M^{x_s}$, тогда выражение $|A^u - M^u|$ определяет непрерывную функцию от переменных t^x_{ij}, значение которой равно 0 в том случае когда значения всех переменных t^x_{ij} равны 0, и т.д.)

Поскольку число строк в $X^{\leq k}$ конечно, то в качестве искомого ε можно взять $\min\{\delta_1, \frac{c}{2}, \varepsilon_u \ (u \in X^{\leq k})\}$. $\blacksquare$

Нижеследующая теорема является усилением теоремы 28.

Теорема 29.
Пусть задан ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$, и

$$\exists\, l : \forall\, u \in X^l \ \ A^u > 0.$$

Тогда $\forall\, a \in I(A) \ \ A$ устойчив относительно a.

Доказательство.

Пусть $a \in I_\delta(A)$, и $|S_A| = n$.

Как и в доказательстве теоремы 28, для доказательства теоремы 29 достаточно доказать, что $\exists\, \varepsilon \in (0, \delta_1)$, где $\delta_1 \stackrel{\text{def}}{=} \frac{\delta}{4n|\lambda|}$:

$$\forall\, B \in O_\varepsilon(A), \ \ \forall\, u \in X^* \quad |A^u - B^u| < \delta_1. \tag{3.62}$$

По предположению, $\forall\, u \in X^l \; \exists\, c_u > 0 : A^u \geq c_u$.

Определим $c \stackrel{\text{def}}{=} \min\limits_{u \in X^l} c_u$. Таким образом, $\forall\, u \in X^l \; A^u \geq c$.

Выберем k так, чтобы было верно неравенство $(1-c)^{[k/l]} < \frac{\delta_1}{3}$.

Аналогично доказательству теоремы 28, доказываются свойства

- $\forall\, u \in X^k \;\; \|A^u\| \leq (1-c)^{[k/l]}$, и
- $\exists\, \varepsilon > 0$: $\forall\, B \in O_\varepsilon(A)$, $\forall\, u \in X^{\leq k} \;\; |A^u - B^u| \leq \frac{\delta_1}{3}$,

из которых следует (3.62). ■

Можно доказать, что если $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ – ВА, то $\forall\, a \in I(A)$ A устойчив относительно a тогда и только тогда, когда $\exists\, \delta < 1 : \forall\, x \in X \;\; \|A^x\| < \delta$.

Глава 4

Вероятностные языки

4.1 Понятие вероятностного языка

Пусть A – ВА, и $a \in [0, 1)$.

Вероятностным языком (ВЯ) ВА A с точкой сечения a называется множество строк $A_a \subseteq X^*$ (где X – множество входных сигналов A), определяемое следующим образом:

$$A_a \stackrel{\text{def}}{=} \{u \in X^* \mid f_A(u) > a\}.$$

Понятие ВЯ может использоваться для решения различных задач, в частности для моделирования процесса обучения. Одна из моделей обучения имеет следующий вид. Задано конечное множество S, элементы которого называются **реакциями** на некоторый стимул, причём каждая реакция $s \in S$ рассматривается либо как правильная, либо как неправильная. Обозначим символом λ вектор-столбец, компоненты которого индексированы реакциями $s \in S$, причем те компоненты λ, индексы которых являются правильными реакциями, равны 1, а остальные компоненты равны 0. С обучаемой системой в каждый момент времени $t = 0, 1, \ldots$ связано некоторое распределение $\xi \in S^\triangle$, в соответствии с которым она реагирует на стимул в этот момент: для каждого $s \in S$ вероятность того, что система выдаст реакцию s в момент t, равна s^ξ. Из данных определений следует, что вероятность выдачи правильной реакции в момент t имеет вид $\xi\lambda$. Предполагается, что задано конечное множество $\{A^x \mid x \in X\}$

обучающих операторов, каждый из которых является стохастической матрицей порядка $|S|$ и определяет изменение реакции системы на стимул по следующему правилу: если

- в текущий момент времени система реагировала на стимул в соответствии с распределением $\xi \in S^{\triangle}$, и
- в этот момент был применён обучающий оператор A^x,

то в следующий момент времени система будет реагировать на стимул в соответствии с распределением ξA^x. Таким образом, путем применения к обучаемой системе обучающих операторов можно добиться изменения вероятности правильной реакции на стимул. Система считается **обученной** в некоторой момент времени, если вероятность того, что она выдаст в этот момент правильную реакцию на стимул, превышает некоторое заданное значение $a \in [0, 1)$. Одна из задач, связанных с обучением, имеет следующий вид: задано начальное распределение $\xi^0 \in S^{\triangle}$, требуется описать все последовательности $(A^{x_1}, \ldots, A^{x_n})$ обучающих операторов, каждая из которых обладает следующим свойством: после последовательного применения операторов, входящих в эту последовательность, система станет обученной. Нетрудно видеть, что каждая такая последовательность соответствует строке $x_1 \ldots x_n \in X^*$, такой, что $\xi^0 A^{x_1} \ldots A^{x_n} \lambda > a$, т.е. строке из ВЯ A_a, где $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$.

4.2 Свойства вероятностных языков

Теорема 30.

Пусть $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ – ВА, и $a \in [0, 1)$.

Тогда $A_a = B_a$, где ВА B имеет вид

$$B = (\vec{e}_1, \begin{pmatrix} 0 & \xi^0 A^x \\ \mathbf{0} & A^x \end{pmatrix}, \begin{pmatrix} \xi^0 \lambda \\ \lambda \end{pmatrix}),$$

где $\vec{e}_1$ – вектор-строка длины $|S_A| + 1$, у которой первая компонента равна 1, а остальные компоненты равны 0, и $\mathbf{0}$ – вектор-столбец длины $|S_A|$, все компоненты которого равны 0.

Доказательство.

Нетрудно доказать, что $\forall\, u \in X^* \setminus \{\varepsilon\} \quad B^u = \begin{pmatrix} 0 & \xi^0 A^u \\ \mathbf{0} & A^u \end{pmatrix}$, откуда следует равенство $f_A = f_B$. ■

Теорема 31.

Пусть задан ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$, где каждая компонента λ_j вектора λ принадлежит отрезку $[0, 1]$, и $a \in [0, 1)$.

Тогда $A_a = B_a$, где $B = (\xi^0_B, \{B^x \mid x \in X\}, \lambda_B)$, $|S_B| = 2|S_A|$, и

- ξ^0_B получается из ξ^0 добавлением 0 после каждой компоненты,
- $\forall\, x \in X$ матрица B^x получается из матрицы A^x заменой каждого её элемента A^x_{ij} на матрицу $\begin{pmatrix} \lambda_j A^x_{ij} & (1-\lambda_j) A^x_{ij} \\ \lambda_j A^x_{ij} & (1-\lambda_j) A^x_{ij} \end{pmatrix}$,
- $\lambda_B = (1\,0\,\ldots\,1\,0)^\sim$.

Доказательство.

Нетрудно доказать, что $\forall\, u \in X^* \setminus \{\varepsilon\}$ матрица B^u получается из матрицы A^u заменой каждого её элемента A^u_{ij} на матрицу $\begin{pmatrix} \lambda_j A^u_{ij} & (1-\lambda_j) A^u_{ij} \\ \lambda_j A^u_{ij} & (1-\lambda_j) A^u_{ij} \end{pmatrix}$, откуда следует равенство $f_A = f_B$. ■

Теорема 32.

Пусть A – ВА, и $a, b \in [0, 1)$. Тогда $\exists$ ВА B :

$$A_a = B_b. \tag{4.1}$$

Доказательство.

Пусть ВА A имеет вид $(\xi^0_A, \{A^x \mid x \in X\}, \lambda_A)$.
Рассмотрим отдельно случаи $b < a$ и $b > a$.

- Если $b < a$, то $b = \alpha a$, где $\alpha \in [0, 1)$.

  В этом случае $|S_B| = 2|S_A|$, и компоненты ξ^0_B, B^x, λ_B искомого ВА B имеют следующий вид:

  - $\xi^0_B = (\xi^0_A, \mathbf{0})$, где $\mathbf{0}$ – вектор-строка длины $|S_A|$, все компоненты которого равны 0,

– $\forall\, x \in X \quad B^x \stackrel{\text{def}}{=} \begin{pmatrix} \alpha A^x & (1-\alpha)A^x \\ \alpha A^x & (1-\alpha)A^x \end{pmatrix}$,

– $\lambda_B = \begin{pmatrix} \lambda_A \\ \mathbf{0} \end{pmatrix}$, где $\mathbf{0}$ – вектор-столбец длины $|S_A|$, все компоненты которого равны 0.

Нетрудно доказать, что

$$\forall\, u \in X^* \setminus \{\varepsilon\} \quad B^u = \begin{pmatrix} \alpha A^u & (1-\alpha)A^u \\ \alpha A^u & (1-\alpha)A^u \end{pmatrix},$$

откуда следует, что $\forall\, u \in X^* \;\; f_B(u) = \alpha f_A(u) = \frac{b}{a} f_A(u)$.

Из последнего соотношения следует (4.1).

- Если $b > a$, то $b = \alpha + (1-\alpha)a$, где $\alpha \in (0,1)$.

  В этом случае $|S_B| = 1 + |S_A|$, и компоненты ξ^0_B, B^x, λ_B искомого ВА B имеют следующий вид:

  – $\xi^0_B = (\alpha, (1-\alpha)\xi^0_A)$,

  – $\forall\, x \in X \quad B^x = \begin{pmatrix} 1 & 0 \\ 0 & A^x \end{pmatrix}$,

  – $\lambda_B = \begin{pmatrix} 1 \\ \lambda_A \end{pmatrix}$.

  Нетрудно доказать, что $\forall\, u \in X^*$

$$f_B(u) = \alpha + (1-\alpha) f_A(u) = \tfrac{b-a}{1-a} + \tfrac{1-b}{1-a} f_A(u),$$

  откуда следует (4.1). ■

4.3 Языки, представимые вероятностными автоматами общего вида

Пусть $A = (X, Y, S, P, \xi^0)$ – ВА общего вида, $y \in Y$, и $a \in [0,1)$. Обозначим записью $A_{y,a}$ множество

$$A_{y,a} \stackrel{\text{def}}{=} \{ux \mid u \in X^*, x \in X, \xi^0 A^u A^{xy} I > a\},$$

где $A^u \stackrel{\text{def}}{=} \sum\limits_{v \in Y^*} A^{u,v}$.

Если строка $u \in X^*$ имеет вид $x_0 \ldots x_{k-1}$, то значение $\xi^0 A^u A^{xy} I$ можно интерпретировать как вероятность того, что если

- в моменты времени $0, 1, \ldots, k-1$ на вход A последовательно поступали элементы строки u (т.е. в момент 0 поступил сигнал x_0, в момент 1 поступил сигнал x_1, ..., в момент $k-1$ поступил сигнал x_{k-1}), и

- в момент k поступил сигнал x,

то в момент времени k выходный сигнал A равен y.

Таким образом, множество $A_{y,a}$ можно интерпретировать как совокупность всех строк $u \in X^* \setminus \{\varepsilon\}$, обладающих следующим свойством: вероятность того, что

- если, начиная с момента 0, на вход A последовательно поступали элементы строки u

- то при подаче на вход A последнего входного сигнала из u выходной сигнал A в этот момент времени равен y,

больше a.

Множество $A_{y,a}$ называется **языком, представимым ВА общего вида** A выходным сигналом y и точкой сечения a.

Теорема 33.
Пусть $A = (X, Y, S, P, \xi^0)$ – ВА общего вида, $y \in Y$, и $a \in [0, 1)$. Тогда $A_{y,a} = B_a \setminus \{\varepsilon\}$, где B – ВА вида

$$B = \left((\xi^0, \mathbf{0}), \{ \begin{pmatrix} \sum\limits_{y' \neq y} A^{xy'} & A^{xy} \\ \sum\limits_{y' \neq y} A^{xy'} & A^{xy} \end{pmatrix} \mid x \in X \}, (\mathbf{0}, \mathbf{1})^{\sim} \right)$$

где $\mathbf{0}$ и $\mathbf{1}$ – вектор-строки длины $|S_A|$, все компоненты которых равны 0 и 1, соответственно.

Доказательство.

Нетрудно доказать, что

$$\forall\, u \in X^*, \forall\, x \in X \quad B^{ux} = \begin{pmatrix} A^u \sum\limits_{y' \neq y} A^{xy'} & A^u A^{xy} \\ A^u \sum\limits_{y' \neq y} A^{xy'} & A^u A^{xy} \end{pmatrix},$$

откуда следует соотношение

$$\forall\, u \in X^*, \forall\, x \in X \quad f_B(ux) = \xi^0 A^u A^{xy} I,$$

которое влечёт доказываемое утверждение. ■

4.4 Регулярность вероятностных языков

4.4.1 Понятие регулярного языка

Регулярный язык над конечным множеством X – это подмножество $U \subseteq X^*$, такое, что существует автомат Мура M вида (1.6), удовлетворяющий условиям:

$$|S| < \infty, \;\; Y = \{0, 1\}, \;\; U = \{u \in X^* \mid f_M(u) = 1\}. \qquad (4.2)$$

Весом регулярного языка U называется наименьшее число n, такое, что $\exists$ автомат Мура M вида (1.6), такой, что $|S| = n$ и выполнены условия (4.2). Мы будем обозначать вес регулярного языка U записью $w(U)$.

Можно доказать, что для любого языка $U \subseteq X^*$ следующие утверждения эквивалентны:

- язык U является регулярным,
- число классов эквивалентности $\underset{U}{\sim}$ на множестве X^*, определяемой следующим образом: $\forall\, u, u' \in X^*$

  $$u \underset{U}{\sim} u' \;\Leftrightarrow\; \forall\, v \in X^* \; (uv \in U \Leftrightarrow u'v \in U) \qquad (4.3)$$

  является конечным,

и если U регулярен, то $w(U)$ совпадает с числом классов эквивалентности $\underset{U}{\sim}$.

Теорема 34.

Пусть X – конечное множество, и $U \subseteq X^*$ – регулярный язык. Тогда U – ВЯ.

Доказательство.

Пусть $M = (X, \{0,1\}, S, \delta, \lambda, s^0)$ – автомат Мура такой, что $U = \{u \in X^* \mid f_M(u) = 1\}$.

Обозначим записью $\tilde{M}$ ВА $(\xi_{s^0}, \{A^x \mid x \in X\}, \lambda)$, множество состояний и отображение λ которого совпадают с соответствующими компонентами автомата M, и

$$\forall\, x \in X, \forall\, s, s' \in S \quad A^x_{s,s'} \stackrel{\text{def}}{=} \begin{cases} 1, & \text{если } \delta(s,x) = s', \\ 0, & \text{иначе.} \end{cases}$$

Нетрудно доказать, что $U = \tilde{M}_0$. ■

4.4.2 Изолированные точки сечения

Пусть задан ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$.

Число $a \in [0,1)$ называется **изолированной точкой сечения (ИТС)** для ВА A, если $\exists\, \delta > 0$:

$$\forall\, u \in X^* \quad |f_A(u) - a| > \delta. \tag{4.4}$$

Для каждого ВА A и каждого $\delta > 0$ запись $I_\delta(A)$ обозначает совокупность всех ИТС a для A, обладающих свойством (4.4). Запись $I(A)$ обозначает множество $\bigcup\limits_{\delta>0} I_\delta(A)$.

Можно доказать, что проблема выяснения истинности утверждения $a \in I(A)$ (где a и все численные значения ВА A – рациональные числа) алгоритмически неразрешима.

Теорема 35.

Пусть задан ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$, где $\lambda \in [0,1]^n$, $n = |S_A|$, и $\exists\, \delta > 0 : a \in I_\delta(A)$.

Тогда A_a – регулярный язык, и

$$w(A_a) \leq (1 + \tfrac{1}{\delta})^{n-1}. \tag{4.5}$$

Доказательство.

Обозначим символом R множество, содержащее по одному представителю каждого класса эквивалентности $\underset{A_a}{\sim}$. Из определения R следует, что $\forall\, u, u' \in U,\ u \neq u' \Rightarrow \exists\, v \in X^*$:

$$uv \in A_a \not\Leftrightarrow u'v \in A_a. \tag{4.6}$$

Т.к. $a \in I_\delta(A)$, то из (4.6) и (4.4) следует неравенство

$$|f_A(uv) - f_A(u'v)| > 2\delta,$$

т.е.

$$|\xi^0(A^u - A^{u'})A^v\lambda| > 2\delta. \tag{4.7}$$

Поскольку $A^vI = I$ и $\lambda \in [0,1]^n$, то $\lambda' \stackrel{\text{def}}{=} A^v\lambda \in [0,1]^n$.

Можно доказать, что существует частичная функция Vol_{n-1} с неотрицательными значениями на подмножествах множества $\mathbf{R}^n$, называемая $n-1$–**мерным объемом**, и обладающая следующими свойствами:

- значение $Vol_{n-1}(\{1,\ldots,n\}^\triangle)$ определено и положительно, (обозначим это значение символом C)
- если $Vol_{n-1}(M)$ определено, то $\forall\, \vec{x} = (x_1,\ldots,x_n) \in \mathbf{R}^n$

  $$Vol_{n-1}(\vec{x} + M) = Vol_{n-1}(M),$$

  где $\vec{x} + M \stackrel{\text{def}}{=} \{(x_1+y_1,\ldots,x_n+y_n) \mid (y_1,\ldots,y_n) \in M\}$,
- если $Vol_{n-1}(M)$ определено, то $\forall\, d > 0$

  $$Vol_{n-1}(dM) = d^{n-1}Vol_n(M),$$

  где $dM \stackrel{\text{def}}{=} \{(dx_1,\ldots,dx_n) \mid (x_1,\ldots,x_n) \in M\}$,
- если $Vol_{n-1}(M)$ и $Vol_{n-1}(M')$ определены, и $M \cap M' = \emptyset$, то

  $$Vol_{n-1}(M \cup M') = Vol_{n-1}(M) + Vol_{n-1}(M'),$$
- если $Vol_{n-1}(M)$ и $Vol_{n-1}(M')$ определены, и $M \subseteq M'$, то

  $$Vol_{n-1}(M) \leq Vol_{n-1}(M').$$

Введём следующие обозначения: пусть $d > 0$ и $u \in R$, тогда записи Δ_d и Δ_d^u обозначают множество $d\{1, \ldots, n\}^{\triangle}$ и $\xi^0 A^u + \Delta_d$ соответственно. Из сказанного выше следует, что

$$\forall\, d > 0 \quad Vol_{n-1}(\Delta_d) = Vol_{n-1}(\Delta_d^u) = d^{n-1}C. \tag{4.8}$$

Нетрудно видеть, что

(A) $\forall\, u \in R \quad \Delta_\delta^u \subseteq \Delta_{1+\delta}$, т.к. если $\vec{x} \in \Delta_\delta^u = \xi^0 A^u + \Delta_\delta$, то

$$\vec{x} = \xi^0 A^u + \delta\vec{y}, \text{ где } \vec{y} \in \Delta_1.$$

Все компоненты $\vec{y}$ неотрицательны и $\vec{y}I = 1$, поэтому все компоненты $\vec{x}$ неотрицательны и

$$\vec{x}I = (\xi^0 A^u + \delta\vec{y})I = \xi^0 A^u I + \delta\vec{y}I = 1 + \delta,$$

откуда следует, что $\vec{x} \in \Delta_{1+\delta}$.

(B) $\forall\, u, u' \in R,\ u \neq u' \ \Rightarrow\ \Delta_\delta^u \cap \Delta_\delta^{u'} = \emptyset$, т.к. если $\exists\, \vec{x} \in \Delta_\delta^u \cap \Delta_\delta^{u'}$, то $\exists\, \vec{y}, \vec{z} \in \Delta_1$:

$$\vec{x} = \xi^0 A^u + \delta\vec{y}, \quad \vec{x} = \xi^0 A^{u'} + \delta\vec{z}. \tag{4.9}$$

Из (4.9) следует, что $\xi^0(A^u - A^{u'}) = \delta(\vec{z} - \vec{y})$, откуда на основании (4.7) получаем: $|\delta(\vec{z} - \vec{y})\lambda'| > 2\delta$, или

$$|(\vec{z} - \vec{y})\lambda'| > 2, \text{ где } \lambda' \in [0, 1]^n. \tag{4.10}$$

Если $\vec{z} = (z_1, \ldots, z_n), \vec{y} = (y_1, \ldots, y_n), \lambda' = (\lambda'_1, \ldots, \lambda'_n)^{\sim}$, то

$$\begin{aligned}|(\vec{z} - \vec{y})\lambda'| = |\sum_{i=1}^{n}(z_i - y_i)\lambda'_i| \leq \sum_{i=1}^{n}|z_i - y_i|\lambda'_i \leq \\ \leq \sum_{i=1}^{n}|z_i - y_i| \leq \sum_{i=1}^{n} z_i + \sum_{i=1}^{n} y_i = 2,\end{aligned}$$

что противоречит соотношению (4.10).

Из (A), (B) и из перечисленных выше свойств функции Vol_{n-1} следует, что если R содержит k элементов $u_1, \ldots, u_k$, то

$$\sum_{i=1}^{k} Vol_{n-1}(\Delta_\delta^{u_i}) \leq Vol_{n-1}(\Delta_{1+\delta}),$$

откуда, ввиду (4.8), следует неравенство

$$k\delta^{n-1}C \leq (1+\delta)^{n-1}C,$$

которое эквивалентно неравенству $k \leq (1+\frac{1}{\delta})^{n-1}$, откуда следует регулярность A_a и неравенство (4.5). ■

Следующая теорема показывает, что в ВА ограниченного размера можно представлять сколь угодно сложные регулярные языки.

Теорема 36.

Пусть $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ – ВА, где

$$\begin{array}{l} |S_A| = 2,\ X = \{0, 2\},\ \xi^0 = (1\ \ 0), \\ A^0 = \begin{pmatrix} 1 & 0 \\ 2/3 & 1/3 \end{pmatrix},\ A^2 = \begin{pmatrix} 1/3 & 2/3 \\ 0 & 1 \end{pmatrix},\ \lambda = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \end{array}$$

и $\forall n \geq 1\ \ a_n$ – число из $(0, 1)$, имеющее в троичной записи вид $0.2\ldots211$ (количество двоек $= n-1$).

Тогда $\forall n \geq 1\ \ a_n \in I(A)$ и $w(A_{a_n})\ \geq n$.

Доказательство.

Индукцией по длине строки $u \in X^*$ доказывается, что если u имеет вид $x_1 \ldots x_k$, то $f_A(u) = 0.x_k \ldots x_1$ (в троичной записи).

Обозначим символом D топологическое замыкание множества $\{f_A(u) \mid u \in X^*\}$. Множество D называется **канторовым дисконтинуумом**. Нетрудно доказать, что $[0, 1] \setminus D \subseteq I(A)$.

Из определения a_n следует, что $a_n < f_A(u)\ \Leftrightarrow\ u = u_1 2 \ldots 2$ (количество двоек $\geq n$), поэтому если $u \in A_{a_n}$, то $|u| \geq n$, откуда следует свойство $w(A_{a_n}) \geq n$. ■

Нижеследующая теорема является обобщением теоремы 35.

Теорема 37.

Пусть M – автомат Мура вида $(X, \{0, 1\}, S, \delta, \lambda, s^0)$, причем S – компактное метрическое пространство с метрикой ρ, такой, что для любой пары s_1, s_2 достижимых состояний автомата M

выполнены условия:

$$\forall\, x \in X \quad \rho(s_1x, s_2x) \le \rho(s_1, s_2), \tag{4.11}$$

$$\exists\, \delta > 0 :\ \lambda(s_1) \ne \lambda(s_2) \ \Rightarrow\ \rho(s_1, s_2) \ge \delta. \tag{4.12}$$

Тогда язык $U \stackrel{\text{def}}{=} \{u \in X^* \mid f_M(u) = 1\}$ регулярен, и $w(U)$ не превосходит числа элементов в минимальном (по числу элементов) покрытии множества S открытыми шарами радиуса $\delta/2$.

Доказательство.

Обозначим символом R множество, содержащее по одному представителю каждого класса эквивалентности $\underset{U}{\sim}$. Из (4.12) и (4.11) следует, что $\forall\, u, u' \in R$: $u \ne u'$ $\exists\, v \in X^*$:

$$\delta \le \rho(s^0uv, s^0u'v) \le \rho(s^0u, s^0u'). \tag{4.13}$$

Пусть $\mathcal{C}$ – минимальное (по числу элементов) конечное покрытие S открытыми шарами радиуса $\delta/2$. Если u, u' – различные элементы R, то из (4.13) следует, что s^0u и s^0u' не могут попасть в один и тот же шар из $\mathcal{C}$, поэтому $|R| \le |\mathcal{C}|$. ■

Теорема 35 является следствием теоремы 37, т.к. если заданы ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ и число $a \in [0, 1)$, удовлетворяющие условиям теоремы 35, то, полагая

$$M \stackrel{\text{def}}{=} (X, \{0, 1\}, S_A^{\triangle}, \delta, \lambda_M, \xi^0),$$

где $\delta(\xi, x) \stackrel{\text{def}}{=} \xi A^x$, $\quad \lambda_M(\xi) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{если } \xi\lambda > a, \\ 0, & \text{иначе,} \end{cases}$ получаем:

$$\forall\, u \in X^* \quad u \in A_a \ \Leftrightarrow\ f_M(u) = 1.$$

Множество $S_A^{\triangle}$ является компактным метрическим пространством относительно метрики

$$\rho(\vec{x}, \vec{y}\,) = \max_{i=1\ldots n} |x_i - y_i|, \tag{4.14}$$

т.к. оно является ограниченным подмножеством $\mathbf{R}^n$. Нетрудно доказать, что для любой пары s_1, s_2 достижимых состояний автомата M условия (4.11) и (4.12) выполнены.

Следующая теорема также является следствием теоремы 37.

Теорема 38.
Пусть M – автомат Мура вида

$$(X, \{0,1\}, \{1,\ldots,n\}^{\triangle}, \delta, \lambda, \xi^0),$$

где $|X| < \infty$, и для любой пары s_1, s_2 достижимых состояний автомата M выполнены условия (4.11) и (4.12), где метрика ρ определяется соотношением (4.14).

Тогда язык $U \stackrel{\text{def}}{=} \{u \in X^* \mid f_M(u) = 1\}$ регулярен, и

$$w(U) \leq C^m_{n+m-1}, \quad \text{где } m =]\tfrac{2}{\delta}[. \tag{4.15}$$

Доказательство.
Регулярность U непосредственно следует из теоремы 37.

Для доказательства неравенства (4.15) определим покрытие множества $\{1,\ldots,n\}^{\triangle}$ открытыми шарами радиуса $\frac{1}{m}$ (поскольку $\frac{1}{m} \leq \frac{\delta}{2}$, то, согласно теореме 37, $w(U)$ не превосходит числа элементов в этом покрытии).

Обозначим записью F^n_m совокупность точек $\{1,\ldots,n\}^{\triangle}$, имеющих вид $(\frac{k_1}{m},\ldots,\frac{k_n}{m})$, где $k_1,\ldots,k_n$ – неотрицательные целые числа, сумма которых равна m. Нетрудно видеть, что

$$\forall\, \vec{x} \in \{1,\ldots,n\}^{\triangle}\ \exists\, \vec{y} \in F^n_m : \rho(\vec{x},\vec{y}) < \tfrac{1}{m}$$

(если $\vec{x} = (x_1, x_2, \ldots)$, то $\vec{y} = (]x_1 m[,](x_1+x_2)m[-]x_1 m[, \ldots))$, поэтому множество открытых шаров радиуса $\frac{1}{m}$ с центрами в точках из F^n_m является покрытием множества $\{1,\ldots,n\}^{\triangle}$. Число элементов в этом покрытии равно $|F^n_m| = C^m_{n+m-1}$. ■

4.5 Дефинитные языки

Пусть задано конечное множество X.

Мы будем использовать следующие обозначения.

- $\forall\, S \subseteq X^*$, $\forall\, k \geq 0$ записи S_k, $S_{\geq k}$, и т.д. обозначают множества $S \cap X^k$, $S \cap X^{\geq k}$, и т.д., соответственно.

- $\forall S_1, S_2 \subseteq X^*$ записи $S_1 + S_2$ и $S_1 \cdot S_2$ (точка в этой записи обычно опускается) обозначают множества

$$S_1 \cup S_2 \quad \text{и} \quad \{uv \mid u \in S_1, v \in S_2\}$$

  соответственно.

$S \subseteq X^*$ называется **дефинитным языком (ДЯ)**, если

$$\exists\, k \geq 0: \quad S_{\geq k} = X^* \cdot S_k.$$

Теорема 39.
Пусть задан ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$, и $\forall\, x \in X \;\; A^x > 0$.
Тогда $\forall\, a \in I(A) \;\; A_a$ – ДЯ.

Доказательство.
По предположению, $\exists\, c > 0$: $\forall\, x \in X \;\; A^x \geq c$, и $\exists\, \delta > 0$:

$$\forall\, u \in X^* \quad |f_A(u) - a| > \delta. \tag{4.16}$$

Выберем k: $(1-2c)^{k-1} < \frac{2\delta}{n|\lambda|}$, где $n = |S_A|$.
По теореме 25,

$$\forall\, u \in X^k \;\; \|A^u\| \leq (1-2c)^{k-1} < \frac{2\delta}{n|\lambda|}. \tag{4.17}$$

$\forall\, u \in X^k$, $\forall\, v \in X^*$

$$|f_A(vu) - f_A(u)| = |\xi^0 A^{vu}\lambda - \xi^0 A^u \lambda| = |\xi^0 (A^{vu} - A^u)\lambda| \leq$$
$$\leq |A^{vu} - A^u| \cdot n \cdot |\lambda| = |A^v A^u - A^u| \cdot n \cdot |\lambda| \leq \|A^u\| \cdot n \cdot |\lambda| < 2\delta.$$

(предпоследнее неравенство верно по теореме 26, а последнее – согласно (4.17)). Таким образом,

$$\forall\, u \in X^k, \forall\, v \in X^* \quad |f_A(vu) - f_A(u)| < 2\delta. \tag{4.18}$$

Докажем, что $(A_a)_{\geq k} = X^* \cdot (A_a)_k$, т.е. $\forall\, u \in X^k$, $\forall\, v \in X^*$

$$vu \in A_a \quad \Leftrightarrow \quad u \in A_a. \tag{4.19}$$

Если для некоторых $u \in X^k, v \in X^*$ соотношение (4.19) неверно, то, согласно определению ВЯ A_a и соотношению (4.16),

$$|f_A(vu) - f_A(u)| > 2\delta$$

что противоречит соотношению (4.18). ■

Теорема 40.
Пусть ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ – эргодичный.
Тогда $\forall\, a \in I(A) \quad A_a$ – ДЯ.

Доказательство.
Пусть $a \in I_\delta(A)$, и $|S_A| = n$.
Из эргодичности A следует, что $\|A^u\| \to 0$ при $|u| \to \infty$, поэтому $\exists\, k$:

$$\forall\, u \in X^k \quad \|A^u\| < \tfrac{2\delta}{n|\lambda|}.$$

Оставшаяся часть доказательства совпадает с частью доказательства теоремы 39, идущей после соотношения (4.17). ■

4.6 Языки, представимые линейными автоматами

Пусть задан ЛА L. Для каждого $a \in \mathbf{R}$ запись L_a обозначает подмножество множества X^*, определяемое следующим образом:

$$L_a \stackrel{\text{def}}{=} \{u \in X^* \mid f_L(u) > a\}. \tag{4.20}$$

Множество (4.20) называется **языком**, представимым ЛА L с точкой сечения a.

Теорема 41.
Пусть $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$ – ЛА, и $a \in \mathbf{R}$.
Тогда существует ВА A вида

$$(\vec{e}_1, \{A^x \mid x \in X\}, e_1^{\downarrow}), \tag{4.21}$$

такой, что

$$|S_A| = n + 4, \quad L_a = A_{\frac{1}{n+4}}, \quad \text{где } n = \dim L.$$

Доказательство.

Определим ЛА L' следующим образом:

$$L' \stackrel{\text{def}}{=} \left((\xi^0, 1), \{\begin{pmatrix} L^x & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \mid x \in X\}, \begin{pmatrix} \lambda \\ -a \end{pmatrix}\right),$$

где символы $\mathbf{0}$ обозначают строку и столбец соответствующего размера с нулевыми компонентами. Отметим, что

- $\dim L' = \dim L + 1$, и
- $f_{L'} = f_L - a, \;\Rightarrow\; L_a = L'_0$,

поэтому для доказательства теоремы 41 достаточно доказать следующее утверждение: для каждого ЛА L существует ВА A вида (4.21), такой, что $|S_A| = n+3$ и $L_0 = A_{\frac{1}{n+3}}$, где $n = \dim L$.

Для доказательства этого утверждения рассмотрим отдельно случаи $f_L(\varepsilon) > 0$ и $f_L(\varepsilon) \leq 0$. Пусть $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$.

- Если $f_L(\varepsilon) > 0$, то $\varepsilon \in L_0$. Определим функцию $f \in \mathbf{R}^{X^*}$ следующим образом:

  $$\forall\, u \in X^* \quad f(u) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{если } u = \varepsilon \\ f_L(u), & \text{иначе.} \end{cases}$$

  Обозначим записью L_1 ЛА $(\xi_1^0, \{L_1^x \mid x \in X\}, \lambda_1)$, где

  $$\xi_1^0 = (\xi^0, 1), \quad L_1^x = \begin{pmatrix} L^x & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \quad \lambda_1 = \begin{pmatrix} \lambda \\ 1 - f_L(\varepsilon) \end{pmatrix},$$

  где символы $\mathbf{0}$ обозначают строку и столбец соответствующего размера с нулевыми компонентами.

  Пусть P – невырожденная матрица порядка $\dim L_1$, первый столбец которой равен λ_1, а остальные столбцы ортогональны ξ_1^0.

  Нетрудно видеть, что $Pe_1^{\downarrow} = \lambda_1$, $\xi_1^0 P = \vec{e}_1$.

  Обозначим записью L_2 ЛА $(\vec{e}_1, \{P^{-1} L_1^x P \mid x \in X\}, e_1^{\downarrow})$.

  Нетрудно видеть, что $f_{L_2} = f_{L_1}$.

Затем L_2 преобразуется в искомый ВА A (в соответствии с построением, изложенным в доказательстве теоремы 21), такой, что $\exists\, b > 0$:

$$\forall\, u \in X^* \quad f_A(u) = \begin{cases} 1, & \text{если } u = \varepsilon, \\ b^{|u|+1} f_{L_2}(u) + \frac{1}{n+3}, & \text{иначе.} \end{cases}$$

- Если $f_L(\varepsilon) \leq 0$, то $\varepsilon \notin L_0$. Определим функцию $f \in \mathbf{R}^{X^*}$ следующим образом:

$$\forall\, u \in X^* \quad f(u) \stackrel{\text{def}}{=} \begin{cases} 0, & \text{если } u = \varepsilon \\ f_L(u), & \text{иначе.} \end{cases}$$

Обозначим записью L_1 ЛА $(\xi_1^0, \{L_1^x \mid x \in X\}, \lambda_1)$, где

$$\xi_1^0 = (\xi^0, 1), \quad L_1^x = \begin{pmatrix} L^x & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \quad \lambda_1 = \begin{pmatrix} \lambda \\ -f_L(\varepsilon) \end{pmatrix},$$

Пусть P – невырожденная матрица порядка $\dim L_1$, первый столбец которой равен λ_1, все столбцы кроме последнего ортогональны ξ_1^0, и $\xi_1^0 P_{n+1}^{\downarrow} = 1$, где $n = \dim L$ и $P_{n+1}^{\downarrow}$ – последний столбец P.

Нетрудно видеть, что $Pe_1^{\downarrow} = \lambda_1$, $\xi_1^0 P = \vec{e}_{n+1}$.

Обозначим записью L_2 ЛА $(\vec{e}_{n+1}, \{P^{-1} L_1^x P \mid x \in X\}, e_1^{\downarrow})$.

Нетрудно видеть, что $f_{L_2} = f_{L_1}$.

Затем L_2 преобразуется в искомый ВА A (в соответствии с построением, изложенным в доказательстве теоремы 21), такой, что $\exists\, b > 0$:

$$\forall\, u \in X^* \quad f_A(u) = \begin{cases} 0, & \text{если } u = \varepsilon, \\ b^{|u|+1} f_{L_2}(u) + \frac{1}{n+3}, & \text{иначе.} \;\blacksquare \end{cases}$$

Теорема 42.
Пусть X – конечное множество, и $S \subseteq X^*$.
Следующие условия эквивалентны:

- S – ВЯ,

- $\exists$ ЛАФ $f \in \mathbf{R}^{X^*}$, $\exists\, a \in \mathbf{R}$: $S = f_a \stackrel{\mathrm{def}}{=} \{u \in X^* \mid f(u) > a\}$,
- $\exists$ ЛАФ $f \in \mathbf{R}^{X^*}$: $S = f_0$.

В доказательстве этой теоремы используется утверждение из главы 5, что если f – ЛАФ и $a \in \mathbf{R}$, то $f - a$ – ЛАФ. $\blacksquare$

Глава 5

Алгебраические вопросы теории линейных автоматов

В этой главе мы рассматриваем некоторые алгебраические вопросы, относящиеся к линейным автомамам и связанным с ними функциям на строках. Материал этой главы будет использоваться в последующих частях.

5.1 Алгебраические свойства множества функций на строках

Пусть задано конечное множество X.

В пункте 1.2.2 было введено понятие функции на строках из X^*, и на множестве $\mathbf{R}^{X^*}$ таких функций были определены алгебраические операции суммы, разности, и умножения на числа из $\mathbf{R}$ (соотношениями (1.4) и (1.5)).

Определим другие алгебраические операции на $\mathbf{R}^{X^*}$

- Для функций $f_1, f_2 \in \mathbf{R}^{X^*}$ их **произведение** $f_1 f_2$ и **свёртка** $f_1 \circ f_2$ определяются следующим образом:

$$\forall\, u \in X^* \left\{ \begin{array}{l} (f_1 f_2)(u) \stackrel{\text{def}}{=} f_1(u) f_2(u), \\ (f_1 \circ f_2)(u) \stackrel{\text{def}}{=} \sum\limits_{u_1 u_2 = u} f_1(u_1) f_2(u_2). \end{array} \right. \tag{5.1}$$

- Для каждой функции $f \in \mathbf{R}^{X^*}$ **инвертирование** $\tilde{f}$ определяется следующим образом: $\forall\, u \in X^* \quad \tilde{f}(u) \stackrel{\text{def}}{=} f(\tilde{u})$.

Для каждого подмножества $S \subseteq X^*$ мы будем обозначать записью χ_S характеристическую функцию этого подмножества, т.е. функцию из $\mathbf{R}^{X^*}$, которая отображает каждый элемент из S в 1, и все остальные элементы из S – в 0. Если множество S одноэлементно и имеет вид $\{s\}$, то мы будем обозначать функцию $\chi_{\{s\}}$ более короткой записью χ_s.

Теорема 43.

Множество $\mathbf{R}^{X^*}$ является кольцом, в котором сумма элементов определяется первым соотношением в (1.4) а в качестве умножения выступает операция свёртки. Функция χ_ε является единицей этого кольца.

Доказательство.

Ассоциативность свёртки следует из того, что $\forall\, f_1, f_2, f_3 \in \mathbf{R}^{X^*}$ и $\forall\, u \in X^*$ значения $((f_1 \circ f_2) \circ f_3)(u)$ и $(f_1 \circ (f_2 \circ f_3))(u)$ равны значению выражения

$$\sum_{u_1 u_2 u_3 = u} f_1(u_1) f_2(u_2) f_3(u_3).$$

и, следовательно, функции $(f_1 \circ f_2) \circ f_3$ и $f_1 \circ (f_2 \circ f_3)$ совпадают.

Остальные свойства кольца для $\mathbf{R}^{X^*}$ устанавливаются непосредственной проверкой. ■

Теорема 44.

Если $f \in \mathbf{R}^{X^*}$ и $f(\varepsilon) \neq 0$, то существует единственная функция $g \in \mathbf{R}^{X^*}$, удовлетворяющая условию

$$f \circ g = g \circ f = \chi_\varepsilon. \tag{5.2}$$

Доказательство.

(5.2) эквивалентно конъюнкции следующих соотношений:

$$\begin{array}{l} f(\varepsilon)g(\varepsilon) = 1, \\ \forall\, x \in X \quad f(\varepsilon)g(x) + f(x)g(\varepsilon) = 0, \\ \forall\, x_1, x_2 \in X \quad f(\varepsilon)g(x_1x_2) + f(x_1)g(x_2) + \\ \qquad\qquad\qquad\qquad + f(x_1x_2)g(\varepsilon) = 0, \end{array} \tag{5.3}$$

и т.д.

Определим функцию $g \in \mathbf{R}^{X^*}$ индуктивно следующим образом:

$g(\varepsilon) \stackrel{\text{def}}{=} (f(\varepsilon))^{-1}$,
$\forall\, x \in X \quad g(x) \stackrel{\text{def}}{=} -(f(\varepsilon))^{-1} f(x) g(\varepsilon)$,
$\forall\, x_1, x_2 \in X \quad g(x_1 x_2) \stackrel{\text{def}}{=} -(f(\varepsilon))^{-1}(f(x_1)g(x_2) + f(x_1 x_2) g(\varepsilon))$,
и т.д.

Нетрудно видеть, что определённая таким образом функция g является единственной функцией, удовлетворяющей соотношениям (5.3). ■

Мы будем использовать следующие обозначения.

- Функцию g, удовлетворяющую условию (5.2), мы будем обозначать записью f^{-1}.
- Если $f \in \mathbf{R}^{X^*}$ и $k \geq 1$, то запись $f^{\circ k}$ обозначает
  - функцию f, если $k = 1$, и
  - функцию $f^{\circ(k-1)} \circ f$, если $k > 1$.

Лемма.
Если $f(\varepsilon) = 0$ и $u \in X^{<k}$, то $f^{\circ k}(u) = 0$.

Доказательство.
Нетрудно видеть, что $f^{\circ k}(u)$ равно сумме

$$\sum_{u_1 \ldots u_k = u} f(u_1) \ldots f(u_k). \tag{5.4}$$

Каждое слагаемое в (5.4) равно 0, т.к. если $u_1 \ldots u_k = u$ и $|u| < k$, то $\exists\, i \in \{1, \ldots, k\} : u_i = \varepsilon$, откуда по предположению следует, что $f(u_i) = f(\varepsilon) = 0$. ■

Если функция $f \in \mathbf{R}^{X^*}$ такова, что $f(\varepsilon) = 0$, то запись $\sum\limits_{k \geq 1} f^{\circ k}$ обозначает функцию из $\mathbf{R}^{X^*}$, значение которой на каждой строке $u \in X^*$ равно сумме всех чисел вида $f^{\circ k}(u)$, где $k \geq 1$. Из вышесказанного следует, что лишь конечное число чисел такого вида будет отлично от 0.

Пусть $f \in \mathbf{R}^{X^*}$ и $f(\varepsilon) \neq 1$. Поскольку функция $\chi_\varepsilon - f$ принимает на ε значение, не равное 0, то, согласно теореме 44, определена функция $(\chi_\varepsilon - f)^{-1}$. Обозначим записью f^+ функцию

$$f^+ \stackrel{\text{def}}{=} (\chi_\varepsilon - f)^{-1} - \chi_\varepsilon. \tag{5.5}$$

Мы будем называть функцию f^+ **итерацией** функции f.

Теорема 45.
Если функция $f \in \mathbf{R}^{X^*}$ такова, что $f(\varepsilon) = 0$, то $f^+ = \sum\limits_{k\geq 1} f^{\circ k}$.

Доказательство.
Доказываемое равенство эквивалентно равенству

$$(\chi_\varepsilon - f) \circ (\chi_\varepsilon + \sum_{k\geq 1} f^{\circ k}) = \chi_\varepsilon, \tag{5.6}$$

которое, в силу дистрибутивности свёртки относительно сложения и вычитания, эквивалентно равенству

$$\chi_\varepsilon - f + \sum_{k\geq 1} f^{\circ k} - f \circ \sum_{k\geq 1} f^{\circ k} = \chi_\varepsilon. \tag{5.7}$$

(5.7) можно переписать в виде

$$f + f \circ \sum_{k\geq 1} f^{\circ k} = \sum_{k\geq 1} f^{\circ k}.$$

Нетрудно доказать, что последнее равенство верно. ■

5.2 Операции на линейных автоматах

Пусть задано конечное множество X.

В пункте 1.4 было введено понятие линейного автомата (ЛА) над X. На множестве всех ЛА над X можно определить операции, аналогичные тем, которые были определены в пунктах 1.2.2 и 5.1 на множестве $\mathbf{R}^{X^*}$.

Для определения этих операций мы будем использовать

- обозначения, введенные в пункте 2.1.5,

- а также следующее обозначение: если A и B – матрицы, и A имеет вид $\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$, то запись $A \otimes B$ обозначает матрицу
$$\begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \dots & \dots & \dots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}.$$

Пусть $L_i = (\xi_i^0, \{L_i^x \mid x \in X\}, \lambda_i) \quad (i = 1, 2)$ – ЛА над X. Определим их **сумму** $L_1 + L_2$, **произведение** $L_1 L_2$ и **свёртку** $L_1 \circ L_2$ следующим образом.

- $L_1 + L_2 \stackrel{\text{def}}{=} ((\xi_1{}^0, \xi_2{}^0), \{\begin{pmatrix} L_1^x & 0 \\ 0 & L_2^x \end{pmatrix} \mid x \in X\}, \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix})$.

- $L_1 L_2 \stackrel{\text{def}}{=} (\xi_1{}^0 \otimes \xi_2{}^0, \{L_1^x \otimes L_2^x \mid x \in X\}, \lambda_1 \otimes \lambda_2)$.

- $L_1 \circ L_2 \stackrel{\text{def}}{=}$
  $\stackrel{\text{def}}{=} ((\xi_1{}^0, \xi_1{}^0 M), \{\begin{pmatrix} L_1^x & L_1^x M \\ 0 & L_2^x \end{pmatrix} \mid x \in X\}, \begin{pmatrix} 0^{\downarrow} \\ \lambda_2 \end{pmatrix})$,
  где $M = \lambda_1 \xi_2{}^0$, и $0^{\downarrow}$ – вектор-столбец с нулевыми компонентами, размерность которого равна размерности λ_1.

Пусть $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$ – ЛА над X, и $a \in \mathbf{R}$. Определим ЛА aL, $\tilde{L}$ и L^+ (называемые соответственно **произведением** a на L, **инвертированием** L и **итерацией** L) следующим образом:

- $aL \stackrel{\text{def}}{=} (\xi^0, \{L^x \mid x \in X\}, a\lambda)$,

- $\tilde{L} \stackrel{\text{def}}{=} (\tilde{\lambda}, \{(L^x)^{\sim} \mid x \in X\}, \tilde{\xi}^0)$,

- $L^+ \stackrel{\text{def}}{=} (\xi^0, \{L^x(E + \lambda\xi^0) \mid x \in X\}, \lambda)$, где E – единичная матрица.

Теорема 46.

1. Пусть X – конечное множество, и L_1, L_2 – ЛА над X. Тогда

$$f_{L_1+L_2} = f_{L_1} + f_{L_2}, \tag{5.8}$$

$$f_{L_1L_2} = f_{L_1} f_{L_2}, \tag{5.9}$$

$$f_{L_1 \circ L_2} = f_{L_1} \circ f_{L_2}. \tag{5.10}$$

2. Для любого ЛА L

$$\forall\, a \in \mathbf{R} \quad f_{aL} = a f_L, \tag{5.11}$$

$$f_{\tilde{L}} = (f_L)^{\sim}, \tag{5.12}$$

$$\text{если } f_L(\varepsilon) = 0, \text{ то } f_{L^+} = (f_L)^+. \tag{5.13}$$

Доказательство.

1. Пусть L_i $(i = 1, 2)$ имеют вид $(\xi_i^0, \{L_i^x \mid x \in X\}, \lambda_i)$.

   (a) (5.8) доказывается непосредственной проверкой, с использованием того, что

   $$\forall\, u \in X^* \quad (L_1 + L_2)^u = \begin{pmatrix} L_1^u & 0 \\ 0 & L_2^u \end{pmatrix}.$$

   (b) (5.9) непосредствено вытекает из следующего утверждения: если матрицы A, B, C, D таковы, что определены произведения AC и BD, то верно равенство

   $$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

   (c) Докажем (5.10).
   Нетрудно доказать (индукцией по $|u|$), что $\forall\, u \in X^*$

   $$(L_1 \circ L_2)^u = \begin{pmatrix} L_1^u & \sum\limits_{u_1u_2=u} L_1^{u_1} M L_2^{u_2} - M L_2^u \\ 0 & L_2^u \end{pmatrix} \tag{5.14}$$

где $M = \lambda_1 \xi_0^2$. Поэтому $\forall\, u \in X^*$

$$\begin{aligned}
&f_{L_1 \circ L_2}(u) = (\xi_1^0,\ \xi_1^0 M)(L_1 \circ L_2)^u \begin{pmatrix} 0^{\downarrow} \\ \lambda_2 \end{pmatrix} = \\
&= (\xi_1^0,\ \xi_1^0 M) \begin{pmatrix} \Big(\sum\limits_{u_1 u_2 = u} L_1^{u_1} M L_2^{u_2} - M L_2^u\Big)\lambda_2 \\ L_2^u \lambda_2 \end{pmatrix} = \\
&= \xi_1^0 \Big(\sum\limits_{u_1 u_2 = u} L_1^{u_1} M L_2^{u_2} - M L_2^u\Big)\lambda_2 + \xi_1^0 M L_2^u \lambda_2 = \\
&= \xi_1^0 \Big(\sum\limits_{u_1 u_2 = u} L_1^{u_1} M L_2^{u_2}\Big)\lambda_2 = \sum\limits_{u_1 u_2 = u} \xi_1^0 L_1^{u_1} \lambda_1 \xi_0^2 L_2^{u_2} \lambda_2 = \\
&= \sum\limits_{u_1 u_2 = u} f_{L_1}(u_1) f_{L_2}(u_2) = (f_{L_1} \circ f_{L_2})(u).
\end{aligned}$$

2. Пусть L имеет вид $(\xi^0, \{L^x \mid x \in X\}, \lambda)$.

   (a) (5.11) доказывается непосредственной проверкой:

   $$\begin{aligned} \forall\, u \in X^* \quad f_{aL}(u) &= \xi^0 L^u (a\lambda) = a \xi^0 L^u \lambda = \\ &= a f_L(u) = (a f_L)(u). \end{aligned}$$

   (b) Для доказательства (5.12) мы используем утверждение о том, что если A, B – матрицы, для которых определено произведение AB, то $(AB)^{\sim} = \tilde{B}\tilde{A}$:

   $$\begin{aligned} \forall\, u \in X^* \quad f_{\tilde{L}}(u) &= \tilde{\lambda} \tilde{L}^u \tilde{\xi}^0 = \tilde{\lambda} (L^{\tilde{u}})^{\sim} \tilde{\xi}^0 = \\ &= (\xi^0 L^{\tilde{u}} \lambda)^{\sim} = (f_L(\tilde{u}))^{\sim} = f_L(\tilde{u}). \end{aligned}$$

   (c) Докажем соотношение (5.13), т.е. $\forall\, u \in X^*$

   $$f_{L^+}(u) = (f_L)^+(u). \tag{5.15}$$

   Если $u = \varepsilon$, то левая часть (5.15) равна

   $$f_{L^+}(\varepsilon) = \xi^0 \lambda = f_L(\varepsilon) = 0.$$

   Нетрудно доказать, что правая часть (5.15) в случае $u = \varepsilon$ тоже равна 0.

   Пусть $u \neq \varepsilon$. Рассмотрим обе части (5.15) в этом случае, и докажем, что они совпадают.

i. Согласно теореме 45, из $f_L(\varepsilon) = 0$ следует, что правая часть (5.15) равна сумме $\sum\limits_{k\geq 1} f_L^{\circ k}(u)$, которая равна сумме

$$\sum_{\substack{k\geq 1,\ u_1\ldots u_k = u\\ \forall\, i\in\{1,\ldots,k\}\ \ u_i\neq\varepsilon}} f_L(u_1)\ldots f_L(u_k). \tag{5.16}$$

ii. Левая часть (5.15) имеет вид

$$\xi^0 L^{x_1}(E+\lambda\xi^0)\ldots L^{x_k}(E+\lambda\xi^0)\lambda \tag{5.17}$$

где $u = x_1 \ldots x_k$, $\forall\, i = 1,\ldots,k\ \ x_i \in X$.
Т.к. $\xi^0\lambda = 0$, то $(E+\lambda\xi^0)\lambda = \lambda$, поэтому можно переписать (5.17) в виде

$$\xi^0 L^{x_1}(E+\lambda\xi^0)\ldots L^{x_k}\lambda. \tag{5.18}$$

Раскроем все скобки (5.18). Нетрудно видеть, что каждое слагаемое в получившейся сумме имеет вид

$$\xi^0 L^{u_1}\lambda\ldots\xi^0 L^{u_k}\lambda \tag{5.19}$$

где $u_1\ldots u_k = u$ и $\forall\, i\in\{1,\ldots,k\}\ \ u_i\neq\varepsilon$. Согласно определению функции f_L, (5.19) совпадает с соответствующим слагаемым (определяемым тем же представлением u в виде конкатенации непустых подстрок $u_1\ldots u_k$) в сумме (5.16).

Таким образом, каждому слагаемому в разложении выражения (5.18) взаимно однозначно соответствует равное ему слагаемое в сумме (5.16). Следовательно, значения выражений (5.18) и (5.16) совпадают. ■

5.3 Алгебраические свойства множества линейно-автоматных функций

Пусть задано конечное множество X.

Обозначим символом $\mathcal{L}^X$ множество всех ЛАФ из $\mathbf{R}^{X^*}$.

Отметим, что $\chi_\varepsilon \in \mathcal{L}^X$, т.к. $\chi_\varepsilon = f_L$, где L – ЛА вида (1.7) размерности 1, у которого $\xi^0 = (1)$, $\forall\, x \in X$ $L^x = (0)$, $\lambda = (1)$.

Согласно вышесказанному и теореме 43, $\mathcal{L}^X$ является кольцом относительно определённых в пункте 1.2.2 операций сложения и свёртки (рассматриваемой в данном случае как умножение). Единицей этого кольца является ЛАФ χ_ε. Ниже мы будем опускать символ свёртки $\circ$ в обозначении произведения элементов кольца $\mathcal{L}^X$ (т.е. для любых элементов a, b кольца $\mathcal{L}^X$ будем обозначать их произведение $a \circ b$ записью ab).

Обозначим записью $\mathcal{L}^X_n$ кольцо квадратных матриц порядка n над $\mathcal{L}^X$. Единицей этого кольца является диагональная матрица E_ε, все компоненты диагонали которой равны χ_ε.

$\forall\, M \in \mathcal{L}^X_n$ и $\forall\, u \in X^*$ мы будем обозначать записью $M(u)$ матрицу над $\mathbf{R}$, каждая компонента $M(u)_{ij}$ которой равна значению соответствующей функции M_{ij} на строке u.

Теорема 47.

Пусть X – конечное множество, и матрица $A \in \mathcal{L}^X_n$ такова, что матрица $A(\varepsilon)$ – нулевая. Тогда существует матрица из $\mathcal{L}^X_n$, обозначаемая записью $\sum\limits_{k\geq 1} A^k$, для которой верно равенство

$$\Big(E_\varepsilon - A\Big)\Big(E_\varepsilon + \sum_{k\geq 1} A^k\Big) = E_\varepsilon. \tag{5.20}$$

Доказательство.

$\forall\, u \in X^*$, $\forall\, k > |u|$ матрица $A^k(u)$ является нулевой, т.к. каждая компонента матрицы A^k является суммой произведений вида $a_{i_1 j_1} \ldots a_{i_k j_k}$, где $a_{i_s j_s}$ $(s = 1, \ldots, k)$ – компоненты матрицы A. Нетрудно видеть, что

$$(a_{i_1 j_1} \ldots a_{i_k j_k})(u) = \sum_{u_1 \ldots u_k = u} a_{i_1 j_1}(u_1) \ldots a_{i_k j_k}(u_k). \tag{5.21}$$

Каждое слагаемое в сумме в правой части (5.21) равно 0, т.к. если $u_1 \ldots u_k = u$ и $|u| < k$, то $\exists\, s \in \{1, \ldots, k\} : u_s = \varepsilon$, откуда по предположению следует, что $a_{i_s j_s}(u_s) = a_{i_s j_s}(\varepsilon) = 0$.

Искомая матрица $\sum\limits_{k\geq 1} A^k$ определяется как матрица из $\mathcal{L}^X_n$, значение которой на каждой строке $u \in X^*$ равно сумме всех матриц вида $A^k(u)$, где $k \geq 1$. Из вышесказанного следует, что лишь конечное число матриц такого вида отлично от нулевой.

Равенство (5.20) доказывается так же, как доказывается аналогичное ему равенство (5.6), и следует из того, что $\forall\, u \in X^*$

$$\Big(A + A\Big(\sum_{k\geq 1} A^k\Big)\Big)(u) = \Big(\sum_{k\geq 1} A^k\Big)(u). \tag{5.22}$$

Обоснуем (5.22).

Пусть $B = \sum\limits_{k=1}^{|u|} A^k$ и $C = \sum\limits_{k>|u|} A^k$. Нетрудно видеть, что

- матрица $(AC)(u)$ – нулевая, т.к. каждая компонента матрицы AC является суммой произведений вида $a_{ij}c_{jk}$, где a_{ij} и c_{jk} – компоненты A и C соответственно, и поскольку

$$(a_{ij}c_{jk})(u) = \sum_{u_1u_2=u} a_{ij}(u_1)c_{jk}(u_2) \tag{5.23}$$

то из $|u_2| \leq |u|$ следует, что матрица $C(u_2)$ является нулевой, в частности $c_{jk}(u_2) = 0$, поэтому правая (а значит и левая) часть (5.23) равна 0,

- $\sum\limits_{k\geq 1} A^k = B + C$, поэтому, используя утверждение из предыдущего пункта, левую часть (5.22) можно переписать в виде

$$\begin{aligned}(A + A(B + C))(u) &= A(u) + (AB)(u) + (AC)(u) = \\ &= A(u) + (AB)(u) + 0 = A(u) + \sum_{k=2}^{|u|+1} A^k(u).\end{aligned} \tag{5.24}$$

Значение последнего выражения в цепочке равенств (5.24) совпадает с правой частью (5.22). ■

Пусть X – конечное множество. Мы будем использовать следующие обозначения.

- Символ Ω обозначает совокупность операций на $\mathbf{R}^{X^*}$, состоящую из определённых в пункте 1.2.2 операций сложения, свёртки, итерации (для тех функций, которые отображают ε в 0) и умножения на числа из $\mathbf{R}$.

- Для произвольного подмножества $F \subseteq \mathbf{R}^{X^*}$ запись ΩF обозначает множество всех функций из $\mathbf{R}^{X^*}$, которые могут быть получены из функций из F при помощи применения операций из Ω.

Теорема 48.

Пусть X – конечное множество, n – натуральное число, A – матрица из $\mathcal{L}_n^X$, и B, C – вектор-столбцы размерности n над $\mathcal{L}^X$, имеющие вид

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ \dots \\ b_n \end{pmatrix}, \quad C = \begin{pmatrix} c_1 \\ \dots \\ c_n \end{pmatrix}$$

причём матрица $A(\varepsilon)$ – нулевая, и верно равенство

$$(E_\varepsilon - A)B = C. \tag{5.25}$$

Тогда

$$\forall\, i = 1, \dots, n \quad b_i \in \Omega\{\chi_\varepsilon, a_{ij}, c_i \mid i, j = 1, \dots, n\}. \tag{5.26}$$

Доказательство.

Докажем теорему индукцией по n.

Если $n = 1$, то (5.25) имеет вид $(\chi_\varepsilon - a_{11})b_1 = c_1$. Поскольку по предположению $a_{11}(\varepsilon) = 0$, то к a_{11} можно применить операцию итерации, и, согласно определению (5.5), верны равенства

$$b_1 = (\chi_\varepsilon + a_{11}^+)(\chi_\varepsilon - a_{11})b_1 = (\chi_\varepsilon + a_{11}^+)c_1$$

т.е. в данном случае (5.26) верно.

Пусть $n > 1$. Перепишем (5.25) в виде системы равенств

$$\left\{ \begin{array}{ll} (\chi_\varepsilon - a_{11})b_1 - \ldots - a_{1,n-1}b_{n-1} - a_{1n}b_n & = c_1 \\ \dots & \\ -a_{n1}b_1 - \ldots - a_{n,n-1}b_{n-1} + (\chi_\varepsilon - a_{nn})b_n & = c_n \end{array} \right. \tag{5.27}$$

Используя рассуждения, аналогичные вышеизложенным, получаем, что последнее равенство в (5.27) равносильно равенству

$$b_n = (\chi_\varepsilon + a_{nn}^+)(c_n + a_{n1}b_1 + \ldots + a_{n,n-1}b_{n-1}). \tag{5.28}$$

Заменим в (5.27) во всех равенствах, кроме последнего, выражение b_n на правую часть равенства (5.28). После раскрытия скобок и приведения подобных членов совокупность этих $n-1$ равенств будет представлять собой систему, которая в матричной записи имеет вид

$$(E_\varepsilon - A')B' = C', \tag{5.29}$$

где компоненты a'_{ij}, b'_i, c'_i $(i, j = 1, \ldots, n-1)$ матриц A', B', C' имеют следующий вид:

$$\begin{aligned} a'_{ij} &= a_{ij} + a_{in}(\chi_\varepsilon + a^+_{nn})a_{nj}, \\ b'_i &= b_i, \\ c'_i &= c_i + a_{in}(\chi_\varepsilon + a^+_{nn})c_n. \end{aligned} \tag{5.30}$$

Поскольку $\forall\, f, g \in \mathbf{R}^{X^*}$ верно равенство $(fg)(\varepsilon) = f(\varepsilon)g(\varepsilon)$, то из (5.30) и из того, что матрица $A(\varepsilon)$ нулевая, следует, что матрица $A'(\varepsilon)$ нулевая. Таким образом, к матрицам A', B', C' можно применить индуктивное предположение, согласно которому будет верно утверждение теоремы 48, в котором матрицы A, B, C заменены на A', B', C' соответственно, т.е.

$$\forall\, i = 1, \ldots, n-1 \quad b_i \in \Omega\{\chi_\varepsilon, a'_{ij}, c'_i \mid i, j = 1, \ldots, n-1\}. \tag{5.31}$$

Из (5.30) следует, что

$$\forall\, i, j = 1, \ldots, n-1 \quad a'_{ij}, c'_i \in \Omega\{\chi_\varepsilon, a_{ij}, c_i \mid i, j = 1, \ldots, n\}. \tag{5.32}$$

Из (5.31) и (5.32) следует, что

$$\forall\, i = 1, \ldots, n-1 \quad b_i \in \Omega\{\chi_\varepsilon, a_{ij}, c_i \mid i, j = 1, \ldots, n\}. \tag{5.33}$$

Из (5.28) и (5.33) следует, что

$$b_n \in \Omega\{\chi_\varepsilon, a_{ij}, c_i \mid i, j = 1, \ldots, n\}. \tag{5.34}$$

Из (5.33) и (5.34) следует утверждение теоремы 48. ■

Теорема 49.

Пусть X – конечное множество. Тогда

$$\mathcal{L}^X = \Omega\{\chi_\varepsilon, \chi_x \mid x \in X\}. \tag{5.35}$$

Доказательство.

Как было отмечено в начале пункта 5.3, $\chi_\varepsilon \in \mathcal{L}^X$. Кроме того, $\forall\, x \in X \;\; \chi_x \in \mathcal{L}^X$, т.к. $\chi_x = f_L$, где L – ЛА вида (1.7) размерности 2, компоненты которого имеют следующий вид:

$$\xi^0 = (1,0),\;\; L^x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\; L^y = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \text{ для } y \neq x,\; \lambda = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Согласно теореме 46, $\mathcal{L}^X$ замкнуто относительно операций из Ω. Следовательно, правая часть соотношения (5.35) содержится в левой его части.

Докажем, что верно и обратное включение.

Пусть $f = f_L$, где L – ЛА вида $(\xi^0, \{L^x \mid x \in X\}, \lambda)$.

$\forall\, u \in X^*$ обозначим записью L_u матрицу из $\mathcal{L}_n^X$ (где n – размерность ЛА L), имеющую вид $L^u \chi_u$ (т.е. каждая компонента L_u представляет собой произведение соответствующей компоненты матрицы L^u и ЛАФ χ_u).

Обозначим символом A матрицу $\sum\limits_{x \in X} L_x$. Нетрудно видеть, что

$$\forall\, k \geq 1 \;\; A^k = \sum_{x_1,\ldots,x_k \in X} L_{x_1} \ldots L_{x_k} = \sum_{u \in X^k} L_u. \tag{5.36}$$

Последнее равенство в (5.36) верно потому, что

$$\forall\, x_1, \ldots, x_k \in X \quad L_{x_1} \ldots L_{x_k} = L_{x_1 \ldots x_k}. \tag{5.37}$$

(5.37) следует из равенства $\chi_{x_1} \ldots \chi_{x_k} = \chi_{x_1 \ldots x_k}$.

Поскольку матрица $A(\varepsilon)$ – нулевая, то, согласно теореме (47), существует матрица $\sum\limits_{k \geq 1} A^k$, для которой верно равенство (5.20).

Компоненты матрицы A представляют собой суммы ЛАФ вида $a\chi_x$, где $a \in \mathbf{R}$ и $x \in X$. Следовательно, используя равенство (5.20) и теорему 48, применяемую к столбцам матрицы $E_\varepsilon + \Big(\sum\limits_{k \geq 1} A^k\Big)$, можно заключить, что компоненты этой матрицы принадлежат множеству $\Omega\{\chi_\varepsilon, \chi_x \mid x \in X\}$.

Из (5.36) следует, что

$$\forall\, u \in X^* \quad \Big(E_\varepsilon + \Big(\sum_{k \geq 1} A^k\Big)\Big)(u) = L^u. \tag{5.38}$$

Напомним, что $\forall\, u \in X^*$

$$f_L(u) = \xi^0 L^u \lambda. \tag{5.39}$$

Обозначим компоненты векторов ξ^0, λ и матриц L^u, $E_\varepsilon + \Big(\sum\limits_{k\geq 1} A^k\Big)$ записями ξ^0_i, λ_j, L^u_{ij} и a_{ij} соответственно (где $i, j = 1, \ldots, n$). В этих обозначениях равенство (5.39) можно переписать в виде

$$f_L(u) = \sum_{i=1}^n \sum_{j=1}^n \xi^0_i L^u_{ij} \lambda_j \tag{5.40}$$

Определим функцию f как линейную комбинацию функций a_{ij} вида

$$f \stackrel{\rm def}{=} \sum_{i=1}^n \sum_{j=1}^n \xi^0_i \lambda_j a_{ij}. \tag{5.41}$$

Докажем, что f совпадает с f_L. По определению f, $\forall\, u \in X^*$

$$f(u) = \sum_{i=1}^n \sum_{j=1}^n \xi^0_i \lambda_j a_{ij}(u). \tag{5.42}$$

Из (5.38) следует, что $\forall\, i, j = 1, \ldots, n \;\; a_{ij}(u) = L^u_{ij}$, поэтому можно переписать (5.42) в виде

$$f(u) = \sum_{i=1}^n \sum_{j=1}^n \xi^0_i \lambda_j L^u_{ij}. \tag{5.43}$$

Поскольку правая часть (5.43) совпадает с правой частью (5.40), то, следовательно, и левые части этих равенств совпадают, т.е. верно равенство $f(u) = f_L(u)$, что и требовалось доказать.

Как было отмечено выше, все функции a_{ij} принадлежат множеству $\Omega\{\chi_\varepsilon, \chi_x \mid x \in X\}$, поэтому, согласно определению (5.41), функция f (т.е. f_L) тоже принадлежит этому множеству. ■

5.4 Линейные пространства, связанные с линейно-автоматными функциями

Пусть X – конечное множество, и $f \in \mathbf{R}^{X^*}$.

Мы будем использовать следующие обозначения: $\forall\, u \in X^*$ f_u обозначает функцию из $\mathbf{R}^{X^*}$, определяемую следующим образом:

$$\forall\, u_1 \in X^* \quad f_u(u_1) \stackrel{\text{def}}{=} f(uu_1),$$

и E_f обозначает подпространство линейного пространства $\mathbf{R}^{X^*}$, порождённое множеством $\{f_u \mid u \in X^*\}$.

Теорема 50.

Пусть X – конечное множество, $f \in \mathbf{R}^{X^*}$, и n – натуральное число. Следующие условия эквивалентны:

1. существует ЛА L размерности не больше n над X, такой, что $f_L = f$,
2. $\dim E_f \leq n$.

Доказательство.

Докажем, что из условия 1 следует условие 2.

Пусть ЛА $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$ имеет размерность $k \leq n$ и удовлетворяет условию $f_L = f$, т.е.

$$\forall\, u \in X^* \quad f(u) = \xi^0 L^u \lambda,$$

откуда следует, что $\forall\, u \in X^*$ функция f_u удовлетворяет условию

$$\forall\, u_1 \in X^* \quad f_u(u_1) = f(uu_1) = \xi^0 L^{uu_1} \lambda = (\xi^0 L^u) L^{u_1} \lambda.$$

$\forall\, i = 1, \ldots, k$ обозначим

- записью e_i вектор–строку из $\mathbf{R}^k$, i–я компонента которой равна 1, а все остальные компоненты равны 0, и
- записью f_i – функцию из $\mathbf{R}^{X^*}$, определяемую следующим образом:
$$\forall\, u \in X^* \quad f_i(u) = e_i L^u \lambda.$$

Поскольку $\{e_1, \ldots, e_k\}$ – базис в $\mathbf{R}^k$, то существуют числа $a_1, \ldots, a_k \in \mathbf{R}$, такие, что $\xi^0 L^u = \sum\limits_{i=1}^{k} a_i e_i$. Следовательно,

$$\forall\, u_1 \in X^* \quad f_u(u_1) = \sum_{i=1}^{k} a_i e_i L^{u_1} \lambda = \sum_{i=1}^{k} a_i f_i(u_1),$$

т.е. $f_u = \sum\limits_{i=1}^{k} a_i f_i$.

Таким образом, $\forall\, u \in X^*$ функция f_u принадлежит линейному подпространству в $\mathbf{R}^{X^*}$, порождённому функциями $f_1, \ldots, f_k$. Отсюда непосредственно следует условие 2.

Теперь докажем, что из условия 2 следует условие 1.

Определим компоненты ξ^0, L^x $(x \in X)$, λ искомого ЛА L следующим образом. Выберем базис в пространстве E_f, имеющий вид $\{f_{u_1}, \ldots, f_{u_k}\}$, где $u_1, \ldots, u_k \in X^*$.

- Т.к. $f = f_\varepsilon \in E_f$, то $\exists\, \xi_1^0, \ldots, \xi_k^0 \in \mathbf{R} : f = \sum\limits_{i=1}^{k} \xi_i^0 f_{u_i}$.

  Определим $\xi^0 \stackrel{\text{def}}{=} (\xi_1^0, \ldots, \xi_k^0)$.

- $\forall\, x \in X \quad L^x \stackrel{\text{def}}{=} \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ \ldots & \ldots & \ldots \\ a_{k1} & \ldots & a_{kk} \end{pmatrix}$, где $\forall\, i = 1, \ldots, k$ числа a_{i1}, …, a_{ik} являются коэффициентами разложения функции $f_{u_i x}$ по базису $\{f_{u_1}, \ldots, f_{u_k}\}$: $f_{u_i x} = \sum\limits_{j=1}^{k} a_{ij} f_{u_j}$.

- $\lambda \stackrel{\text{def}}{=} \begin{pmatrix} f_{u_1}(\varepsilon) \\ \ldots \\ f_{u_k}(\varepsilon) \end{pmatrix}$.

Докажем (индукцией по $|u|$), что $\forall\, u, v \in X^*$ верно равенство

$$L^u \begin{pmatrix} f_{u_1}(v) \\ \ldots \\ f_{u_k}(v) \end{pmatrix} = \begin{pmatrix} f_{u_1}(uv) \\ \ldots \\ f_{u_k}(uv) \end{pmatrix}. \qquad (5.44)$$

- Если $u = \varepsilon$, то (5.44) верно по определению λ.

- Пусть (5.44) верно для некоторого u и произвольного v.

Тогда $\forall\, x \in X$

$$L^{ux}\begin{pmatrix} f_{u_1}(v) \\ \dots \\ f_{u_k}(v) \end{pmatrix} = L^u L^x \begin{pmatrix} f_{u_1}(v) \\ \dots \\ f_{u_k}(v) \end{pmatrix} = L^u \begin{pmatrix} \sum\limits_{j=1}^{k} a_{1j} f_{u_j}(v) \\ \dots \\ \sum\limits_{j=1}^{k} a_{kj} f_{u_j}(v) \end{pmatrix} =$$

$$= L^u \begin{pmatrix} f_{u_1x}(v) \\ \dots \\ f_{u_kx}(v) \end{pmatrix} = L^u \begin{pmatrix} f_{u_1}(xv) \\ \dots \\ f_{u_k}(xv) \end{pmatrix} = \begin{pmatrix} f_{u_1}(uxv) \\ \dots \\ f_{u_k}(uxv) \end{pmatrix}.$$

Таким образом, (5.44) верно для произвольных $u, v \in X^*$.

Используя (5.44), докажем, что $f_L = f$.

$$\forall\, u \in X^* \quad f_L(u) = \xi^0 L^u \lambda = \xi^0 L^u \begin{pmatrix} f_{u_1}(\varepsilon) \\ \dots \\ f_{u_k}(\varepsilon) \end{pmatrix} =$$

$$= \xi^0 \begin{pmatrix} f_{u_1}(u) \\ \dots \\ f_{u_k}(u) \end{pmatrix} = \sum_{i=1}^{k} \xi^0_i f_{u_i}(u) = f(u). \blacksquare$$

Из теоремы 50 следует, что если для функции $f \in \mathbf{R}^{X^*}$ существует ЛА L над X, такой, что $f_L = f$, то $\dim E_f$ является наименьшей возможной размерностью такого ЛА.

Пусть X – конечное множество. Мы будем обозначать записью $\mathbf{R}\langle X\rangle$ кольцо многочленов с коэффициентами из $\mathbf{R}$ от некоммутирующих переменных из множества X. Каждый элемент $p \in \mathbf{R}\langle X\rangle$ можно рассматривать как формальную сумму вида $\sum\limits_{i=1}^{n} a_i u_i$, где $\forall\, i = 1, \dots, n \;\; a_i \in \mathbf{R}, u_i \in X^*$.

Для каждой функции $f \in \mathbf{R}^{X^*}$ мы будем обозначать тем же символом f линейное продолжение этой функции на $\mathbf{R}\langle X\rangle$, т.е. функцию вида $\mathbf{R}\langle X\rangle \to \mathbf{R}$ определяемую следующим образом:

$$\forall\, p \in \mathbf{R}\langle X\rangle, \;\; \text{если } p = \sum_{i=1}^{n} a_i u_i, \; \text{ то } f(p) \stackrel{\text{def}}{=} \sum_{i=1}^{n} a_i f(u_i).$$

$\forall\, f \in \mathbf{R}^{X^*}$ мы будем обозначать записью $\sim_f$ отношение эквивалентности на $\mathbf{R}\langle X\rangle$, определяемое следующим образом:

$$\forall\, p_1, p_2 \in \mathbf{R}\langle X\rangle \;\; p_1 \sim_f p_2 \;\Leftrightarrow\; \forall\, u \in X^* \;\; f(p_1 u) = f(p_2 u).$$

Для каждого $p \in \mathbf{R}\langle X\rangle$ мы будем обозначать записью $[p]$ класс эквивалентности $\sim_f$, содержащий p. Нетрудно видеть, что $\sim_f$ сохраняет операции сложения и умножения на числа из $\mathbf{R}$, поэтому определено факторпространство $\mathbf{R}\langle X\rangle/_{\sim_f}$.

Теорема 51.
Пусть X – конечное множество, $f \in \mathbf{R}^{X^*}$. Тогда

$$\dim E_f = \dim\Big(\mathbf{R}\langle X\rangle/_{\sim_f}\Big).$$

Доказательство.
Докажем, что для каждого натурального числа n верно соотношение

$$\dim E_f \leq n \quad \Leftrightarrow \quad \dim \mathbf{R}\langle X\rangle/_{\sim_f} \leq n. \tag{5.45}$$

Левая часть (5.45) равносильна условию

$$\begin{array}{l} \exists\, u_1, \ldots, u_n \in X^* : \forall\, u \in X^*\ \exists\, a_1, \ldots, a_n \in \mathbf{R} : \\ \forall\, v \in X^* \quad f(uv) = \sum\limits_{i=1}^{n} a_i f(u_i v). \end{array} \tag{5.46}$$

Правая часть (5.45) равносильна условию

$$\begin{array}{l} \exists\, p_1, \ldots, p_n \in \mathbf{R}\langle X\rangle : \forall\, p \in \mathbf{R}\langle X\rangle\ \exists\, a_1, \ldots, a_n \in \mathbf{R} : \\ [p] = \sum\limits_{i=1}^{n} a_i [p_i]. \end{array} \tag{5.47}$$

Нетрудно доказать, что (5.47) эквивалентно условию

$$\begin{array}{l} \exists\, u_1, \ldots, u_n \in X^* : \forall\, u \in X^*\ \exists\, a_1, \ldots, a_n \in \mathbf{R} : \\ [u] = \sum\limits_{i=1}^{n} a_i [u_i] = [\sum\limits_{i=1}^{n} a_i u_i]. \end{array} \tag{5.48}$$

Равенство $[u] = [\sum\limits_{i=1}^{n} a_i u_i]$ равносильно условию $u \sim_f \sum\limits_{i=1}^{n} a_i u_i$, т.е. соотношению

$$\forall\, v \in X^* \quad f(uv) = f\Big(\Big(\sum_{i=1}^{n} a_i u_i\Big) v\Big) = \sum_{i=1}^{n} a_i f(u_i v).$$

Используя последнее замечание, заключаем, что условия (5.46) и (5.48) эквивалентны, откуда следует доказываемое соотношение (5.45). ■

Пусть X – конечное множество, и $f \in \mathbf{R}^{X^*}$.
Мы будем использовать следующие обозначения.

- $\forall\, u \in X^*$ запись $\hat{f}_u$ обозначает функцию из $\mathbf{R}^{X^* \times X^*}$, т.е.

$$\hat{f}_u : X^* \times X^* \to \mathbf{R},$$

  определяемую следующим образом:

$$\forall\, v_1, v_2 \in X^* \quad \hat{f}_u(v_1, v_2) \stackrel{\text{def}}{=} f(v_1 u v_2).$$

- Запись $E_{\hat{f}}$ обозначает подпространство линейного пространства $\mathbf{R}^{X^* \times X^*}$, порождённое множеством $\{\hat{f}_u \mid u \in X^*\}$.

- Запись I_f обозначает идеал кольца $\mathbf{R}\langle X\rangle$, определяемый следующим образом:

$$I_f \stackrel{\text{def}}{=} \{p \in \mathbf{R}\langle X\rangle \mid \forall\, p_1, p_2 \in \mathbf{R}\langle X\rangle\ f(p_1 p p_2) = 0\}$$

  Нетрудно доказать, что I_f действительно является идеалом, и, следовательно, определено фактор-кольцо $\mathbf{R}\langle X\rangle / I_f$, которое также является линейным пространством над $\mathbf{R}$. $\forall\, p \in \mathbf{R}\langle X\rangle$ мы будем обозначать соответствующий элемент фактор-кольца $\mathbf{R}\langle X\rangle / I_f$ записью $[p]$.

Теорема 52.
Пусть X – конечное множество, $f \in \mathbf{R}^{X^*}$. Тогда

$$\dim E_{\hat{f}} = \dim \mathbf{R}\langle X\rangle / I_f.$$

Доказательство.
Докажем, что для каждого натурального числа n верно соотношение

$$\dim E_{\hat{f}} \leq n \quad \Leftrightarrow \quad \dim \mathbf{R}\langle X\rangle / I_f \leq n. \tag{5.49}$$

Левая часть (5.49) равносильна условию

$$\begin{array}{l} \exists\, u_1, \ldots, u_n \in X^* : \forall\, u \in X^*\ \exists\, a_1, \ldots, a_n \in \mathbf{R} : \\ \forall\, v_1, v_2 \in X^* \quad f(v_1 u v_2) = \sum\limits_{i=1}^{n} a_i f(v_1 u_i v_2). \end{array} \tag{5.50}$$

Правая часть (5.49) равносильна условию

$$\begin{aligned}&\exists\, p_1, \ldots, p_n \in \mathbf{R}\langle X\rangle : \forall\, p \in \mathbf{R}\langle X\rangle\ \exists\, a_1, \ldots, a_n \in \mathbf{R} :\\ &[p] = \sum_{i=1}^{n} a_i [p_i].\end{aligned} \tag{5.51}$$

Нетрудно доказать, что (5.51) эквивалентно условию

$$\begin{aligned}&\exists\, u_1, \ldots, u_n \in X^* : \forall\, u \in X^*\ \exists\, a_1, \ldots, a_n \in \mathbf{R} :\\ &[u] = \sum_{i=1}^{n} a_i [u_i] = [\sum_{i=1}^{n} a_i u_i].\end{aligned} \tag{5.52}$$

Равенство $[u] = [\sum_{i=1}^{n} a_i u_i]$ равносильно условию $u - \sum_{i=1}^{n} a_i u_i \in I_f$, т.е. соотношению

$$\forall\, v_1, v_2 \in X^* \quad f(v_1(u - \sum_{i=1}^{n} a_i u_i) v_2) = 0$$

которое можно переписать в виде

$$\forall\, v_1, v_2 \in X^* \quad f(v_1 u v_2) = \sum_{i=1}^{n} a_i f(v_1 u_i v_2).$$

Используя последнее замечание, заключаем, что условия (5.50) и (5.52) эквивалентны, откуда следует доказываемое соотношение (5.49). ■

Теорема 53.

Пусть X – конечное множество, $f \in \mathbf{R}^{X^*}$. Тогда для каждого натурального числа n верны импликации

$$\dim E_{\hat{f}} \leq n \quad\Rightarrow\quad \dim E_f \leq n, \tag{5.53}$$

$$\dim E_f \leq n \quad\Rightarrow\quad \dim E_{\hat{f}} \leq n^2. \tag{5.54}$$

Доказательство.

Докажем импликацию (5.53).

Как было отмечено в доказательстве теоремы 52, левая часть (5.53) равносильна условию (5.50), из которого (полагая $v_1 \stackrel{\text{def}}{=} \varepsilon$) получаем условие (5.46), которое, как было отмечено в доказательстве теоремы 51, равносильно правой части (5.53).

Теперь докажем импликацию (5.54).

Согласно теореме 50, из левой части (5.54) следует, что существует ЛА L размерности $k \leq n$ над X, такой, что $f_L = f$.

Пусть L имеет вид $(\xi^0, \{L^x \mid x \in X\}, \lambda)$, и u – произвольная строка из X^*, тогда

$$\begin{aligned} &\forall\, v_1, v_2 \in X^* \\ &\hat{f}_u(v_1, v_2) = f(v_1 u v_2) = \xi^0 L^{v_1 u v_2} \lambda = \xi^0 L^{v_1} L^u L^{v_2} \lambda. \end{aligned} \tag{5.55}$$

Обозначим записью E_{ij} (где $i, j = 1, \ldots, k$) квадратную матрицу порядка k, в которой компонента в строке i столбце j равна 1, а все остальные компоненты равны 0. Используя эти обозначения, можно представить матрицу L^u в виде суммы

$$L^u = \sum_{i,j=1,\ldots,k} a_{ij} E_{ij}, \tag{5.56}$$

где a_{ij} – компонента матрицы L^u в строке i столбце j.

Из (5.55) и (5.56) следует, что

$$\forall\, v_1, v_2 \in X^* \quad \hat{f}_u(v_1, v_2) = \sum_{i,j=1,\ldots,k} a_{ij} \xi^0 L^{v_1} E_{ij} L^{v_2} \lambda. \tag{5.57}$$

Обозначим записью f_{ij} (где $i, j = 1, \ldots, k$) функцию из $\mathbf{R}^{X^* \times X^*}$, определяемую следующим образом:

$$\forall\, v_1, v_2 \in X^* \quad f_{ij}(v_1, v_2) = \xi^0 L^{v_1} E_{ij} L^{v_2} \lambda. \tag{5.58}$$

Из (5.57) и (5.58) следует, что

$$\forall\, v_1, v_2 \in X^* \quad \hat{f}_u(v_1, v_2) = \sum_{i,j=1,\ldots,k} a_{ij} f_{ij}(v_1, v_2). \tag{5.59}$$

Из (5.59) следует, что функция $\hat{f}_u$ принадлежит подпространству линейного пространства $\mathbf{R}^{X^* \times X^*}$, порождённому функциями f_{ij} (где $i, j = 1, \ldots, k$). Это подпространство одинаково для всех $u \in X^*$. Поскольку размерность этого подпространства не превосходит k^2, и $k \leq n$, то, следовательно, верна правая часть импликации (5.54). ■

Можно доказать, что $\forall\, f \in \mathbf{R}^{X^*} \;\; \dim E_{\hat{f}} = (\dim E_f)^2$.

5.5 Счетномерные линейные автоматы и их языки

5.5.1 Вспомогательные понятия

В этом пункте мы будем использовать следующие обозначения.

- Символ $\mathbf{N}$ обозначает множество положительных натуральных чисел $(1, 2, \ldots)$.

- Запись $\mathbf{R}^{\mathbf{N}}$ обозначает множество последовательностей действительных чисел, $\forall f \in \mathbf{R}^{\mathbf{N}}$, $\forall i \in \mathbf{N}$ i–й член последовательности f обозначается записью f_i.

- Символ Ξ обозначает множество всех последовательностей $\xi \in \mathbf{R}^{\mathbf{N}}$, удовлетворяющих условию $\sum\limits_{i \geq 1} |\xi_i| < \infty$.

  Мы будем рассматривать элементы Ξ как счетномерные вектор-строки.

- Символ Λ обозначает множество всех последовательностей $\lambda \in \mathbf{R}^{\mathbf{N}}$, удовлетворяющих условию $\sup\limits_{i \geq 1} |\lambda_i| < \infty$.

  Мы будем рассматривать элементы Λ как счетномерные вектор-столбцы.

- Символ $\mathcal{H}$ обозначает множество всех бесконечных матриц H, компоненты которых индексированы парами натуральных чисел, и удовлетворяют условию $\sup\limits_{i \geq 1} \sum\limits_{j \geq 1} |H_{ij}| < \infty$.

  $\forall\, H, H' \in \mathcal{H}$ запись HH' обозначает матрицу из $\mathcal{H}$, компоненты которой определяются следующим образом:

  $$\forall\, i, j \in \mathbf{N} \quad (HH')_{ij} = \sum_{k \geq 1} H_{ik} H_{kj}.$$

  Нетрудно видеть, что

  – операция умножения матриц из $\mathcal{H}$ ассоциативна, и

– нейтральным элементом относительно этой операции является матрица E, определяемая следующим образом: $\forall\, i, j \in \mathbf{N}\ \ E_{ij} \stackrel{\mathrm{def}}{=} 1$, если $i = j$, и 0, иначе.

- $\forall\, \xi \in \Xi$, $\forall\, H \in \mathcal{H}$ запись ξH обозначает вектор-строку из Ξ, компоненты которой определяются следующим образом:

$$\forall\, i \in \mathbf{N} \quad (\xi H)_i = \sum_{k \geq 1} \xi_k H_{ki}.$$

- $\forall\, \lambda \in \Lambda$, $\forall\, H \in \mathcal{H}$ запись $H\lambda$ обозначает вектор-столбец из Λ, компоненты которого определяются следующим образом:

$$\forall\, i \in \mathbf{N} \quad (H\lambda)_i = \sum_{k \geq 1} H_{ik} \lambda_k.$$

5.5.2 Понятие счетномерного линейного автомата и связанные с ним понятия

Пусть задано конечное множество X.

Счетномерным линейным автоматом (СЛА) мы будем называть тройку L вида

$$L = (\xi^0, \{L^x \mid x \in X\}, \lambda) \tag{5.60}$$

где $\xi^0 \in \Xi$, $\forall\, x \in X\ \ L^x \in \mathcal{H}, \lambda \in \Lambda$.

Так же, как и для обычных ЛА, для каждого СЛА L можно определить понятие реакции $f_L \in \mathbf{R}^{X^*}$ и языка $L_a \subseteq X^*$, представимого СЛА с заданной точкой сечения $a \in \mathbf{R}$:

- $\forall\, u \in X^*\ \ f_L(u) \stackrel{\mathrm{def}}{=} \xi^0 L^u \lambda$, где

$$L^u \stackrel{\mathrm{def}}{=} \begin{cases} E, & \text{если } u = \varepsilon, \\ L^{x_1} \dots L^{x_n}, & \text{если } u = x_1 \dots x_n, \text{где } x_1, \dots, x_n \in X, \end{cases}$$

- язык L_a определяется соотношением (4.20).

Пусть $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$ – СЛА.

Базисные матрицы СЛА L – это бесконечные матрицы M_L и N_L, обладающие следующими свойствами:

- строки матрицы M_L образуют базис линейного подпространства пространства Ξ, порожденного строками вида $\xi^0 L^u$, где $u \in X^*$,

- столбцы матрицы N_L образуют базис линейного подпространства пространства Λ, порожденного столбцами вида $L^u\lambda$, где $u \in X^*$.

Базисные матрицы M_L и N_L определяют конгруэнции на линейных пространствах Ξ и Λ соответственно, обозначаемые символом $\sim_L$ и определяемые следующим образом:

$$\begin{array}{lll} \forall\, \xi, \xi' \in \Xi \quad \xi \sim_L \xi' & \Leftrightarrow & \xi N_L = \xi' N_L, \\ \forall\, \lambda, \lambda' \in \Lambda \quad \lambda \sim_L \lambda' & \Leftrightarrow & M_L\lambda = M_L\lambda'. \end{array}$$

Нетрудно доказать, что если матрица M_L (N_L) имеет конечное число строк (столбцов), то конгруэнция $\sim_L$ на Ξ (Λ) имеет конечный индекс.

5.5.3 Свойства счетномерных линейных автоматов

Теорема 54.

Пусть L – СЛА вида (5.60), $a \in \mathbf{R}$, и $E = \Xi$ или Λ.

Тогда L_a – ВЯ $\Leftrightarrow \exists$ линейная конгруэнция $\sim$ конечного ранга на E, обладающая следующими свойствами:

- $$\forall\, x \in X\ \exists\, L^x_\sim : \quad \begin{tikzcd} E \arrow[r, "L^x"] \arrow[d] & E \arrow[d] \\ E_\sim \arrow[r, dashed, "L^x_\sim"] & E_\sim \end{tikzcd} \tag{5.61}$$

  где $E_\sim$ – фактор-пространство, и $E \to E_\sim$ – каноническая проекция,

- $\exists$ линейное отображение $\lambda_\sim : E_\sim \to \mathbf{R}$, $\exists\ \xi^0_\sim \in E_\sim$, $\exists\, b \in \mathbf{R}$:

$$u \in L_a \quad \Leftrightarrow \quad \lambda_\sim(\xi^0_\sim A^u_\sim) > b. \tag{5.62}$$

Доказательство.

Обоснуем лишь импликацию “$\Rightarrow$”.
Пусть $\exists$ ВА B, $\exists\, b \in \mathbf{R} : L_a = B_b$.
Обозначим символом π эпиморфизм $E \to \mathbf{R}^{|S_B|}$.
Искомая конгруэнция $\sim$ определяется следующим образом:

- если $E = \Xi$, то $\xi' \sim \xi'' \;\Leftrightarrow\; \pi(\xi')B^u = \pi(\xi'')B^u$,
- если $E = \Lambda$, то $\lambda' \sim \lambda'' \;\Leftrightarrow\; B^u\pi(\lambda') = B^u\pi(\lambda'')$. $\blacksquare$

Отметим, что для каждого СЛА L

- конгруэнция $\sim_L$ (определённая в конце пункта 5.5.2) обладает свойством (5.61), и
- если конгруэнция $\sim_L$ имеет конечный индекс, то для языка $(f_L)_0$ выполнены условия теоремы 54.
  (напомним, что $\forall\, f \in \mathbf{R}^{X^*}\; f_0 \stackrel{\text{def}}{=} \{u \in X^* \mid f(u) > 0\}$)

Пусть G – некоторая полугруппа. Мы будем использовать следующие обозначения.

- Запись $\mathbf{R}^G$ обозначает множество функций вида $f : G \to \mathbf{R}$.
  $\mathbf{R}^G$ является линейным пространством, в котором операции сложения и умножения на действительные числа определяются стандартным образом.
- $\forall\, f \in \mathbf{R}^G$, $\forall\, u \in G$ запись f_u обозначает функцию из $\mathbf{R}^G$, которая сопоставляет каждому $u' \in G$ число $f(uu')$.
- $\forall\, f \in \mathbf{R}^G$ запись T_f обозначает множество $\{f_u \mid u \in G\}$.
- $\forall\, f \in \mathbf{R}^G$ запись $\langle T_f\rangle$ обозначает подпространство линейного пространства $\mathbf{R}^G$, порожденное функциями из T_f.

Теорема 55.
Пусть G – полугруппа, $f \in \mathbf{R}^G$, и $|Im(f)| < \infty$. Тогда

$$\dim\langle T_f\rangle < \infty \;\Leftrightarrow\; |T_f| < \infty.$$

Доказательство.
Импликация “$\Leftarrow$” является очевидной.
Докажем импликацию “$\Rightarrow$”.

$\langle T_f \rangle$ можно рассматривать как нормированное конечномерное линейное пространство, где $\|f\| = \max\limits_{u \in G} |f(u)|$.

T_f – ограниченное подмножество $\langle T_f \rangle$, поэтому T_f – компакт.

Из $|Im(f)| < \infty$ следует, что $\exists\, c > 0$:

$$\forall\, f_{u_1} \neq f_{u_2} \in T_f \quad \|f_{u_1} - f_{u_2}\| \geq c \tag{5.63}$$

Из компактности T_f следует, что множество открытых шаров с центрами в точках из T_f радиуса $\frac{c}{2}$ содержит конечное подмножество, покрывающее T_f. Согласно (5.63), отсюда следует, что $|T_f| < \infty$. ■

Теорема 56.

Пусть L – СЛА вида (5.60), $a \in \mathbf{R}$, и $E = \Xi$ или Λ.

Тогда язык L_a регулярен $\Leftrightarrow$ $\exists$ линейная конгруэнция $\sim$ конечного ранга на E, такая, что

- $\sim$ обладает свойствами, изложенными в теореме 54,
- $|\{\lambda_\sim(\xi^0_\sim A^u_\sim) \mid u \in X^*\}| < \infty$.

Доказательство.

Импликация "$\Rightarrow$" следует из теоремы 54, в данном случае B – КДА, и

$$\forall\, u \in X^* \quad \lambda_\sim(\xi^0_\sim A^u_\sim) \in \{0, 1\}.$$

Докажем импликацию "$\Leftarrow$".

Обозначим символом f функцию из $\mathbf{R}^{X^*}$, которая сопоставляет каждому $u \in X^*$ число $\lambda_\sim(\xi^0_\sim A^u_\sim)$. Поскольку f – ЛАФ, то $\dim\langle L_f \rangle < \infty$, поэтому по теореме 55 $|L_f| < \infty$.

Нетрудно видеть, что язык L_a совпадает с множеством строк, которые принимаются конечным детерминированным автоматом $A = (X, S, s^0, \delta, F)$, где

- $S \stackrel{\text{def}}{=} \{f_u \mid u \in X^*\}$,
- $s^0 \stackrel{\text{def}}{=} f_\varepsilon = f$,
- $\delta : S \times X \to S, \;\; \delta(f_u, x) \stackrel{\text{def}}{=} f_{ux}$,
- $F \stackrel{\text{def}}{=} \{f_u \mid f(u) > a\}$. ■

Теорема 57.

Пусть $f \in \mathbf{R}^{X^*}, |Im(f)| < \infty$.

Тогда f – ЛАФ $\Leftrightarrow$ $|\{f_u \mid u \in X^*\}| < \infty$. $\blacksquare$

Теорема 58.

Пусть $S \subseteq X^*$. Тогда χ_S (характеристическая функция подмножества S) – ЛАФ $\Leftrightarrow$ S регулярен. $\blacksquare$

Теорема 59.

Пусть $f \in \mathbf{R}^{X^*}$, $|Im(f)| < \infty$.

Тогда f – ЛАФ $\Leftrightarrow \forall\, a \in \mathbf{R}$ $\{u \in X^* \mid f(u) = a\}$ регулярен.

Доказательство.

Импликация "$\Rightarrow$" следует из теоремы 56.

Докажем импликацию "$\Leftarrow$".

Пусть $Im(f) = \{a_1, \ldots, a_m\}$, и $\forall\, i = 1, \ldots, m$ существует ДА A_i, который представляет язык $\{u \in X^* \mid f(u) = a_i\}$. Представим этот ДА в виде ЛА $(\xi_i^0, \{B_i^x \mid x \in X\}, \lambda_i)$, т.е.

- ξ_i^0 содержит единицу в позиции, соответствующей начальному состоянию A_i, остальные компоненты ξ_i^0 равны 0,
- $(B_i^x)_{s,s'} = 1$, если $s' = \delta_i(s, x)$, и 0 – иначе,
- компоненты λ_i, соответствующие терминальным состояниям A_i, равны 1, остальные компоненты λ_i равны 0.

Тогда $f(u) = (\xi_1^0 \ldots \xi_m^0) \begin{pmatrix} B_1^u & & \\ & \ldots & \\ & & B_m^u \end{pmatrix} \begin{pmatrix} a_1\lambda_1 \\ \ldots \\ a_m\lambda_m \end{pmatrix}$. $\blacksquare$

5.6 Достижимость и различимость в линейных автоматах

5.6.1 Достижимость

Пусть задан ЛА $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$, и $n = \dim L$.

Вектор $\xi \in \mathbf{R}^n$ называется **достижимым** в L, если

$$\exists\, u \in X^* :\ \xi = \xi^0 L^u.$$

Степенью достижимости ЛА L называется число

$$\delta(L) \stackrel{\mathrm{def}}{=} \min k : \langle \xi^0 L^u \mid u \in X^* \rangle = \langle \xi^0 L^u \mid u \in X^{\leq k} \rangle,$$

где $\forall\, V \subseteq \mathbf{R}^n$ запись $\langle V \rangle$ обозначает подпространство, порожденное векторами из множества V.

Обозначим записью M_L бесконечную матрицу, строками которой являются вектора вида $\xi^0 L^u$ ($u \in X^*$), и записью $r(M_L)$ – ранг этой матрицы. Нетрудно видеть, что $r(M_L) \leq n$.

Теорема 60.
Если L – ЛА конечной размерности, то $\delta(L) \leq r(M_L) - 1$.

Доказательство.
Цепочка подпространств

$$V_0 \subseteq V_1 \subseteq V_2 \subseteq \ldots \subseteq \mathbf{R}^n,$$

где $V_0 \stackrel{\mathrm{def}}{=} \langle \xi^0 \rangle$ и $\forall\, k \geq 0 \quad V_{k+1} \stackrel{\mathrm{def}}{=} V_k + \langle \{\xi L^x \mid \xi \in V_k, x \in X\} \rangle$, не может неограниченно возрастать. Нетрудно видеть, что

$$\forall\, k \geq 0 \quad V_k = \langle \{\xi^0 L^u \mid u \in X^{\leq k}\} \rangle,$$

и минимальное k, такое, что $V_k = V_{k+1}$, совпадает с $\delta(L)$, и удовлетворяет неравенству

$$k + 1 \leq \dim V_k = r(M_L),$$

откуда следует утверждение теоремы. ■

5.6.2 Различимость

Пусть задан ЛА $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$, и $n = \dim L$.

Векторы $\xi_1, \xi_2 \in \mathbf{R}^n$ называются **различимыми** в L, если

$$\exists\, u \in X^* : \ \xi_1 L^u \lambda \neq \xi_2 L^u \lambda. \qquad (5.64)$$

Ниже запись $\xi_1 \not\sim_L \xi_2$ обозначает соотношение (5.64).

Степенью различимости ЛА L называется число

$$\rho(L) \stackrel{\mathrm{def}}{=} \min k : \xi_1 \not\sim_L \xi_2 \Rightarrow \ \exists\, u \in X^{\leq k} : \xi_1 L^u \lambda \neq \xi_2 L^u \lambda.$$

Из определения $\rho(L)$ следует, что

$$\langle L^u\lambda \mid u \in X^*\rangle = \langle L^u\lambda \mid u \in X^{\leq \rho(L)}\rangle.$$

Обозначим записью N_L бесконечную матрицу, столбцами которой являются вектора вида $L^u\lambda$ ($u \in X^*$), и записью $r(N_L)$ – ранг этой матрицы. Нетрудно видеть, что $r(N_L) \leq n$.

Теорема 61.
Если L – ЛА конечной размерности, то $\rho(L) \leq r(N_L) - 1$.

Доказательство.
Теорема доказывается аналогично теореме 60: определяется цепочка вложенных подпространств $\{V_k \mid k \geq 0\}$, где

$$\forall\, i \geq 0 \quad V_k = \langle\{L^u\lambda \mid u \in X^{\leq k}\}\rangle,$$

и нетрудно доказать, что тот номер k, на котором эта цепочка стабилизируется (т.е. минимальное k, такое, что $V_k = V_{k+1}$), совпадает с $\rho(L)$. ■

5.7 Реализация функций на строках линейными автоматами

Пусть X – конечное множество.

Ганкелева матрица функции $f \in \mathbf{R}^{X^*}$ – это бесконечная матрица H^f, строки и столбцы которой индексированы элементами множества X^* и

$$\forall\, u, v \in X^* \quad H^f_{u,v} \stackrel{\text{def}}{=} f(uv),$$

где $H^f_{u,v}$ – элемент в строке u и столбце v матрицы H^f.

Для каждого $u \in X^*$ записи $\vec{H}^f_u$ и $H^{f\downarrow}_u$ обозначают строку u и столбец u соответственно матрицы H^f (данные понятия определяются так же, как аналогичные понятия из пункта 2.1.2).

Мы будем обозначать записью $r(f)$ ранг матрицы H^f, т.е. размерность линейного пространства, порожденного множеством её строк (или столбцов).

Подмножество B строк (или столбцов) матрицы H^f называется **базисным**, если порождаемое ими линейное пространство $\langle B\rangle$ совпадает с пространством, порождаемым всеми строками (или столбцами) H^f, и ни один элемент B не является линейной комбинацией других элементов B.

Теорема 62.

Если $f \in \mathbf{R}^{X^*}$, и $r(f) < \infty$, то в матрице H^f можно выбрать такое базисное множество B строк (или столбцов), что индекс u каждой строки (или столбца) из B удовлетворяет условию $|u| \le r(f) - 1$. ■

Будем использовать следующие обозначения. Пусть

$$r \stackrel{\text{def}}{=} r(f) < \infty,$$

и U и V – последовательности строк из X^* вида

$$U = (u_1, \ldots, u_r), \quad V = (v_1, \ldots, v_r) \tag{5.65}$$

соответственно, обладающие следующими свойствами:

- $u_1 = v_1 = \varepsilon$,
- $\{\vec{H}^f_u \mid u \in U\}$ и $\{H^{f\downarrow}_v \mid v \in V\}$ – базисные множества строк и столбцов матрицы H^f соответственно.

Тогда

- запись $(U, V)^f$ обозначает квадратную матрицу порядка r, строки и столбцы которой индексированы элементами U и V соответственно (мы будем называть эту матрицу **базисной матрицей** фунции f), и

  $$\forall\, u \in U,\ \forall\, v \in V \quad (U, V)^f_{u,v} \stackrel{\text{def}}{=} f(uv),$$

  т.е. $(U, V)^f$ – подматрица порядка r матрицы H^f (нетрудно доказать, что $(U, V)^f$ – невырожденная подматрица максимального порядка матрицы H^f),

- $\forall\, w \in X^*$ запись $(U,V)^{f,w}$ обозначает квадратную матрицу порядка r, строки и столбцы которой индексированы элементами U и V соответственно, и

$$\forall\, u \in U,\ \forall\, v \in V \quad (U,V)^{f,w}_{u,v} \stackrel{\text{def}}{=} f(uwv).$$

Теорема 63.

Пусть $f \in \mathbf{R}^{X^*}$, $r = r(f) < \infty$, и U, V – последовательности строк из X^*, такие, что $(U,V)^f$ базисная матрица функции f.

Тогда

$$(U,V)^{f,uv} = (U,V)^{f,u}\Big((U,V)^f\Big)^{-1}(U,V)^{f,v}.$$

Доказательство.

Нетрудно видеть, что

$$\begin{pmatrix} (U,V)^f & (U,V)^{f,v} \\ (U,V)^{f,u} & (U,V)^{f,uv} \end{pmatrix} =$$
$$= ((u_1,\ldots,u_r,u_1u,\ldots,u_ru),(v_1,\ldots,v_r,vv_1,\ldots,vv_r))^f.$$

Доказываемое равенство является следствием того, что ранг этой матрицы равен r. ■

Из теоремы 63 вытекает нижеследующая теорема.

Теорема 64.

Пусть $f \in \mathbf{R}^{X^*}$, $r = r(f) < \infty$, и U, V – последовательности строк из X^*, такие, что $(U,V)^f$ базисная матрица функции f.

Тогда $\forall\, u \in X^*$, если $u = x_1 \ldots x_s$, где $x_1, \ldots, x_s \in X$, то

$$(U,V)^{f,u} = (U,V)^{f,x_1}\Big((U,V)^f\Big)^{-1}\ldots\Big((U,V)^f\Big)^{-1}(U,V)^{f,x_s}. \ ■$$

Теорема 65.

Пусть $f \in \mathbf{R}^{X^*}$, $r = r(f) < \infty$, и U, V – последовательности строк из X^*, такие, что $(U,V)^f$ – базисная матрица функции f.

Тогда $f = f_L$, где L – ЛА, определяемый следующим образом:

$$L \stackrel{\text{def}}{=} \Big(\vec{e}_1,\ \Big\{(U,V)^{f,x}\Big((U,V)^f\Big)^{-1} \mid x \in X\Big\},\ (U,V)^f e_1^{\downarrow}\Big). \ ■$$

Литература

[1] *Rabin M. O.* Probabilistic automata // Information and Control, 1963. Vol. 6. No. 3. P. 230–245.

[2] *Хопкрофт Д., Мотвани Р., Ульман Дж.* Введение в теорию автоматов, языков и вычислений. – М.: Вильямс, 2002. 528 с.

[3] *Кемени Дж., Снелл Дж.* Конечные цепи Маркова. – М.: Наука, 1970. 274 с.

[4] *Carlyle J. W.* Reduced forms for stochastic sequential machines // Journal of Mathematical Analysis and Applications, 1963. Vol. 7. No. 2. P. 167–175.

[5] *Бухараев Р. Г.* Некоторые эквивалентности в теории вероятностных автоматов // Учёные записки Казанского университета, 1964. Т. 124. № 2. С. 45–65.

[6] *Starke P. H.* Theorie stochastischen Automaten // Elektronische Informationsverarbeitung und Kybernetik, 1965. Vol. 1. No. 2. P. 5–32.

[7] *Paz A.* Introduction to Probabilistic Automata. – New York: Academic Press, 1971. 228 p.

[8] *Бухараев Р. Г.* Основы теории вероятностных автоматов. – М.: Наука, 1985. 288 с.

[9] *Segala R., Lynch N. A.* Probabilistic simulations for probabilistic processes // Nordic Journal of Computing, 1995. Vol. 2. No. 2. P. 250–273.

[10] *Stoelinga M.* An introduction to probabilistic automata // Bulletin of the European Association for Theoretical Computer Science, 2002. Vol. 78. P. 176–198.

[11] *Sokolova A., de Vink E. P.* Probabilistic Automata: System Types, Parallel Composition and Comparison // Lecture Notes in Computer Science, 2004. Vol. 2925. P. 1–43.

[12] *Rabiner L. R.* A tutorial on Hidden Markov Models and selected applications in speech recognition // Proceedings of the IEEE, 1989. Vol. 77. No. 2. P. 257–286.

[13] *Darwiche A.* Modeling and Reasoning with Bayesian Networks. – Cambridge: Cambridge University Press, 2009. 562 p.

[14] *Koller D., Friedman N.* Probabilistic Graphical Models. Principles and Techniques. – Massachusetts: MIT Press, 2009. 1280 p.

[15] *Handbook of Markov Decision Processes* / Eds. E. A. Feinberg, A. Shwartz. – Boston: Kluwer, 2002. 562 p.

[16] *Wu S.-H., Smolka S. A., Stark E. W.* Composition and behaviors of probabilistic I/O automata // Theoretical Computer Science, 1997. Vol. 176. P. 1–38.

[17] *Delahaye B., Katoen J.-P., Larsen K.G., Legay A., Pedersen M. L., Sher F., Wasowski A.* Abstract Probabilistic Automata // Information and Computation, 2013. Vol. 232. P. 66–116.

[18] *Kudlek M.* Probability in Petri nets // Fundamenta Informaticae, 2005. Vol. 67. No. 1. P. 121–130.

[19] *Liu Y., Miao H., Zeng H., Li Z.* Probabilistic Petri net and its logical semantics. // Proceedings of Ninth International Conference on Software Engineering Research, Management and Applications. – Baltimore: IEEE Computer Society, 2011. P. 73–78.

[20] *Eisentraut C., Hermanns H., Zhang L.* On probabilistic automata in continuous time // Proceedings of the 25th Annual

IEEE Symposium on Logic in Computer Science (LICS), 2010. P. 342–351.

[21] *Jonsson B., Larsen K. G., Yi W.* Probabilistic extensions of process algebras // Handbook of Process Algebras. – North Holland: Elsevier, 2001. P. 685-710.

[22] *Бухараев Р. Г.* Теория абстрактных вероятностных автоматов // Проблемы кибернетики, 1975. Вып. 30. С. 147–198.

[23] *Homuth H. H.* A type of stochastic automation applicable to the communication channel // Angewandte Informatik, 1971. No. 8. P. 362–372.

[24] *Бухараев Р. Г.* Сети вероятностных процессоров // Математические вопросы кибернетики, 2007. Вып. 16. С. 57–72.

[25] *Мур Э. Ф.* Умозрительные эксперименты с последовательностными машинами. // Автоматы. – М.: Иностранная литература, 1956. С. 179–210.

[26] *Миронов А. М., Френкель С. Л.* Минимизация вероятностных моделей программ // Фундаментальная и прикладная математика, 2014. Т. 19. Вып. 1. С. 121–163.

[27] *Kiefer S., Wachter B.* Stability and Complexity of Minimising Probabilistic Automata // Lecture Notes in Computer Science, 2014. Vol. 8573. P. 268–279.

[28] *Mateus P., Qiu D., Li L.* On the complexity of minimizing probabilistic and quantum automata // Information and Computation, 2012. Vol. 218. P. 36–53.