

УДК 519.71

DOI: 10.35330/1991-6639-2024-26-6-165-174

EDN: NYEZQR

Научная статья

## Анализ целесообразного поведения различных типов автоматов

Д. П. Димитриченко

Институт прикладной математики и автоматизации –  
филиал Кабардино-Балкарского научного центра Российской академии наук  
360000, Россия, г. Нальчик, ул. Шортанова, 89 А

**Аннотация.** Коллективное поведение автоматов является одним из направлений развития методов машинного обучения. Такие автоматы реализуют функцию целенаправленного поведения. Автомат выполняет действие, в ответ на которое окружающая среда подает свой выходной сигнал на вход автомата. Автомат, в соответствии со своей конструкцией, реагирует на этот входной сигнал очередным действием. Таким образом, выстраивается замкнутый контур взаимодействия между некоторой окружающей средой и функционирующим в ней автоматом. Сама эта среда во многих случаях допускает автоматную реализацию. Оценка эффективности автомата определяется как оптимизационная задача максимизации суммы положительных сигналов (поощрений) или минимизации отрицательных сигналов (штрафов), получаемых от окружающей среды за рассматриваемый промежуток времени. Формализация как свойств окружающей среды, так и действий автоматов, а также обработка полученных результатов производится при помощи аппарата теории игр. В этом случае сигналы от окружающей среды удобно представить, как суммы выигрышей и проигрышей игроков-автоматов. В настоящей работе проведено сравнение автоматов различных конструкций, так как эффективность реакций автоматов определяется не только свойствами окружающей среды, но и такими параметрами, как тип и глубина памяти.

**Ключевые слова:** автомат, целесообразное поведение, оптимальная стратегия, глубина памяти, теория игр, формализация окружающей среды, динамическая среда

Поступила 11.10.2024, одобрена после рецензирования 14.11.2024, принята к публикации 02.12.2024

**Для цитирования.** Димитриченко Д. П. Анализ целесообразного поведения различных типов автоматов // Известия Кабардино-Балкарского научного центра РАН. 2024. Т. 26. № 6. С. 165–174. DOI: 10.35330/1991-6639-2024-26-6-165-174

MSC: 68P01

Original article

## Analysis of expedient behavior of various types of automata

D.P. Dimitrichenko

Institute of Applied Mathematics and Automation –  
branch of Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences  
360000, Russia, Nalchik, 89 A Shortanov street

**Abstract.** Collective behavior of automata is one of the directions of development of machine learning methods. Such machines fulfil the function of goal-oriented behavior. The machine performs an action, in response to which the environment sends its output signal to the input of the machine. The machine, in accordance with its design, responds to this input signal with the next action. Thus, a closed loop of interaction is built between a certain environment and the machine operating in it. This environment

itself in many cases allows for machine implementation. Effectiveness evaluation of the machine is defined as an optimization problem of maximizing the sum of positive signals (rewards), or minimizing negative signals (penalties), received from the environment, over the considered period of time. Formalization of both the properties of the environment and the actions of the machines, as well as processing of the obtained results is performed using the apparatus of game theory. In this case, signals from the environment are conveniently represented as the sums of the winnings and losses of the players-machines. In this paper, a comparison of machines of different designs is carried out, since the efficiency of machine reactions is determined not only by the properties of the environment, but also by such parameters as the type and depth of memory.

**Keywords:** automaton, expedient behavior, optimal strategy, memory depth, game theory, formalization of the environment, dynamic environment

Submitted 11.10.2024,

approved after reviewing 14.11.2024,

accepted for publication 02.12.2024

**For citation.** Dimitrichenko D.P. Analysis of expedient behavior of various types of automata. *News of the Kabardino-Balkarian Scientific Center of RAS*. 2024. Vol. 26. No. 6. Pp. 165–174. DOI: 10.35330/1991-6639-2024-26-6-165-174

## ВВЕДЕНИЕ

Организация сложного поведения децентрализованных систем, построенных на базе простых элементов, находит свое отражение в таком разделе машинного обучения, как «Коллективное поведение автоматов» [1]. При этом связь между нейросетевым и автоматным подходами бывает очень тесной [2].

Анализ поведения автоматов будет проводиться при помощи аппарата теории игр. Такой выбор определяется тем, что теория игр позволяет легко создавать формализуемые системы правил и стратегий поведения, определяющие характер взаимодействия отдельного автомата (или совокупности автоматов) и окружающей среды [3].

Это обстоятельство позволяет сформировать замкнутый контур Среда-Автомат, в котором действия автомата определяют реакцию окружающей среды, которая в свою очередь через обратную связь оказывает влияние на выбор последующего действия автомата.

Первые интересные модели такого типа были созданы М. Л. Цетлиным (1924–1966) [4]. Он выступил создателем целого направления исследований, получившего название «коллективное поведение автоматов» [5–7]. Им были сформулированы основные принципы, лежащие в основе подобных моделей, и способы их реализации.

Формализация понятия целенаправленного поведения позволила М. Л. Цетлину сформулировать базовый принцип организации сложного поведения совокупности децентрализованных систем, образующих коллективы автоматов.

Основные положения этого подхода состоят в следующем:

- 1) принцип суперпозиции;
- 2) принцип соизмеримости;
- 3) принцип универсальности структуры автомата.

*Принцип суперпозиции.* Любое достаточно сложное поведение складывается из совокупности простых поведенческих актов. Совместная реализация таких простых поведенческих актов и простейшее взаимодействие приводят в результате к весьма сложным поведенческим процессам.

*Принцип соизмеримости.* Степень целесообразного поведения анализируемой кибернетической системы рассматривается как величина математического ожидания совокупности штрафов и поощрений, полученных этой системой от окружающей среды за наблюдаемый промежуток времени  $t$ . Эта величина находится в интервале со следующими границами:

Левая граница – это значение математического ожидания количества штрафов и поощрений, полученных системой, реализующей стратегию случайного выбора доступных в данной среде действий из множества  $D = d_1, \dots, d_m, m \geq 2$ .

Правая граница – это величина математического ожидания штрафов и поощрений, полученных системой, которая в любой момент времени  $t, t \gg 1$  всегда реализует заведомо оптимальное в данной среде действие  $d_i = d^*, i = 1, \dots, m$ .

Очевидно, что чем ближе величина математического ожидания поощрений анализируемой системы к левой границе определенного таким образом интервала, тем точнее ее поведение соответствует стратегии «случайного выбора». А чем ближе эта величина к правой границе интервала, тем ближе характер поведения такой системы соответствует наилучшей возможной стратегии. Такое определение целесообразного поведения не зависит от структуры анализируемой системы, а опирается только на результат взаимодействия со средой.

*Принцип универсальности.* Конструкция анализируемой системы не должна содержать эмпирических данных об оптимальных (или неоптимальных) действиях в данной среде, т. е. такие действия система выявляет в процессе своего функционирования в течение некоторого заданного времени  $t$ . Первая автоматная реализация таких систем была предложена М. Л. Цетлиным.

#### АВТОМАТНАЯ РЕАЛИЗАЦИЯ

Приведем формальное описание автомата:

$X = x_1, \dots, x_k$  – множество входных сигналов, поступающих от окружающей среды на вход автомата (входной алфавит).

$D = d_1, \dots, d_m$  – конечное множество доступных автомату действий (выходной алфавит).

$S = s_1, \dots, s_n, 2 \leq m \leq n$  – конечное множество внутренних состояний автомата (внутренний алфавит).

Правила функционирования автомата в дискретные моменты времени  $t$  однозначным образом задаются двумя функциями:

функцией переходов внутренних состояний  $s_{t+1} = F(s_t, x_t)$  и начальным внутренним состоянием в нулевой момент времени  $t = t_0: s_0 = S(t_0)$ ;

функцией зависимости выходных сигналов (действий) от внутренних состояний  $d_t = G(s_t)$ .

#### ОБЩАЯ ПОСТАНОВКА ЗАДАЧИ

Пусть задана некоторая среда  $E$ , в которой будут функционировать типы исследуемых автоматов.

Пусть также имеются типы автоматов, для которых возможно установить степень целесообразного поведения в среде  $E$  определенным выше образом.

Требуется:

- определить степень влияния глубины памяти  $q$  каждого из рассматриваемых типов автоматов на изменение степени целесообразности поведения в указанной среде  $E$ ;
- на основе полученных результатов определить тип конструкции автомата, достигающей заданного уровня целесообразности с минимально возможной глубиной памяти  $q = q^*$ ;
- в случае наличия нескольких эквивалентных конструкций указать наиболее простую в реализации конструкцию.

#### БИОЛОГИЧЕСКИЕ ПРЕДПОСЫЛКИ

Результаты из экспериментальной биологии потребовали сформулировать кибернетический «белый ящик», обладающий целесообразным поведением.

Подопытное животное помещалось в основание Т-образного лабиринта с возможностью выбора одного из двух действий:

«Повернуть влево»;  
«Повернуть вправо».

По условиям эксперимента в конце каждого из двух ответвлений реализовывались неизвестные подопытному животному благоприятные (или неблагоприятные) условия с независимыми вероятностями для каждого из ответвлений.

Требовалось установить способность подопытных животных различать условия окружающей среды, носящие вероятностный характер.

Оказалось, что несмотря на ошибки выбора в начале серии экспериментов животные впоследствии верно ассоциировали выбираемый ими поворот (выполняемое действие) с тем, для которого вероятность штрафа (вероятность попасть в неблагоприятную ситуацию) была минимальной.

#### СТАЦИОНАРНАЯ ОКРУЖАЮЩАЯ СРЕДА

Стационарная окружающая среда  $E$  является математическим описанием условий T-образного лабиринта.

В T-образном лабиринте было всего два доступных для выполнения действия: «повернуть влево» и «повернуть вправо». В каждом из выбранных ответвлений ожидалась своя вероятность поощрения  $P_l$  и  $P_r$ .

Обобщением этой ситуации является стационарная среда с  $m$  исходами (действиями).

Для каждого из  $m$  действий задается совокупность значений вероятностей: либо поощрения  $p_i$ , либо штрафа  $1 - p_i$ ,  $i=1, \dots, m$ .

Выбор оптимального действия в такой среде сводится к определению действия с максимальной вероятностью поощрения (минимальной вероятностью штрафа).

При этом ни подопытное животное в биологических экспериментах, ни соответствующий ему по поведению автомат не располагают знаниями об исходных значениях вероятностей, но вынуждены эту информацию получать в опосредованной форме через сигналы поощрения и штрафа.

#### ТИПЫ КОНСТРУКЦИЙ АВТОМАТОВ

Первой конструкцией автомата, способного вести себя целесообразно в описанной выше стационарной среде  $E$ , является автомат с линейной тактикой, предложенный М. Л. Цетлиным.

Автомат L1 (с глубиной памяти  $q=1$ ) состоит из  $m$  состояний  $S=s_1, \dots, s_m$ .

За каждым из  $m$  состояний  $s_i$  закреплено действие  $d_i$ ,  $i=1, \dots, m$ .

Автомат L1 функционирует следующим образом:

В некоторый момент времени  $t=t^*$ , находясь в состоянии  $s_i$ , автомат L1 выполняет выходное действие  $d_i$ ,  $i=1, \dots, m$ .

В ответ на это действие окружающая среда и формирует сигнал, принимающий значение «поощрение» в соответствии с вероятностью  $p_i$  и «штраф» с вероятностью  $1 - p_i$ .

Этот сигнал подается на вход автомата L1.

Если автомат L1 на входе получает сигнал «штраф», то он переходит в состояние  $s_{i+1}$  и, следовательно, переключает внешнее действие с  $d_i$  на  $d_{i+1}$ .

При получении сигнала «поощрение» автомат L1 остается в исходном состоянии  $s_i$ , при этом смена действия не происходит.

Таким образом, в следующий момент времени  $t^*+1$  автомат L1 выполнит действие в соответствии со своим внутренним состоянием, и процесс взаимодействия автомата со средой повторится описанным выше образом.

Закрепляя за каждым из  $m$  действий  $d_i$   $q$  последовательных состояний, мы получим последовательность линейных автоматов с монотонно возрастающей глубиной памяти  $q$ .

М. Л. Цетлину удалось показать, что последовательность таких автоматов  $L1, L2, \dots, Lq$ , функционирующих в стационарной среде  $E$ , является асимптотически оптимальной, т.е. чем больше глубина памяти линейного автомата  $Lq$ , тем дольше такой автомат выполняет самое оптимальное действие (с максимальной вероятностью поощрения) и почти никогда не покидает связанные с ним состояния.

Усложнением автомата с линейной тактикой является автомат Кринского.

Он характеризуется тем, что при получении поощрения этот автомат всегда переходит в самое глубокое состояние, соответствующее текущему (выполненному им) действию.

При получении штрафа от окружающей среды этот автомат реагирует точно так же, как и автомат с линейной тактикой, понижая на единицу номер состояния текущего (выполненного) действия.

Для этого автомата также доказана теорема об асимптотической оптимальности при функционировании в стационарных средах  $E$ .

Следующим по степени усиления свойства инерционности является автомат Роббинса.

Он отличается от автомата Кринского тем, что в отличие от него при смене действия автомат Роббинса сразу переходит в самое глубокое состояние, соответствующее новому действию.

В остальном он ничем не отличается от автомата Кринского.

Для него также верна теорема об асимптотической оптимальности в стационарных средах.

Заметим, что у трех автоматов – автомата с линейной тактикой, автомата Кринского и автомата Роббинса – при глубине памяти  $q = 1$  алгоритмы функционирования полностью совпадают, утрачивая отличия между ними.

Несколько другой подход к усилению инерционности применен в конструкции автомата Крылова.

Он сочетает в себе элементы поведения как детерминированного, так и стохастического автомата.

При получении поощрения автомат функционирует точно так же, как и автомат с линейной тактикой, увеличивая вплоть до самого глубокого номер состояния, соответствующего действию, за выполнение которого этот автомат получил поощрение от среды  $E$ .

При получении штрафа автомат Крылова «подбрасывает монетку».

При выпадении «орла» автомат не меняет своего состояния, а при выпадении «решки» уменьшает номер состояния за действие, за которое от среды получен штраф.

Этот автомат также является асимптотически оптимальным в условиях стационарных сред.

#### ДИНАМИЧЕСКИЕ СРЕДЫ

Первоначально формальным языком описания как стационарных, так и динамических сред с функционирующими в них автоматами послужила теория игр.

Действия автомата из множества  $D$  взаимно однозначно соотносятся со стратегиями, доступными игроку в соответствующей (заданной) игре с конечным числом стратегий.

Например, действия автомата в уже рассмотренной выше стационарной среде  $E$  на языке теории игр формулируется как игра с природой:

- 1) задана одностробцовая платежная матрица с  $m$  неотрицательными значениями;
- 2) элементы матрицы игры нормируются относительно максимального значения платежа и рассматриваются как вероятности поощрений от стационарной среды  $E$ ;
- 3) значение поощрения равно единице, значение штрафа равно нулю;
- 4) действиям автомата (игрок 1) соответствуют номера строк заданной матрицы;
- 5) природа (игрок 2) выбрала свою стратегию (столбец) и не меняет ее в течение всех партий игры;
- 6) целью игры является максимизация выигрыша автомата.

Игрок 1 (автомат) в каждый момент времени  $t$  разыгрывает партию игры с природой и в зависимости от выбранного действия  $d^*=d_i$  и соответствующей этому действию вероятности  $p^*=p_i$  получает сигнал поощрения (или штрафа) от окружающей среды  $I, i=1, \dots, m$ .

Поскольку оптимальная стратегия игры с природой известна и соответствует выбору игроком действия с максимальной ценой  $p_i=p_{max}$ , то создается возможность сравнить эффективность действия осведомленного игрока, которому известна вся платежная матрица, с действиями автомата, которому не известно содержимое этой матрицы.

Из теорем об асимптотической оптимальности рассмотренных выше типов автоматов следует, что чем больше глубина памяти  $q$  автомата, тем точнее его совокупные действия соответствуют оптимальной стратегии.

### ИГРЫ ДВУХ АВТОМАТОВ

Случай матричной игры двух игроков с нулевой суммой – это пример того, как один автомат способен создать среду для другого автомата.

В этом случае действиям первого автомата сопоставляются номера строк, а действиям второго автомата – номера столбцов платежной матрицы.

Если первый автомат с вероятностью  $p$  получает поощрение, то второй автомат получает штраф, и наоборот, реализуя принцип антагонистической игры с нулевой суммой.

Задача первого автомата состоит в максимизации поощрений, задача второго автомата – в минимизации штрафов.

Если игра допускает решение в чистых стратегиях, то автоматы реализуют друг для друга стационарную среду, для которой повышение целесообразности связано с ростом глубины памяти  $q$ .

Решение в смешанных стратегиях требует от автоматов оптимальной глубины памяти, удовлетворяющей свойствам данной динамической среды и типа автомата.

Аналогичным образом строится игра двух автоматов в случае бескоалиционных, неантагонистических, биматричных игр с конечным числом стратегий.

### ИГРА В РАЗМЕЩЕНИЯ

Примером того, как коллектив автоматов формирует динамическую среду для отдельного автомата, является игра в размещения.

Биологической предпосылкой для игры в размещения служит следующая задача:

1) имеется конечное число участков с различной степенью производительности пищи. На этих участках некоторым образом размещаются животные;

2) если два и более животных в текущий момент времени оказываются на одном участке, то производительность пищи участка делится поровну между этими животными.

Задача животного состоит в максимизации количества добываемой им пищи.

Описание игры в размещения легко формулируется на языке теории игр следующим образом:

1) пусть задано  $m$  стратегий  $d = d_1, \dots, d_m$ . Значение каждой стратегии равно математическому ожиданию события «поощрение / штраф» в случае ее выбора;

2) в рамках заданных стратегий функционируют  $n$  автоматов:  $n < m$ . Выигрыш от стратегии  $d_i, i=1, \dots, m$ , в любой момент времени  $t$  поровну делится между всеми выбравшими эту стратегию автоматами;

3) требуется найти условия, обеспечивающие максимальный выигрыш для всей совокупности автоматов.

Эти условия делятся на три основные группы:

– типы исследуемых автоматов, определяющие структуру и особенности их поведения;

- значение глубины памяти  $q$ , определяющее степень инерционности автомата (скорость переключения между действиями);
- структурные изменения автоматов, направленные на организацию различных способов взаимодействия в коллективе.

Мы будем анализировать рассмотренные выше типы автоматов: автомат с линейной тактикой, автомат Кринского, автомат Роббинса и автомат Крылова.

Мы также проанализируем влияние глубины памяти  $q$  на степень инерционности автоматов, т.е. на скорость переключения автоматов между действиями в среде  $E$ .

Взаимодействие автоматов варьируется от полной изолированности автоматов друг от друга до введения внешней по отношению к структуре автоматов процедуры общей кассы.

#### ОЦЕНКА СКОРОСТИ ПЕРЕКЛЮЧЕНИЯ

Скорость переключения действий автомата определяет минимальное время, за которое автомат способен использовать все доступные стратегии.

Требуется определить необходимый интервал времени функционирования автомата для сбора достоверной статистики.

**Утверждение 1.** Средняя скорость переключения между действиями одного из изолированных автоматов в игре в размещении не ниже значения этой скорости автомата, функционирующего в соответствующей стационарной среде.

Действительно, если у автомата при выборе текущего действия  $d^* = d_i$ ,  $i = 1, \dots, m$  в текущей партии отсутствуют соседи, то его новое состояние будет обусловлено только значением отклика окружающей среды на выбор данного действия в соответствии с вероятностью получения поощрения  $p^* = p_i$ . В этом случае его скорость смены действий будет совпадать со скоростью автомата в соответствующей стационарной среде  $E$ .

Если при выборе данной стратегии  $d^* = d_i$  у рассматриваемого автомата имеются соседи (один или более одного) на  $i$ -й площадке, то текущая ситуация распадается на два случая.

Если сигнал поощрения среды достался рассматриваемому автомату, то в текущей партии факт наличия соседей не влияет на его скорость переключения между действиями.

Если, наоборот, при получении от среды сигнала поощрения в рамках дележа пищи рассматриваемому автомату достается штраф, то такой сигнал может лишь уменьшить (по крайней мере не увеличить) время его пребывания на текущей площадке.

Таким образом, средняя скорость переключения между действиями изолированного автомата в игре в размещении не меньше, чем скорость переключения автомата в соответствующей стационарной среде  $E$ .

Теперь рассмотрим фактор влияния процедуры общей кассы.

**Утверждение 2.** Средняя скорость переключения  $n$  автоматов в наилучшем размещении не ниже скорости переключения между действиями одним автоматом, выполняющим наилучшее действие в соответствующей стационарной среде  $E$ .

Действительно, наличие процедуры общей кассы приводит к тому, что вся совокупность полученных от окружающей среды поощрений в любой партии делится в равных долях между всеми  $n$  автоматами.

Технически это может быть реализовано, например, путем последовательной раздачи  $n$  автоматам всех поощрений, получаемых от среды за полное время игры, в порядке циклической очереди.

Очевидно, что для коллектива из  $n$  автоматов наилучшим является такое размещение, при котором все автоматы по одному занимают  $n$  наилучших площадок из  $m$  возможных,  $n < m$ . При таком размещении, благодаря процедуре общей кассы, каждый из  $n$  автоматов получает поощрение с вероятностью, равной среднеарифметическому суммарной вероятностей  $n$  наилучших площадок.

Эта величина меньше (не больше) вероятности поощрения, получаемого одним автоматом, расположенным на наилучшей площадке.

Следовательно, среднее время пребывания коллектива автоматов в наилучшем для себя размещении не меньше времени пребывания одного автомата на наилучшей площадке.

А суммарное время последовательного пребывания одного автомата на всех площадках тем более не превосходит времени пребывания коллектива автоматов в более худших размещениях, где средняя скорость переключения между действиями еще выше.

Поочередно вычеркивая из списка наилучшую площадку, мы можем провести данное доказательство для всех  $m$  площадок.

#### ПРОВЕДЕНИЕ ВЫЧИСЛИТЕЛЬНОГО ЭКСПЕРИМЕНТА

Для проведения игры в размещения задано 10 площадок. В качестве доступных для автоматов действий.

Величина поощрения равна (+1), величина штрафа (−1).

Эти величины определяют следующие значения вероятностей получения поощрений:

**Таблица 1.** Платежная матрица

**Table 1.** Payoff matrix

Цена	0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
Поощрения	0,50	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95

Для каждого типа автомата была проведена серия из 10 независимых экспериментов, после чего полученные данные усреднялись.

Для каждой серии экспериментов генератор случайных чисел запускался с фиксированного начального значения, порождая для всех типов автоматов идентичную псевдослучайную стационарную среду  $E$ .

Обозначим типы автоматов следующим образом:

автомат с линейной тактикой – АЛТ;

автомат Кринского – АКРН;

автомат Роббинса – АРББ;

автомат Крылова – АКРЛ.

**Таблица 2.** Математические ожидания поощрений и штрафов

**Table 2.** Mathematical expectations of rewards and penalties

Автомат	Память 1	Память 2	Память 3	Память 4	Память 5	Память 6	Память 7	Память 8
АЛТ	0,63679	0,76809	0,80862	0,87499	0,88372	0,88475	0,79856	0,88803
АКРН	0,63679	0,76809	0,80474	0,87788	0,87707	0,88170	0,88255	0,87800
АРББ	0,63679	0,75081	0,80495	0,87277	0,87602	0,89240	0,89255	0,89688
АКРЛ	0,630618	0,760276	0,848816	0,884885	0,886225	0,89742	0,89549	–

Мы видим, что даже в рамках вероятностной стационарной среды конструкция автомата не является абсолютным преимуществом при всех значениях глубины памяти  $q$ .

При этом общая тенденция увеличения эффективности функционирования автомата с ростом глубины памяти  $q$  сохраняется.

Считая значение математического ожидания каждого автомата как результат выполнения некоторого алгоритма, воспользуемся величиной среднего времени работы автомата как оценкой быстродействия этого алгоритма для достижения полученного результата.

**Таблица 3.** Среднее время работы автомата

**Table 3.** Average machine operating time

Автомат	Память 1	Память 2	Память 3	Память 4	Память 5	Память 6	Память 7	Память 8
АЛТ	74	695	3761	251394	3648854	19393688	927170104	22414244336
АКРН	74	695	4527	264297	3795658	–	48562927	927170104
АРББ	74	759	5179	264297	3795658	53461387	1029193271	26407904074
АКРЛ	116	2152	50814	2426164	89700530	3627473535	126294097025	–

Таким образом, мы видим две автоматные стратегии оптимизации поведения:

- 1) получение оптимального результата за счет более высокой скорости смены действий;
- 2) замедление скорости переключения действий для увеличения времени нахождения в рамках оптимального в среде  $E$  действия.

#### ЗАКЛЮЧЕНИЕ

Рассмотренные выше системы, состоящие из достаточно простых автоматов, обладают свойством целесообразного поведения. Целесообразное поведение таких систем наблюдается уже при небольшом количестве входящих в них автоматов. Разнообразие поведения также достигается уже при незначительных изменениях в способе обработки сигналов автоматами. Более того, во всех случаях, кроме стационарной случайной среды, чрезмерная емкость памяти  $q$  оказывалась даже вредной. Внешняя процедура общей кассы в игре в размещения может быть реализована при помощи различных способов взаимодействия между участвующими в ней автоматами, которое варьируется от информационного до материального (полученные от среды поощрения).

#### СПИСОК ЛИТЕРАТУРЫ

1. Стефанюк В. Л. Локальная организация интеллектуальных систем. М.: ФИЗМАТЛИТ, 2004. 328 с.
2. Димитриченко Д. П. Оптимизация рекуррентной нейронной сети при помощи автоматов с переменной структурой // Программные системы и вычислительные методы. 2023. № 4. С. 30–43. DOI: 10.7256/2454-0714.2023.4.69011
3. Поспелов Д. А. Игры и автоматы. М.: Энергия, 1966. 136 с.
4. Цетлин М. Л. Исследования по теории автоматов и моделированию биологических систем. М. 1969. 316 с.
5. Поспелов Д. А. Вероятностные автоматы. М.: Энергия, 1970. 88 с.
6. Варшавский В. И. Коллективное поведение автоматов. М.: Наука, 1973. 408 с.
7. Варшавский В. И., Поспелов Д. А. Оркестр играет без дирижера: размышления об эволюции некоторых технических систем и управление ими. М.: Наука, 1984. 208 с.

## REFERENCES

1. Stefanyuk V.L. *Lokal'naya organizatsiya intellektual'nykh sistem* [Local organization of intelligent systems]. Moscow: FIZMATLIT, 2004. 328 p. (In Russian)
2. Dimitrichenko D.P. Optimization of a recurrent neural network using automata with variable structure. *Programmnyye sistemy i vychislitel'nyye metody* [Software systems and computational methods]. 2023. No. 4. Pp. 30–43. DOI: 10.7256/2454-0714.2023.4.69011. (In Russian)
3. Pospelov D.A. *Igry i avtomaty* [Games and automata]. Moscow: Energiya, 1966. 136 p. (In Russian)
4. Tsetlin M.L. *Issledovaniya po teorii avtomatov i modelirovaniyu biologicheskikh sistem* [Studies in the theory of automata and modeling of biological systems]. Moscow, 1969. 316 p. (In Russian)
5. Pospelov D.A. *Veroyatnostnyye avtomaty* [Probabilistic automata]. Moscow: Energiya, 1970. 88 p. (In Russian)
6. Varshavskii V.I. *Kollektivnoye povedeniye avtomatov* [Collective behavior of automata]. Moscow: Nauka, 1973. 408 p. (In Russian)
7. Varshavskii V.I., Pospelov D.A. *Orkestr igrayet bez dirizhera: razmyshleniya ob evolyutsii nekotorykh tekhnicheskikh sistem i upravleniye imi* [The orchestra plays without a conductor: reflections on the evolution of some technical systems and their control]. Moscow: Nauka, 1984. 208 p. (In Russian)

**Финансирование.** Исследование проведено без спонсорской поддержки.

**Funding.** The study was performed without external funding.

#### Информация об авторе

**Димитриченко Дмитрий Петрович**, канд. техн. наук, ст. науч. сотр., отдел нейроинформатики и машинного обучения, Институт прикладной математики и автоматизации – филиал Кабардино-Балкарского научного центра Российской академии наук;  
360000, Россия, г. Нальчик, ул. Шортанова, 89 А;  
dimdp@rambler.ru, ORCID: <https://orcid.org/0000-0003-2399-3538>, SPIN-код: 3272-3520

#### Information about the author

**Dmitry P. Dimitrichenko**, Candidate of Technical Sciences, Senior Researcher, the Department of Neuroinformatics and Machine Learning, Institute of Applied Mathematics and Automation – branch of Kabardino-Balkarian Scientific Center of the Russian Academy of Sciences;  
360000, Russia, Nalchik, 89 A Shortanov street;  
dimdp@rambler.ru, ORCID: <https://orcid.org/0000-0003-2399-3538>, SPIN-code: 3272-3520