О реализуемости функций на строках вероятностными автоматами

А. М. Миронов

Аннотация. В работе формулируется и доказывается необходимое и достаточное условие реализуемости функций на строках вероятностными автоматами Мура с числовым выходом.

Ключевые слова: *вероятностные автоматы, автоматы Мура, реакции, случайные функции.*

Введение

Понятие **вероятностного автомата** впервые было сформулировано в 1963 г. в основополагающей работе М. Рабина [1]. Данное понятие было предназначено главным образом для изучения вопросов представимости регулярных языков вероятностными автоматами. Затем это понятие было обобщено до такого понятия, которое позволило моделировать вероятностные преобразователи информации. Определение вероятностного автомата в общей форме было введено независимо в работах Дж. Карлайла [2], Р. Г. Бухараева [3], и П. Штарке [4]. Одним из частных случав общего понятия вероятностного автомата является рассматриваемое в настоящей работе понятие вероятностного автомата Мура с числовым выходом.

Главный результат работы (теорема 1) представляет собой необходимое и достаточное условие реализуемости функций на строках вероятностными автоматами Мура с числовым выходом. Вопрос о реализуемости вероятностных функций на строках вероятностными автоматами общего вида исследовался Р. Г. Бухараевым, и им был найден критерий такой реализуемости, который изложен в [5] (теорема 2.4.3, с. 76). Суть этого критерия заключается в следующем. Вводятся понятия множества S_f состояний функции на строках f и опорного множества для S_f. Критерий имеет следующий вид: функция на строках f реализуется в конечном вероятностном автомате общего вида тогда и только тогда, когда существует конечное опорное множество для S_f, выпуклое относительно полугруппы всех вращений. Проверка данного критерия представляется затруднительной по той причине, что для этого необходимо вычислить всё множество S_f, что сложно сделать в том случае, когда данное множество является бесконечным. Доказательство этого критерия основано на представлении функций на строках как точек счетномерного линейного пространства с индексацией координат элементами множества строк. В доказательстве используются бесконечномерные матрицы, индексированные парами строк.

Сформулированный и доказанный в настоящей работе критерий реализуемости функций на строках в вероятностных автоматах Мура с числовым выходом имеет следующий вид: функция f на строках реализуема конечным вероятностным автоматом Мура с числовым выходом тогда и только тогда, когда она принадлежит выпуклой оболочке некоторого конечного множества функций на строках, устойчивого относительно сдвигов. Данный критерий существенно проще критерия из [5] (описанного в предыдущем абзаце), поскольку он требует проверки лишь для одной функции f, а не для бесконечного множества функций из S_f. Кроме того, доказательство этого критерия имеет существенно более простой вид, чем доказательство критерия из [5], и не использует бесконечномерных матриц.

1. Вспомогательные понятия

1.1. Случайные функции и распределения

Пусть задана пара множеств X, Y. **Случайной функцией (СФ)** из X в Y называется произвольная функция f вида

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

такая, что $\forall x \in X$ множество $\{y \in Y \mid f(x,y) > 0\}$ конечно или счётно, и $\forall x \in X \quad \sum_{y \in Y} f(x,y) = 1$.

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 $f : X \xrightarrow[r]{} Y$. Мы будем называть X **областью определения** СФ f, а Y — **областью значений** СФ f.

СФ (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), и x, y — такие элементы X и Y соответственно, что $f(x,y) = 1$, то мы будем говорить, что f **отображает** x **в** y.

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

Пусть задана КСФ $f : X \xrightarrow[r]{} Y$, 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)

$$f = \begin{pmatrix} f(x_1, y_1) & \ldots & f(x_1, y_n) \\ \ldots & \ldots & \ldots \\ f(x_m, y_1) & \ldots & f(x_m, y_n) \end{pmatrix}. \qquad (2)$$

Ниже мы будем отождествлять каждую КСФ f с соответствующей ей матрицей (2). 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.

Вероятностным распределением (или просто **распределением**) на множестве X называется СФ вида $\xi : 1 \xrightarrow[r]{} X$, где — 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: $\forall\, y \in X \;\; y^{\xi_x} \stackrel{\text{def}}{=} 1$, если $y = x$, и $y^{\xi_x} \stackrel{\text{def}}{=} 0$, если $y \neq x$.

1.2. Строки и функции на строках

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 каждой пары строк $u, v \in X^*$ их **конкатенацией** называется строка, обозначаемая записью uv и определяемая следующим образом: $u\varepsilon \stackrel{\text{def}}{=} \varepsilon u \stackrel{\text{def}}{=} u$, и если $u = x_1 \ldots x_n$ и $v = x'_1 \ldots x'_m$, то $uv \stackrel{\text{def}}{=} x_1 \ldots x_n x'_1 \ldots x'_m$.

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** $f_1 + f_2$ определяется следующим образом:

$$\forall\, u \in X^* \quad (f_1 + f_2)(u) \stackrel{\text{def}}{=} f_1(u) + f_2(u).$$

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

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

Множество $\mathbf{R}^{X^*}$ можно рассматривать как векторное пространство над $\mathbf{R}$ относительно определённых выше операций сложения и умножения на числа из $\mathbf{R}$. Для каждого $\Gamma \subseteq \mathbf{R}^{X^*}$ **выпуклой оболочкой** множества Γ называется подмножество $C(\Gamma) \subseteq \mathbf{R}^{X^*}$, состоящее всех выпуклых комбинаций $\sum_{i=1}^{n} a_i f_i$, где $\forall\, i = 1, \ldots, n \quad a_i \in [0,1],\, f_i \in \Gamma, \quad \sum_{i=1}^{n} a_i = 1$.

2. Вероятностные автоматы Мура

2.1. Понятие вероятностного автомата Мура

Вероятностный автомат (ВА) Мура — это шестерка A вида

$$A = (X, Y, S, \delta, \lambda, \xi^0), \tag{3}$$

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. δ — СФ вида $P: S \times X \xrightarrow[r]{} S$, называемая **функцией перехода ВА** A. $\forall (s, x, s') \in S \times X \times S$ значение $\delta(s, x, s')$ 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'.

3. λ — СФ вида $P: S \xrightarrow[r]{} Y$, называемая **фунцией выхода ВА** A. $\forall (s, y) \in S \times Y$ значение $\lambda(s, y)$ понимается как вероятность того, что если в текущий момент времени A находится в состоянии s, то в этот момент времени выходной сигнал A равен y.

4. 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.

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

$$\begin{pmatrix} \delta(s_1, x, s_1) & \dots & \delta(s_1, x, s_n) \\ \dots & \dots & \dots \\ \delta(s_n, x, s_1) & \dots & \delta(s_n, x, s_n) \end{pmatrix}, \tag{4}$$

где δ — функция перехода **ВА** A.

$\forall x \in X$, $\forall s, s' \in S$ мы будем обозначать записью $A^x_{s,s'}$ коэффициент 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 коэффициенты матрицы A^x обладают свойствами

$$\begin{gathered} \forall s, s' \in S \;\; A^x_{s,s'} \geq 0, \\ \forall s \in S \;\; \sum_{s' \in S} A^x_{s,s'} = 1 \;\; (\text{т.е. } A^x I = I). \end{gathered} \tag{5}$$

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

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

Нетрудно доказать, что матрица A^u — стохастическая.

$\forall u \in X^*$, $\forall s, s' \in S$ мы будем обозначать записью $A^u_{s,s'}$ коэффициент 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 $(X, Y, S, \delta, \lambda, \xi^0)$, функция выходов λ которого является детерминированной (т. е. можно считать, что λ имеет вид $S \to Y$). Мы будем обозначать вектор-столбец $\begin{pmatrix} \lambda(s_1) \\ \dots \\ \lambda(s_n) \end{pmatrix}$ значений функции выхода $\lambda: S \to Y$ **ВА** A (где $(s_1, ..., s_n)$ — фиксированное упорядочение множества S) тем же символом λ.

ВА Мура с числовым выходом — **ВА** Мура с детерминированным выходом, множество выходных сигналов которого является подмножеством множества **R** действительных чисел.

Пусть $A = (X, \mathbf{R}, S, \delta, \lambda, \xi^0)$ — **ВА** Мура с числовым выходом, и $\xi \in S^\Delta$. Мы будем говорить, что A в момент времени t имеет распределение ξ, если для каждого состояния $s \in S$ вероятность того, что A в момент t находится в состоянии s, равна s^ξ.

Реакция ВА $A = (X, \mathbf{R}, S, \delta, \lambda, \xi^0)$ в распределении $\xi \in S^\triangle$ — это функция $A^\xi : X^* \to \mathbf{R}$, определяемая следующим образом:

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

Реакцию **ВА** A в его начальном распределении мы будем называть просто **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 $t+k+1$.

3. Вероятностная реализуемость функций на строках

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

Пусть X — конечное множество. Мы будем использовать следующие определения и обозначения.

- $\forall x \in X$ 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

$$\forall u \in X^* \quad (fD^x)(u) \overset{\text{def}}{=} f(xu). \tag{6}$$

- 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 1.

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

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

Доказательство.

Пусть f вероятностно реализуема, т. е. $\exists$ **ВА** Мура с числовым выходом $A = (X, \mathbf{R}, S, \delta, \lambda, \xi^0)$:

$$\forall u \in X^* \quad f(u) = \xi^0 A^u \lambda.$$

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 $A_i \overset{\text{def}}{=} (X, \mathbf{R}, S, \delta, \lambda, \xi_i)$.

Полагаем $\Gamma_f \overset{\text{def}}{=} \{f_{A_i} \mid i = 1, \ldots, n\}$. $f \in C(\Gamma_f)$, т. к. $f = \sum_{i=1}^{n} i^{\xi^0} f_{A_i}$.

Докажем, что

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

Согласно (6),

$$\forall u \in X^* \quad (f_{A_i} D^x)(u) = f_{A_i}(xu) = \xi_i A^{xu} \lambda = \xi_i A^x A^u \lambda. \tag{7}$$

Нетрудно видеть, что

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

Таким образом, $f_{A_i} D^x = \sum_{j=1}^{n} A_{ij}^x f_{A_j} \in C(\Gamma_f)$.

Обратно, пусть $f \in C(\Gamma_f)$, где $\Gamma_f = \{f_1, \ldots, f_n\}$, и Γ_f устойчиво относительно сдвигов. Определим **ВА** Мура с числовым выходом

$$A \overset{\text{def}}{=} (X, \mathbf{R}, S, \delta, \lambda, \xi^0), \tag{8}$$

где

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

$$f = \sum_{i=1}^{n} i^{\xi^0} f_i, \tag{9}$$

- $\forall x \in X, \forall i = 1, \ldots, n$ строка i матрицы A^x (где $A_{ij}^x = \delta(i, x, j)$) состоит из коэффициентов представления функции $f_i D^x$ в виде выпуклой комбинации функций из Γ_f, т. е.

$$f_i D^x = \sum_{j=1}^{n} A^x_{ij} f_j, \tag{10}$$

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

Докажем, что реакция **ВА** (8) совпадает с f, т. е.

$$\forall u \in X^* \quad \xi^0 A^u \lambda = f(u). \tag{11}$$

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

$$A^u \lambda = \begin{pmatrix} f_1(u) \\ \dots \\ f_n(u) \end{pmatrix}. \tag{12}$$

Если $u = \varepsilon$, то обе части (12) совпадают по определению λ.

Если $u = xu'$, то, предполагая верным равенство (12), в котором u заменено на u', имеем:

$$A^u \lambda = A^{xu'} \lambda = A^x A^{u'} \lambda = A^x \begin{pmatrix} f_1(u') \\ \dots \\ f_n(u') \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{i=1}^{n} A^x_{1j} f_j(u') \\ \dots \\ \sum_{i=1}^{n} A^x_{nj} f_j(u') \end{pmatrix}. \tag{13}$$

Из (10) следует, что правую часть в (13) можно переписать в виде

$$\begin{pmatrix} (f_1 D^x)(u') \\ \dots \\ (f_n D^x)(u') \end{pmatrix}. \tag{14}$$

Согласно определению (6) функций вида fD^x, столбец (14) совпадает с правой частью доказываемого равенства (12).

Таким образом, равенство (12) доказано. Согласно этому равенству, левая часть доказываемого равенства (11) равна

$$\xi^0 \begin{pmatrix} f_1(u) \\ \dots \\ f_n(u) \end{pmatrix} = \sum_{i=1}^{n} i^{\xi^0} f_i(u). \tag{15}$$

По определению ξ^0 (см. (9)), правая часть (15) равна $f(u)$, т. е. правой части доказываемого равенства (11).

Заключение

Критерий реализуемости функций на строках конечными вероятностными автоматами Мура с числовым выходом, изложенный в настоящей работе, является более простым, чем соответствующий критерий, сформулированный в теореме 2.4.3 книги [5]. Однако проверка этого критерия для заданной функции на строках f может представлять некоторые трудности, поскольку для доказательства реализуемости f необходимо построить конечное множество функций Γ_f, удовлетворяющее условию теоремы 1. Одним из направлений развития изложенного в настоящей работе результата может быть нахождение стратегий построения для заданной функции на строках f соответствущего множества Γ_f.

Кроме того, поскольку множество Γ_f для заданной функции f можно рассматривать как множество состояний одного из **ВА**, реакция которого совпадает с f, то, следовательно, к проблеме построения для заданной функции f соответствующего множества Γ_f с наименьшим возможным числом элементов сводится проблема построения для заданного **ВА** A такого **ВА**, реакция которого совпадает с реакцией **ВА** A, и который содержит наименьшее возможное число состояний (поскольку в качестве исходной функции f можно рассматривать реакцию **ВА** A). Данная проблема известна в литературе по теории автоматов как проблема минимизации автоматов, и является одним из наиболее популярных предметов исследований в области теории вероятностных автоматов. Среди последних результатов, относящихся к решению данной проблемы, отметим работы [6]–[8]. Одним из путей развития данных результатов может быть разработка на их основе методов построения для заданной функции f соответствующего множества Γ_f, которое содержит как можно меньшее число элементов.

Литература

1. *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).
2. *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).

3. *Бухараев Р. Г.* Некоторые эквивалентности в теории вероятностных автоматов // Уч. записки Казан. ун-та. 1964. 124. № 2. С. 45–65.
4. *Starke P. H.* Theorie stochastischen Automaten, I, II // Elektronische Informationsverarbeitung und Kybernetik. 1965. 1. № 2.
5. *Бухараев Р. Г.* Основы теории вероятностных автоматов. М.: Наука, 1985.
6. *Миронов А. М.*, *Френкель С. Л.* Минимизация вероятностных моделей программ // Фундаментальная и прикладная математика. Т. 19. Вып. 1. С. 121–163 (2014).
7. *Kiefer S.*, *Wachter B.* Stability and Complexity of Minimising Probabilistic Automata // J. Esparza et al. (Eds.): ICALP 2014, Part II, LNCS 8573. Springer-Verlag Berlin Heidelberg, 2014. P. 268–279.
8. *Mateus P.*, *Qiu D.*, *Li L.* On the complexity of minimizing probabilistic and quantum automata // Information and Computation. 218, 36–53 (2012).

Миронов Андрей Михайлович. С. н. с. ФИЦ ИУ РАН. К. ф.-м. н. Окончил в 1989 году мехмат МГУ. Количество печатных работ и монографий: 43. Область научных интересов: теоретическая информатика, математическая логика, теория автоматов, верификация программ. E-mail: amironov66@gmail.com