Основные понятия теории вероятностных автоматов (часть 2)

А. М. Миронов

Настоящая работа является продолжением статьи [1] и использует понятия и обозначения, введённые в [1]. В работе излагаются основные понятия теории вероятностных автоматов Мура с числовым выходом и теории вероятностных языков. Приводятся новые доказательства классических результатов теории вероятностных автоматов, связанных с эквивалентностью и редукцией вероятностных автоматов Мура с числовым выходом, а также с регулярностью вероятностных языков.

Ключевые слова: вероятностные автоматы, вероятностные языки.

Вероятностные автоматы Мура с числовым выходом

Вероятностные автоматы Мили и Мура

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

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

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 $\lambda : S \underset{r}{\to} Y$).

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

$$\begin{pmatrix} \delta(s_1, x, s_1) & \ldots & \delta(s_1, x, s_n) \\ \ldots & \ldots & \ldots \\ \delta(s_n, x, s_1) & \ldots & \delta(s_n, x, s_n) \end{pmatrix}, \tag{1}$$

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

$$\begin{array}{ll} \forall\, s, s' \in S & A^x_{s,s'} \geq 0, \\ \forall\, s \in S & \sum\limits_{s' \in S} A^x_{s,s'} = 1 \ \ (\text{т.е. } A^x I = I). \end{array} \tag{2}$$

(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 $u = x_1 \ldots x_k$, то $A^u \stackrel{\text{def}}{=} A^{x_1} \ldots A^{x_k}$.

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 $x_0 \ldots x_k$, 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

$$(\xi^0, \{A^x \mid x \in X\}, \lambda), \tag{3}$$

компоненты которой определяются следующим образом.

- Компонента ξ^0 в (3) является вектор-строкой, соответствующей начальному распределению ξ^0 ВА A.
- Компонента $\{A^x \mid x \in X\}$ в (3) является совокупностью матриц перехода ВА A.

- Компонента λ в (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.

Вероятностные автоматы Мура с числовым выходом

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 3.3 статьи [1]. Мы будем называть это понятие **усреднённой реакцией**.

Усреднённые реакции

Пусть A – ВА Мура с числовым выходом вида (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:

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

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 $x_0 \dots x_k$, 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.

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

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 3.4 статьи [1].

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 3.4 статьи [1].

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

$$[\![A]\!]_s = \sum_{s' \in S_A \setminus \{s\}} (s')^\xi [\![A]\!]_{s'}. \tag{4}$$

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 параграфе 4 статьи [1]. Мы рассмотрим лишь метод редукции путем удаления выпуклых комбинаций. Данный метод основан на нижеследующей теореме.

Теорема 1.

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 (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

$$\begin{aligned} &\xi_B^0 = \xi_A^0 \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix}, \\ &B^x = (E_{n-1} \ \ \mathbf{0}) A^x \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix}, \\ &\lambda_B = (E_{n-1} \ \ \mathbf{0}) \lambda_A, \end{aligned} \tag{5}$$

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 (4) в виде

$$[\![A]\!]_{s_n} = \sum_{i=1}^{n-1} s_i^\xi [\![A]\!]_{s_i}. \tag{6}$$

Согласно определению матрицы $[\![A]\!]$, равенство (6) равносильно соотношению

$$\forall\, u \in X^* \quad A_n^u \lambda_A = \sum_{i=1}^{n-1} s_i^\xi A_i^u \lambda_A, \tag{7}$$

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

Можно переписать (7) в матричном виде:

$$\forall\, u \in X^* \quad \vec{e}_n A^u \lambda_A = \xi (E_{n-1} \ \ \mathbf{0}) A^u \lambda_A \tag{8}$$

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

В частности, (8) верно при $u = \varepsilon$, т.е.

$$\vec{e}_n \lambda_A = \xi (E_{n-1} \ \ \mathbf{0}) \lambda_A. \tag{9}$$

Докажем, что $A \approx B$, т.е. $\forall\, u \in X^*$ $\xi_A^0 A^u \lambda_A = \xi_B^0 B^u \lambda_B$. Согласно (5), для этого достаточно доказать, что $\forall\, u \in X^*$

$$A^u \lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} B^u (E_{n-1} \ \ \mathbf{0}) \lambda_A. \tag{10}$$

Докажем (10) индукцией по $|u|$.

Если $u = \varepsilon$, то (10) имеет вид

$$\lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} (E_{n-1} \ \ \mathbf{0})\lambda_A. \tag{11}$$

Согласно правилам матричного умножения, и учитывая (9), можно переписать правую часть (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 (10) верно.

Пусть (10) верно для некоторого u. Докажем, что $\forall\, x \in X$

$$A^{xu}\lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} B^{xu}(E_{n-1} \ \ \mathbf{0})\lambda_A. \tag{12}$$

Используя определение B^x из (5), перепишем (12) в виде

$$A^{xu}\lambda_A = \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} (E_{n-1} \ \ \mathbf{0})A^x \begin{pmatrix} E_{n-1} \\ \xi \end{pmatrix} B^u(E_{n-1} \ \ \mathbf{0})\lambda_A. \tag{13}$$

Учитывая индуктивное предположение (10), и используя правила матричного умножения перепишем (13) в виде

$$\begin{aligned} A^{xu}\lambda_A &= \begin{pmatrix} E_{n-1}(E_{n-1} \ \ \mathbf{0}) \\ \xi(E_{n-1} \ \ \mathbf{0}) \end{pmatrix} A^x A^u \lambda_A = \\ &= \begin{pmatrix} (E_{n-1} \ \ \mathbf{0}) \\ \xi(E_{n-1} \ \ \mathbf{0}) \end{pmatrix} A^{xu}\lambda_A = \begin{pmatrix} (E_{n-1} \ \ \mathbf{0})A^{xu}\lambda_A \\ \xi(E_{n-1} \ \ \mathbf{0})A^{xu}\lambda_A \end{pmatrix}. \end{aligned} \tag{14}$$

Используя (8), перепишем правую часть (14) в виде

$$\begin{pmatrix} (E_{n-1} \ \ \mathbf{0})A^{xu}\lambda_A \\ \vec{e}_n A^{xu}\lambda_A \end{pmatrix} = \begin{pmatrix} (E_{n-1} \ \ \mathbf{0}) \\ \vec{e}_n \end{pmatrix} A^{xu}\lambda_A = A^{xu}\lambda_A. \tag{15}$$

Таким образом, (12) верно.

Следовательно, (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 4.3 статьи [1].

Соглашение

Начиная со следующего пункта, 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), 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 параграфе 3 статьи [1], мы будем называть **ВА общего вида**. Для каждого рассматриваемого ВА A запись S_A обозначает множество состояний этого ВА.

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
$$\forall\, u \in X^* \quad (fD^x)(u) \stackrel{\text{def}}{=} f(xu). \tag{16}$$

- 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 2.

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 (16),

$$\forall\, u \in X^* \quad (f_{A_i} D^x)(u) = f_{A_i}(xu) = \xi_i A^{xu} \lambda = \xi_i A^x A^u \lambda. \qquad (17)$$

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

$$A \stackrel{\text{def}}{=} (\xi^0, \{A^x \mid x \in X\}, \lambda), \tag{18}$$

где

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

$$f = \sum_{i=1}^{n} i^{\xi^0} f_i, \tag{19}$$

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

$$f_i D^x = \sum_{j=1}^{n} A^x_{ij} f_j, \tag{20}$$

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

Докажем, что реакция ВА (18) совпадает с f, т.е.

$$\forall\, u \in X^* \quad \xi^0 A^u \lambda = f(u). \tag{21}$$

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

$$A^u \lambda = \begin{pmatrix} f_1(u) \\ \ldots \\ f_n(u) \end{pmatrix}. \tag{22}$$

Если $u = \varepsilon$, то обе части (22) совпадают по определению λ.

Если $u = xu'$, то, предполагая верным равенство (22), в котором u заменено на u', имеем:

$$A^u\lambda = A^{xu'}\lambda = A^x A^{u'}\lambda = A^x \begin{pmatrix} f_1(u') \\ \dots \\ f_n(u') \end{pmatrix} =$$
$$= \begin{pmatrix} \sum\limits_{i=1}^{n} A^x_{1j} f_j(u') \\ \dots \\ \sum\limits_{i=1}^{n} A^x_{nj} f_j(u') \end{pmatrix}. \tag{23}$$

Из (20) следует, что правую часть в (23) можно переписать в виде
$$\begin{pmatrix} (f_1 D^x)(u') \\ \dots \\ (f_n D^x)(u') \end{pmatrix}. \tag{24}$$
Согласно определению (16) функций вида fD^x, столбец (24) совпадает с правой частью доказываемого равенства (22).

Таким образом, равенство (22) доказано. Согласно этому равенству, левая часть доказываемого равенства (21) равна

$$\xi^0 \begin{pmatrix} f_1(u) \\ \dots \\ f_n(u) \end{pmatrix} = \sum_{i=1}^{n} i^{\xi^0} f_i(u). \tag{25}$$

По определению ξ^0 (см. (19)), правая часть (25) равна $f(u)$, т.е. правой части доказываемого равенства (21). ■

Связь между линейно-автоматными функциями и реакциями вероятностных автоматов

Теорема 3.
Пусть $L = (\xi^0_L, \{L^x \mid x \in X\}, \lambda_L)$ – ЛА размерности n.
Тогда существуют ВА A и число $a > 0$, такие, что $\forall\, u \in X^*$

$$f_A(u) = a^{|u|+1} f_L(u) + \tfrac{1}{n+2}. \tag{26}$$

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

$$A_1^x I = 0^{\downarrow}, \quad \tilde{I} A_1^x = \vec{0}, \tag{27}$$

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 (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:

$$B \stackrel{\mathrm{def}}{=} \tfrac{1}{n+2}\begin{pmatrix} 1 & \ldots & 1 \\ \ldots & \ldots & \ldots \\ 1 & \ldots & 1 \end{pmatrix}, \quad \xi \stackrel{\mathrm{def}}{=} \tfrac{1}{n+2}(1 \ldots 1).$$

Нетрудно видеть, что

$$\begin{array}{l} B^2 = B,\ BI = I,\ \xi I = 1,\ \xi B = \xi, \\ A_1^x B = 0,\ BA_1^x = 0,\ \xi_1^0 B = 0,\ \xi A_1^x = 0, \end{array} \tag{28}$$

где символ 0 в (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

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

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

Докажем, что верно равенство (26).

Если $u = \varepsilon$, то левая часть (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 (26).

Если $u \neq \varepsilon$, то из (28) следует, что $A^u = a^{|u|}A_1^u + B$, и, используя (28), получаем:

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

Эргодичные автоматы

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 $u = x_1 \ldots x_k$, то $A^u \stackrel{\text{def}}{=} A^{x_1} \ldots A^{x_k}$, и

– $\forall\, u \in X^*, \forall\, i, j \in \{1, \ldots, n\}$ 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 4.

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

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

$$\forall\, i, i' \in \{1, \ldots, n\} \;\exists\, j : A^u_{ij} > 0, A^u_{i'j} > 0. \tag{29}$$

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 (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

$$\|A^u B\| \leq (1-c)\cdot\|B\|. \tag{30}$$

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

$$|d_{ij} - d_{i'j}| \leq (1-c)(M_j - m_j), \tag{31}$$

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

$$d_{ij} - d_{i'j} \leq M_j - A^u_{ij_0}(M_j - b_{j_0 j}) - m_j - A^u_{i'j_0}(b_{j_0 j} - m_j). \tag{32}$$

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 (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

$$d_{ij} - d_{i'j} \leq (1-c)(M_j - m_j). \tag{33}$$

В силу произвольности выбора индексов i, i' верно неравенство

$$d_{i'j} - d_{ij} \leq (1-c)(M_j - m_j). \tag{34}$$

Из (33) и (34) следует (31). ■

Из леммы следует, что

$$\forall\, u \in X^* \ \|A^u\| \leq (1-c)^{[|u|/k]} \cdot \|A^l\| \quad (|l| < k). \tag{35}$$

Правая часть (35) стремится к нулю при $|u| \to \infty$, поэтому условие (a) верно. ■

Понятие эргодичного вероятностного автомата и критерий эргодичности

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

$$\forall\, s, s' \in S_A \quad |\vec{A}^u_s - \vec{A}^u_{s'}| \to 0 \quad (|u| \to \infty). \tag{36}$$

Теорема 5.
ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ эргодичен тогда и только тогда, когда $\forall\, u \in X^* \backslash \{\varepsilon\}$ матрица A^u **регулярна** (т.е. $\exists\, n: \ (A^u)^n > 0$).

Доказательство.
Если A эргодичен, но для некоторого $u \in X^* \backslash \{\varepsilon\}$ матрица A^u нерегулярна, то $\forall\, k \geq 1$ $(A^u)^k$ нерегулярна. Из эргодичности A следует, что выполнено условие (a) в теореме 4 ($\lim\limits_{|u| \to \infty} \|A^u\| = 0$). 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

$$P \in \mathcal{Q}_A^I \;\Rightarrow\; \forall\, Q \in \mathcal{Q}_A \;\; QP \in \mathcal{Q}_A^I. \tag{37}$$

Докажем, что

$$\forall\, u \in X^{k+1} \;\; Q(A^u) \in \mathcal{Q}_A^I. \tag{38}$$

Пусть $u = (x_1 \ldots x_{k+1})$ и $Q(A^u) \not\in \mathcal{Q}_A^I$. Из (37) следует, что

$$\forall\, i = 1, \ldots, k+1 \quad Q(A^{x_1}) \cdot \ldots \cdot Q(A^{x_{k+1}}) \in \mathcal{Q}_A \setminus \mathcal{Q}_A^I. \tag{39}$$

Поскольку $|\mathcal{Q}_A \setminus \mathcal{Q}_A^I| = k$, то из (39) следует, что

$$\begin{array}{l} \exists\, i, j \in \{1, \ldots, k+1\} : i < j, \\ Q(A^{x_i}) \cdot \ldots \cdot Q(A^{x_{k+1}}) = Q(A^{x_j}) \cdot \ldots \cdot Q(A^{x_{k+1}}) \not\in \mathcal{Q}_A^I. \end{array} \tag{40}$$

Определим $v \stackrel{\text{def}}{=} (x_i \ldots x_{j-1})$, $w \stackrel{\text{def}}{=} (x_j \ldots x_{k+1})$. Из (40) следует, что

$$\forall\, t \geq 1 \quad Q((A^v)^t A^w) = Q(A^w) \not\in \mathcal{Q}_A^I. \tag{41}$$

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 (41).

Таким образом, свойство (38) верно. Это свойство совпадает с условием (b) теоремы 4 для ВА A. Следовательно, верно условие (a) этой теоремы, откуда следует эргодичность A. ■

Устойчивость вероятностных автоматов

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 6.

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,

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

Теорема 7.

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

$$\|A^u\| \leq (1 - 2c)^{|u|-1}. \tag{42}$$

Доказательство.

Обозначим символом n порядок матриц A^x.

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

$$\forall\, \lambda \in \mathbf{R}^n \quad \|A^u\lambda\| \leq (1 - 2c)^{|u|}\|\lambda\|. \tag{43}$$

(43) верно для $u = \varepsilon$.

Если (43) верно для некоторого $u \in X^*$, то $\forall\, x \in X$, используя предположение (43) и теорему 6, получаем:

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

Таким образом, (43) верно для всех $u \in X^*$.

Теперь докажем (42) индукцией по $|u|$.

- Если $|u| = 0$ или 1, то (42) верно.

- Пусть (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 (43) и свойство $\forall\, i = 1, \ldots, n$ $\|\lambda_i\| \leq 1$, получаем:

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

Таким образом, (42) верно для любой строки $u \in X^*$. ∎

Теорема 8.

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 $(B_1^{\downarrow} \ldots B_n^{\downarrow})$, тогда

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

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 $(i, j = 1, \ldots, n)$. Поскольку $\forall\, i = 1, \ldots, n \;\; \sum\limits_j a_{ij} = 1$, 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 9.

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

$$|PB - QB| \leq \|B\|. \tag{44}$$

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:

$$\forall\, i = 1, \ldots, n \quad \sum_{j=1}^{n} a_{ij} = 0, \quad \sum_{j=1}^{n} a_{ij}^{(+)} \leq 1, \tag{45}$$

где $a_{ij}^{(+)} \stackrel{\text{def}}{=} \begin{cases} a_{ij}, & \text{если } a_{ij} \geq 0 \\ 0, & \text{иначе} \end{cases}$, 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 (44) в виде

$$|AB| \leq \|B\|. \tag{46}$$

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

$$|c_{kr}| = \max_{i,j} |c_{ij}| \quad (= |AB|\,). \tag{47}$$

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

$$M - m = \|B_r^{\downarrow}\| \leq \|B\|. \tag{48}$$

Из (47) и (48) следует, что для доказательства неравенства (46) достаточно доказать неравенство

$$|c_{kr}| \leq M - m. \tag{49}$$

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

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

Для доказательства неравенства (49) рассмотрим отдельно случаи $c_{kr} \geq 0$ и $c_{kr} < 0$.

- Если $c_{kr} \geq 0$, то
$$\begin{aligned} &|c_{kr}| = c_{kr} = \sum_j a_{kj} b_{jr} \leq a^{(+)} M + a^{(-)} m = \\ &= a^{(+)}(M - m) \leq M - m, \end{aligned}$$
т.е. в данном случае неравенство (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 (49) также верно. ■

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 10.
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, $\exists\,\delta > 0$: $\forall\,u \in X^*$

$$|f_A(u) - a| > \delta. \tag{50}$$

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

$$|f_A(u) - f_B(u)| \leq \tfrac{\delta}{2}. \tag{51}$$

Отметим, что из (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

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

  согласно (51) и (50) следуют соотношения

  $$\begin{aligned}&|f_B(u) - a| \geq |f_A(u) - a| - |f_A(u) - f_B(u)| > \\ &> \delta - \tfrac{\delta}{2} = \tfrac{\delta}{2},\end{aligned} \tag{52}$$

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

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

  неверно, то, согласно (50) и (52), будет верно неравенство

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

  которое противоречит (51).

Пусть B имеет вид $(\xi^0_B, \{B^x \mid x \in X\}, \lambda)$, тогда можно переписать (51) в виде

$$|\xi^0_A A^u \lambda - \xi^0_B B^u \lambda| \leq \tfrac{\delta}{2}. \tag{53}$$

Поскольку

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

(где $n = |S_A|$), то, следовательно, неравенство (53) будет верно, если будут верны неравенства

$$|A^u - B^u| \leq \delta_1, \quad |\xi_A^0 - \xi_B^0| \leq \delta_1, \quad \text{где } \delta_1 = \tfrac{\delta}{4n|\lambda|}. \tag{54}$$

Таким образом, для доказательства теоремы (10) достаточно доказать, что $\exists\, \varepsilon \in (0, \delta_1)$: $\forall\, B \in O_\varepsilon(A)$, $\forall\, u \in X^*$

$$|A^u - B^u| < \delta_1. \tag{55}$$

По предположению, $\exists\, c > 0$: $\forall\, x \in X \;\; A^x \geq c$.
Выберем k так, чтобы было верно неравенство

$$(1 - c)^{k-1} < \tfrac{\delta_1}{3}. \tag{56}$$

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

$$\forall\, u \in X^{\leq k} \quad |A^u - B^u| < \tfrac{\delta_1}{3}. \tag{57}$$

Тогда $\forall\, u \in X^*$ верно неравенство (55).

Доказательство.
Если $|u| \leq k$, то (55) следует из (57).
Пусть $|u| > k$, тогда $u = u_1 u_2$, где $|u_2| = k$, и

$$\begin{array}{l} |A^u - B^u| = |A^{u_1 u_2} - B^{u_1 u_2}| \leq \\ \leq |A^{u_1} A^{u_2} - A^{u_2}| + |B^{u_1} B^{u_2} - B^{u_2}| + |A^{u_2} - B^{u_2}|. \end{array} \tag{58}$$

Согласно теореме (8), 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 (58) и из истинности $\forall\, u \in X^k$ неравенства в (57) следует неравенство

$$|A^u - B^u| < \|A^{u_2}\| + \|B^{u_2}\| + \tfrac{\delta_1}{3}. \tag{59}$$

Согласно теореме 7 и условию (56),

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

$$\begin{array}{l} \forall\, u \in X^k \\ \|A^u\| \leq (1-2c)^{k-1} < (1-c)^{k-1} < \frac{\delta_1}{3}, \end{array} \tag{60}$$

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

$$\begin{array}{l} \forall\, u \in X^k \\ \|B^u\| \leq (1 - 2 \cdot \frac{c}{2})^{k-1} = (1-c)^{k-1} < \frac{\delta_1}{3}. \end{array} \tag{61}$$

Из (59), (60), и (61) следует, что

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

Таким образом, для доказательства теоремы 10 осталось доказать, что $\exists\, \varepsilon \in (0, \min\{\delta_1, \frac{c}{2}\}) : \forall\, B \in O_\varepsilon(A)$ верно (57).

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

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

(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 $M^u \stackrel{\text{def}}{=} M^{x_1} \ldots M^{x_s}$, 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 10.

Теорема 11.
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 10, для доказательства теоремы 11 достаточно доказать, что $\exists\, \varepsilon \in (0, \delta_1)$, где $\delta_1 \stackrel{\text{def}}{=} \frac{\delta}{4n|\lambda|}$:

$$\forall\, B \in O_\varepsilon(A), \;\; \forall\, u \in X^* \quad |A^u - B^u| < \delta_1. \tag{62}$$

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 10, 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 (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 $\exists\, \delta < 1 : \forall\, x \in X \;\; \|A^x\| < \delta$.

Вероятностные языки

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 $x_1 \dots x_n \in X^*$, такой, что $\xi^0 A^{x_1} \dots A^{x_n} \lambda > a$, т.е. строке из ВЯ A_a, где $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$.

Свойства вероятностных языков

Теорема 12.
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 13.

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 $\begin{pmatrix} \lambda_j A^x_{ij} & (1 - \lambda_j) A^x_{ij} \\ \lambda_j A^x_{ij} & (1 - \lambda_j) A^x_{ij} \end{pmatrix}$,
- $\lambda_B = (1\,0\, \dots\, 1\,0)^\sim$.

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 14.

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

$$A_a = B_b. \tag{63}$$

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 (63).

- 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 (63). ■

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 15.

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. ■

Регулярность вероятностных языков

Понятие регулярного языка

Регулярный язык над конечным множеством X – это подмножество $U \subseteq X^*$, такое, что существует автомат Мура M вида

$$M = (X, Y, S, \delta, \lambda, s^0) \tag{64}$$

удовлетворяющий условиям:

$$|S| < \infty, \;\; Y = \{0,1\}, \;\; U = \{u \in X^* \mid f_M(u) = 1\}. \tag{65}$$

Весом регулярного языка U называется наименьшее число n, такое, что $\exists$ автомат Мура M вида (64), такой, что $|S| = n$ и выполнены условия (65). 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: $\forall\, u, u' \in X^*$

  $$u \underset{U}{\sim} u' \;\Leftrightarrow\; \forall\, v \in X^* \,(uv \in U \Leftrightarrow u'v \in U) \tag{66}$$

  является конечным,

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

Теорема 16.

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 $U = \tilde{M}_0$. ■

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 $\exists\, \delta > 0$:

$$\forall\, u \in X^* \quad |f_A(u) - a| > \delta. \tag{67}$$

Для каждого ВА A и каждого $\delta > 0$ запись $I_\delta(A)$ обозначает совокупность всех ИТС a для A, обладающих свойством (67). 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 17.

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

$$w(A_a) \leq (1 + \tfrac{1}{\delta})^{n-1}. \tag{68}$$

Доказательство.

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

$$uv \in A_a \not\Leftrightarrow u'v \in A_a. \tag{69}$$

Т.к. $a \in I_\delta(A)$, то из (69) и (67) следует неравенство

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

т.е.

$$|\xi^0(A^u - A^{u'})A^v\lambda| > 2\delta. \tag{70}$$

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

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 $\vec{x}+M\overset{\text{def}}{=}\{(x_1+y_1,\ldots,x_n+y_n)\mid(y_1,\ldots,y_n)\in M\}$,
- если $Vol_{n-1}(M)$ определено, то $\forall\,d>0$
$$Vol_{n-1}(dM)=d^{n-1}Vol_n(M),$$
где $dM\overset{\text{def}}{=}\{(dx_1,\ldots,dx_n)\mid(x_1,\ldots,x_n)\in M\}$,
- 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
$$Vol_{n-1}(M)\le Vol_{n-1}(M').$$

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 $\xi^0A^u+\Delta_d$ соответственно. Из сказанного выше следует, что

$$\forall\,d>0\quad Vol_{n-1}(\Delta_d)=Vol_{n-1}(\Delta_d^u)=d^{n-1}C. \tag{71}$$

Нетрудно видеть, что

(A) $\forall\,u\in R\quad \Delta_\delta^u\subseteq\Delta_{1+\delta}$, т.к. если $\vec{x}\in\Delta_\delta^u=\xi^0A^u+\Delta_\delta$, то
$$\vec{x}=\xi^0A^u+\delta\vec{y},\ \text{где}\ \vec{y}\in\Delta_1.$$

Все компоненты $\vec{y}$ неотрицательны и $\vec{y}I=1$, поэтому все компоненты $\vec{x}$ неотрицательны и
$$\vec{x}I=(\xi^0A^u+\delta\vec{y})I=\xi^0A^uI+\delta\vec{y}I=1+\delta,$$
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 $\exists\, \vec{y}, \vec{z} \in \Delta_1$:

$$\vec{x} = \xi^0 A^u + \delta\vec{y}, \quad \vec{x} = \xi^0 A^{u'} + \delta\vec{z}. \tag{72}$$

Из (72) следует, что $\xi^0(A^u - A^{u'}) = \delta(\vec{z} - \vec{y})$, откуда на основании (70) получаем: $|\delta(\vec{z} - \vec{y})\lambda'| > 2\delta$, или

$$|(\vec{z} - \vec{y})\lambda'| > 2, \text{ где } \lambda' \in [0,1]^n. \tag{73}$$

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

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

что противоречит соотношению (73).

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 (71), 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 (68). ■

Следующая теорема показывает, что в ВА ограниченного размера можно представлять сколь угодно сложные регулярные языки.

Теорема 18.
Пусть $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ – ВА, где

$$|S_A| = 2,\ X = \{0, 2\},\ \xi^0 = (1\ \ 0),$$
$$A^0 = \begin{pmatrix} 1 & 0 \\ 2/3 & 1/3 \end{pmatrix},\ A^2 = \begin{pmatrix} 1/3 & 2/3 \\ 0 & 1 \end{pmatrix},\ \lambda = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

и $\forall\, n \geq 1\ \ a_n$ – число из $(0,1)$, имеющее в троичной записи вид $0.2\ldots 211$ (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 17.

Теорема 19.

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:

$$\forall\, x \in X \quad \rho(s_1 x, s_2 x) \leq \rho(s_1, s_2), \tag{74}$$

$$\exists\, \delta > 0 :\ \lambda(s_1) \neq \lambda(s_2)\ \Rightarrow\ \rho(s_1, s_2) \geq \delta. \tag{75}$$

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 (75) и (74) следует, что $\forall\, u, u' \in R$: $u \neq u'\ \exists\, v \in X^*$:

$$\delta \leq \rho(s^0 uv, s^0 u' v) \leq \rho(s^0 u, s^0 u'). \tag{76}$$

Пусть $\mathcal{C}$ – минимальное (по числу элементов) конечное покрытие S открытыми шарами радиуса $\delta/2$. Если u, u' – различные элементы R, то из (76) следует, что $s^0 u$ и $s^0 u'$ не могут попасть в один и тот же шар из $\mathcal{C}$, поэтому $|R| \leq |\mathcal{C}|$. ■

Теорема 17 является следствием теоремы 19, т.к. если заданы ВА $A = (\xi^0, \{A^x \mid x \in X\}, \lambda)$ и число $a \in [0, 1)$, удовлетворяющие условиям теоремы 17, 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

$$\rho(\vec{x}, \vec{y}\,) = \max_{i=1\ldots n} |x_i - y_i|, \tag{77}$$

т.к. оно является ограниченным подмножеством $\mathbf{R}^n$. Нетрудно доказать, что для любой пары s_1, s_2 достижимых состояний автомата M условия (74) и (75) выполнены.

Следующая теорема также является следствием теоремы 19.

Теорема 20.

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 (74) и (75), где метрика ρ определяется соотношением (77).

Тогда язык $U \stackrel{\text{def}}{=} \{u \in X^* \mid f_M(u) = 1\}$ регулярен, и

$$w(U) \leq C_{n+m-1}^m, \quad \text{где } m =]\tfrac{2}{\delta}[. \tag{78}$$

Доказательство.

Регулярность U непосредственно следует из теоремы 19.

Для доказательства неравенства (78) определим покрытие множества $\{1,\ldots,n\}^\triangle$ открытыми шарами радиуса $\frac{1}{m}$ (поскольку $\frac{1}{m} \leq \frac{\delta}{2}$, то, согласно теореме 19, $w(U)$ 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 $|F^n_m| = C^m_{n+m-1}$. ■

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 21.

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 $\exists\, \delta > 0$:

$$\forall\, u \in X^* \quad |f_A(u) - a| > \delta. \tag{79}$$

Выберем k: $(1-2c)^{k-1} < \frac{2\delta}{n|\lambda|}$, где $n = |S_A|$.

По теореме 7,

$$\forall\, u \in X^k \ \ \|A^u\| \leq (1-2c)^{k-1} < \tfrac{2\delta}{n|\lambda|}. \tag{80}$$

$\forall\, u \in X^k$, $\forall\, v \in X^*$

$$|f_A(vu) - f_A(u)| = |\xi^0 A^{vu}\lambda - \xi^0 A^u \lambda| = |\xi^0(A^{vu} - A^u)\lambda| \leq$$
$$\leq |A^{vu} - A^u| \cdot n \cdot |\lambda| = |A^v A^u - A^u| \cdot n \cdot |\lambda| \leq \|A^u\| \cdot n \cdot |\lambda| < 2\delta.$$

(предпоследнее неравенство верно по теореме 8, а последнее – согласно (80)). Таким образом,

$$\forall\, u \in X^k, \forall\, v \in X^* \quad |f_A(vu) - f_A(u)| < 2\delta. \tag{81}$$

Докажем, что $(A_a)_{\geq k} = X^* \cdot (A_a)_k$, т.е. $\forall\, u \in X^k$, $\forall\, v \in X^*$

$$vu \in A_a \quad \Leftrightarrow \quad u \in A_a. \tag{82}$$

Если для некоторых $u \in X^k, v \in X^*$ соотношение (82) неверно, то, согласно определению ВЯ A_a и соотношению (79),

$$|f_A(vu) - f_A(u)| > 2\delta$$

что противоречит соотношению (81). ■

Теорема 22.

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 21, идущей после соотношения (80). ■

Языки, представимые линейными автоматами

Пусть задан ЛА L. Для каждого $a \in \mathbf{R}$ запись L_a обозначает подмножество множества X^*, определяемое следующим образом:

$$L_a \stackrel{\text{def}}{=} \{u \in X^* \mid f_L(u) > a\}. \tag{83}$$

Множество (83) называется **языком**, представимым ЛА L с точкой сечения a.

Теорема 23.
Пусть $L = (\xi^0, \{L^x \mid x \in X\}, \lambda)$ – ЛА, и $a \in \mathbf{R}$.
Тогда существует ВА A вида

$$(\vec{e}_1, \{A^x \mid x \in X\}, e_1^{\downarrow}), \tag{84}$$

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 23 достаточно доказать следующее утверждение: для каждого ЛА L существует ВА A вида (84), 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:

$$\forall\, u \in X^* \quad f(u) \overset{\text{def}}{=} \begin{cases} 1, & \text{если } u = \varepsilon \\ f_L(u), & \text{иначе.} \end{cases}$$

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 3), 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:

$$\forall\, u \in X^* \quad f(u) \overset{\text{def}}{=} \begin{cases} 0, & \text{если } u = \varepsilon \\ f_L(u), & \text{иначе.} \end{cases}$$

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 3), такой, что $\exists\, b > 0$:

$$\forall\, u \in X^* \quad f_A(u) = \begin{cases} 0, & \text{если } u = \varepsilon, \\ b^{|u|+1} f_{L_2}(u) + \frac{1}{n+3}, & \text{иначе.} \end{cases} \blacksquare$$

Теорема 24.
Пусть X – конечное множество, и $S \subseteq X^*$.
Следующие условия эквивалентны:

- S – ВЯ,
- $\exists$ линейно-автоматная функция $f \in \mathbf{R}^{X^*}$, $\exists\, a \in \mathbf{R}$: $S = f_a \overset{\text{def}}{=} \{u \in X^* \mid f(u) > a\}$,
- $\exists$ ЛАФ $f \in \mathbf{R}^{X^*} : S = f_0$.

В доказательстве этой теоремы используется утверждение о том, что если f – ЛАФ и $a \in \mathbf{R}$, то $f - a$ – ЛАФ. ■

Список литературы

[1] Миронов, А.М., Основные понятия теории вероятностных автоматов. Интеллектуальные системы. Теория и приложения. 2015, вып. 4.

[2] Rabin, M.O., Probabilistic automata. Information and Control 6(3), 230–245 (1963). (русский перевод: Рабин М.О. Вероятностные автоматы / Кибернетический сборник, - Вып. 9. -М.: Иностранная литература, 1964.- С. 123-141.)

[3] Хопкрофт Д., Мотвани Р., Ульман Дж. Введение в теорию автоматов, языков и вычислений 2-е изд.: Пер. с англ. - М.: Издательский дом "Вильямс 2002. - 528 с.

[4] Кемени Дж., Снелл Дж. Конечные цепи Маркова.—М.: Наука, 1970.

[5] Carlyle J. W., Reduced forms for stochastic sequential machines, J. Maht. Analysis and Application, 1963, v. 7, № 2 (русский перевод: Карлайл Е.У., Приведенные формы для стохастических последовательностных машин // Кибернетический сборник. Новая серия.- М.: Мир, 1966. - Вып. 3. - С. 101-110.).

[6] Бухараев Р.Г. Некоторые эквивалентности в теории вероятностных автоматов, уч. записки Казан. ун-та, 1964, 124, номер 2, с. 45-65.

[7] Starke P. H., Theorie stochastischen Automaten, I, II, Elektronische Informationsverarbeitung und Kybernetik, 1965, 1, No. 2.

[8] A. Paz, Introduction to Probabilistic Automata. Academic Press, New York, USA, 1971.

[9] Бухараев Р. Г. Основы теории вероятностных автоматов.—М.: Наука, 1985.

[10] R.Segala, N.A.Lynch, Probabilistic simulations for probabilistic processes, Nordic Journal of Computing, 2 (2) (1995), p. 250–273.

[11] M. Stoelinga, An introduction to probabilistic automata, Bulletin of the European Association for Theoretical Computer Science, 2002, p. 176–198.

[12] A. Sokolova, E.P. de Vink, Probabilistic Automata: System Types, Parallel Composition and Comparison, in: C. Baier et al. (Eds.), Validation of Stochastic Systems, LNCS 2925, pp. 1-43, 2004.

[13] L.R. Rabiner, A tutorial on Hidden Markov Models and selected applications in speech recognition, in Proceedings of the IEEE 77 (2): 257–286, 1989.

[14] A. Darwiche, Modeling and Reasoning with Bayesian Networks. Cambridge University Press, 2009. 562 p.

[15] D.Koller and N.Friedman, Probabilistic Graphical Models. Principles and Techniques. Massachusetts: MIT Press, 2009. 1280 p.

[16] E.A. Feinberg and A. Shwartz (eds.), Handbook of Markov Decision Processes, Kluwer, Boston, MA, 2002. 562 p.

[17] S.-H. Wu, S. A. Smolka, and E. W. Stark, Composition and behaviors of probabilistic I/O automata, Theoretical Computer Science 176 (1997), p. 1-38.

[18] B. Delahaye, J.-P. Katoen, K.G. Larsen, A.Legay, M.L. Pedersen, F. Sher, A. Wasowski, Abstract Probabilistic Automata, Information and Computation, Elsevier, 232, (2013), p. 66-116.

[19] M. Kudlek. Probability in Petri nets. Fundamenta Informaticae, 67 (1-3): 121-130, 2005.

[20] Y. Liu, H. Miao, H. Zeng, and Z. Li. Probabilistic Petri net and its logical semantics. In Software Engineering Research, Management and Applications, pages 73–78, 2011.

[21] C. Eisentraut, H. Hermanns, and L. Zhang. On probabilistic automata in continuous time. In Proc. of 25th Annual IEEE Symposium on Logic in Computer Science (LICS), pages 342-351, 2010.

[22] B. Jonsson, K.G. Larsen, and W. Yi, Probabilistic extensions of process algebras, Handbook of Process Algebras, Elsevier, North Holland, 2001.

[23] Бухараев Р. Г. Сети вероятностных процессоров // Математические вопросы кибернетики. Вып. 16. — М.: Физматлит, 2007. — С. 57–72.

URL: http://library.keldysh.ru/mvk.asp?id=2007-57

[24] Э. Ф. Мур, Умозрительные эксперименты с последовательностными машинами. сб. "Автоматы". М.: ИЛ. - 1956. - с. 179-210.

[25] А. М. Миронов, С. Л. Френкель. Минимизация вероятностных моделей программ. Фундаментальная и прикладная математика, т. 19, вып.1, с. 121-163. (2014)

[26] Kiefer, S., Wachter, B., Stability and Complexity of Minimising Probabilistic Automata. J. Esparza et al. (Eds.): ICALP 2014, Part II, LNCS 8573, pp. 268–279, 2014. Springer-Verlag Berlin Heidelberg.

[27] Mateus, P., Qiu, D., Li, L.: On the complexity of minimizing probabilistic and quantum automata. Information and Computation 218, 36–53 (2012)