

Программные системы и вычислительные методы

Правильная ссылка на статью:

Димитриченко Д.П. Анализ целесообразного поведения различных типов автоматов в условиях игры в размещения // Программные системы и вычислительные методы. 2024. № 4. DOI: 10.7256/2454-0714.2024.4.72488 EDN: SQCVFI URL: https://nbpublish.com/library_read_article.php?id=72488

Анализ целесообразного поведения различных типов автоматов в условиях игры в размещения

Димитриченко Дмитрий Петрович

кандидат технических наук

научный сотрудник, Институт прикладной математики и автоматизации

360000, Россия, г. Нальчик, ул. Шортанова, 89а

✉ dimdp@rambler.ru



[Статья из рубрики "Теория, программное обеспечение и языки параллельных вычислений"](#)

DOI:

10.7256/2454-0714.2024.4.72488

EDN:

SQCVFI

Дата направления статьи в редакцию:

27-11-2024

Дата публикации:

05-12-2024

Аннотация: Объектом исследования настоящей работы являются однородные коллективы автоматов, обладающие свойством целенаправленного поведения. Предметом настоящего исследования является сравнение различных конструкций таких автоматов при реализации условий игры в размещения. Целью настоящего исследования является количественная оценка влияния инерционных свойств автоматов и значений глубины памяти на эффективность функционирования коллектива автоматов в рамках игры в размещения. В рассматриваемых коллективах автоматы выполняют действия в заданной среде обитания с различной степенью результативности. Автоматы, в соответствии со своей конструкцией, реагируют на поданный входной сигнал очередным действием. Оценка эффективности автомата определяется, как сумма положительных сигналов, или отрицательных сигналов, полученных автоматом за

рассматриваемый промежуток времени. Указанная характеристика зависит, как от заявленной конструкции автомата, так и от глубины его памяти. Требуется определить наиболее простые конструкции автоматов, позволяющие достигать оптимума эффективности в заданной окружающей среде наискорейшим образом. Формализация, как свойство окружающей среды, так и действий автоматов, а также обработка полученных результатов производится при помощи аппарата теории игр. В этом случае значения эффективности функционирования автоматов представляются, как совокупные суммы выигрышей и проигрышей игроков-автоматов. Новизной проведенного исследования является построение интегральной оценки эффективности поведения коллектива автоматов, позволяющей проследить не только стремление коллектива к максимизации суммы поощрений, но и к минимизации внутриколлективной конкуренции. Полученный результат позволяет проследить влияние инерционных качеств автоматов, реализованных в виде соответствующих конструкций, на эффективность функционирования в заданной окружающей среде, формализованной в виде игры в размещения. Автомат с линейной тактикой и автомат Крылова образуют две предельные реализации автоматной стратегии приближения к оптимуму. Первые за счет высокой скорости смены действий, вторые за счет длительного пребывания в состояниях, близких к оптимальным. Областью применения полученных результатов является дальнейшее исследование более сложных динамических сред при помощи наиболее простых конструкций автоматов, так как синхронные коллективы автоматов в процессе вычислительной реализации плохо поддаются распараллеливанию, что приводит к значительному росту временных и вычислительных затрат при усложнении структуры динамических сред или при увеличении данных оптимизационных задач.

Ключевые слова:

однородный коллектив автоматов, целесообразное поведение, поощрения, штрафы, обучение с подкреплением, автомат, глубина памяти, теория игр, оптимальная стратегия, игра в размещения

Введение

Организация сложного поведения децентрализованных систем, построенных на базе простых элементов находит свое отражение в таком разделе машинного обучения, как «Коллективное поведение автоматов» [\[1\]](#), имеющее широкое прикладное значение от решения задач автоматной оптимизации, интеллектуального управления, построение робототехнических систем [\[2\]](#), и до вопросов психологии [\[3\]](#).

При этом, связь между нейросетевым и автоматным подходами является достаточно тесной [\[4, 5\]](#). В связи с ростом вычислительных возможностей многопользовательских и многопоточных систем возникает объективная потребность в анализе и управлении [\[6, 7\]](#) совокупностью агентов действующих децентрализованно, в соответствии с собственным целеполаганием и взаимодействующих друг с другом и заданной окружающей средой.

При этом часто возникают ситуации, когда одно и тоже действие, выбранное несколькими агентами, может быть физически выполнено в конкретный момент времени только одним из них.

Например, только один из потоков, обрабатывающих актуальные данные, может производить операцию чтения (или записи). Всем остальным претендентам доступ к этим

данным блокируется.

Только один робот на перекрестке, из двух подъехавших по разным путям, может пересекать противоположную трассу. Оба робота не могут одновременно двигаться (будет столкновение), или одновременно ожидать прохождения другого робота, тем самым создав ситуацию взаимоблокировки прохождения перекрестка на неопределенный срок не только текущими, но и следующими за ними роботами.

В биологической интерпретации, совместно выбираемым действиям агентов соотносятся, разделяемые несколькими животными, определенные участки экологической ниши. Тогда количество актуального ресурса на участке делится (не обязательно поровну) между всеми, присутствующими на нем животными.

Аналогичная ситуация возникает при активном использовании интернет-трафика несколькими приложениями.

Это простейшие случаи, иллюстрирующие правила поведения в рамках игры в размещения, когда действие, выбранное несколькими агентами, в отдельно взятый момент времени может выполнить (в нашем случае получить поощрение) только один из них (в нашем случае один из автоматов, выбравших это действие).

Все выше перечисленное позволяет привлечь к анализу поведения мультиагентных систем (конечного множества агентов в искусственной среде) теорию вероятности, теорию игр, теорию автоматов для получения теоретически обоснованных и подтверждаемых методами имитационного моделирования результатов.

Примененный в настоящей работе автоматный подход предоставляет в распоряжение исследователя формализованный (в рамках дискретной математики) метод, позволяющий выполнить формальную постановку задачи и проанализировать поведение агентов и самой заданной среды обитания в терминах автоматной модели (коллектива автоматов), входных, выходных и внутренних алфавитов, а также правил перехода и результатов выбора.

Целью настоящей работы является получение количественных оценок влияния инерционных свойств (различным образом проявляемых) определенных типов автоматов и различных значений глубины памяти на эффективность функционирования коллектива автоматов в стохастической среде, организованной по правилам игры в размещения.

Новизной проведенного исследования является следующее:

1. Обоснование (теоретическое доказательство соответствующего утверждения и его подтверждение при выполнении вычислительных экспериментов) о том, что в стохастической среде игры в размещения в качестве доверительного интервала времени функционирования коллектива автоматов достаточно принять время функционирования отдельного (эталонного) автомата. Это позволило минимизировать время проведения вычислительных экспериментов для исследуемых типов автоматов, что особенно актуально при росте значений глубины памяти и, как следствие, росте количества времени вычислений.
2. Такая оценка доверительного интервала оказывается инвариантной относительно заданного типа автоматов. При этом, чем больше коллектив автоматов, тем точнее оказывается полученная таким образом оценка актуального интервала времени.
3. Применение для оценки эффективности коллектива автоматов интегральной

величины математических ожиданий всех автоматов, которая в условиях не равнозначных по цене действий, является особенно информативной и позволяет точно интерпретировать приближение коллектива автоматов к статистически оптимальному состоянию, так как в силу стохастичности окружающей среды автоматы фактически всегда находятся в движении.

4. Удалось пронаблюдать, что полученная таким образом интегральная величина эффективности функционирования коллектива автоматов, не только количественно стремится к максимизации совокупности поощрений от среды, но и информирует о стремлении коллектива к минимизации уровня внутриколлективной конкуренции.

Одним из традиционных методов анализа поведения автоматов является применение аппарата теории игр [8]. Такой выбор обусловлен тем, что теория игр позволяет легко создавать формализуемые системы правил (среда) и стратегий поведения (доступные действия), определяющие характер взаимодействия отдельного автомата (или совокупности автоматов) и окружающей среды [9].

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 Михаилом Львовичем Цетлиным [9]. Он выступил создателем целого направления исследований, получившего название «коллективное поведение автоматов» [9-12]. 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 простейшей 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 , $t > 1$, 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 «к системе, осведомленной о наилучшей возможной стратегии».

Такое определение целесообразного поведения не зависит от структуры анализируемой системы, а опирается только на статистический 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 времени t .

Первая автоматная реализация такой системы 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: $st+1 = F(st, xt)$, и начальным внутренним состоянием в нулевой момент времени $t=t_0$: $s_0 = S(t_0)$.

Функцией зависимости выходных сигналов (действий) от внутренних состояний: $dt = G(st)$.

Общая постановка задачи

Пусть задана некоторая среда E , в которой будут функционировать автоматы исследуемых типов.

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.

Проводимый эксперимент состоял в следующем [\[9\]](#):

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 двух событий (ответвлений).

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

Оказалось, что, несмотря на ошибки выбора в начале серии экспериментов, животные в последствии 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 Т-образного лабиринта.

В Т-образном лабиринте было всего два доступных для выполнения подопытным животным 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 о исходных значениях вероятностей, но вынуждены эту информацию получать в опосредованной форме через сигналы поощрения и штрафа, получаемые от окружающей среды.

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 L_1 (с глубиной памяти $q=1$) состоит из m состояний $S=s_1, \dots, s_m$.

За каждым из m состояний si закреплено действие $di, i=1, \dots, m$.

Автомат $L1$, функционирует следующим образом:

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

В ответ на это действие окружающая среда и формирует сигнал, принимающий значение «Поощрение» в соответствии с вероятностью pi (или «Штраф» с вероятностью $1 - pi$).

Этот сигнал подается на вход автомата $L1$.

Если автомат $L1$ на входе получает сигнал «Штраф», то он переходит в состояние $si+1$ и, следовательно, переключает внешнее действие с di на $di+1$.

При получении сигнала «поощрение» автомат $L1$ остается в исходном состоянии si , 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 di q последовательных состояний, мы получим последовательность линейных автоматов с монотонно возрастающей глубиной памяти q .

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, выполняемому им, действию.

При получении штрафа от окружающей среды этот автомат реагирует точно также, 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, что будет проиллюстрировано при анализе

их функционирования.

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 также, как и автомат с линейной тактикой, увеличивая вплоть до самого глубокого номер состояния, соответствующего действию, за выполнение которого этот автомат получает поощрение от среды E .

При получении штрафа автомат Крылова «подбрасывает монетку».

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

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, совокупность из m действий, доступных автомату вместе со значениями вероятностей поощрений (штрафов) в стационарной среде E , на языке теории игр формулируется, как игра с природой:

- 1) задана одностолбцовая 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 (единственный 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 $p^*=p_i$, на входе получает сигнал поощрения (или штрафа) от окружающей среды I , $i=1, \dots, m$.

Поскольку оптимальная стратегия игры с природой известно 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 содержимое этой матрицы не известно.

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 (пастбищ, или охотничьих территорий) 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 конкретного типа автомата для сбора достоверной статистики.

Поскольку все автоматы образуют однородный коллектив, и выбор действия каждым из n автоматов является случайным, то, начиная с некоторого момента времени $t=t^*$ суммы поощрений и штрафов автоматов по всему коллективу примут статистически близкие значения. В этой связи в качестве индикатора такого состояния системы «коллектив автоматов» целесообразно принять коэффициент вариации сумм поощрений и штрафов по всему коллективу.

Очевидно, что если коэффициент вариации не увеличивается, то коллектив автоматов принял статистически равновесное (не улучшаемое) состояние.

Вычислительные эксперименты показали, что наилучший результат достигается тогда, когда время функционирования системы, состоящей из n автоматов, определяется временем функционирования однотипного эталонного автомата, не принадлежащего исследуемому коллективу, т. е. не подверженного влиянию коллектива, но функционирующего в той же среде.

В этой связи рассмотрим сравнение времени функционирования отдельного автомата и автомата в коллективе.

Утверждение. Средняя скорость переключения между действиями одного из изолированных автоматов в игре в размещения 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 с скоростью автомата в соответствующей стационарной среде E .

Если при выборе данной стратегии $d^*=di$ у рассматриваемого автомата имеются соседи (один, или более одного) на i -ой 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 достаётся 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 размещения не меньше, чем скорость переключения автомата в соответствующей стационарной среде E .

На практике скорость переключения автомата в коллективе выше скорости переключения эталонного автомата.

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:

Табл. 1. Платежная матрица

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

На первом этапе проведения серии вычислительных экспериментов потребовалось определить величины эффективности исследуемых типов автоматов и статистически достоверные интервалы времени для формирования этих величин.

В качестве опорной величины такого интервала времени было выбрано время (количество ходов) последовательного прохождения фиксированным автоматом с заданной глубиной памяти всех площадок (от наихудшей к наилучшей).

В качестве оценки эффективности функционирования заданного автомата целесообразно принять среднюю величину эффективности за ход, т. е. сумму поощрений и штрафов, полученных за время однократного прохождения автоматом всех площадок (выполнением всех доступных действий).

Фактически эта величина равна математическому ожиданию поощрений и штрафов автомата на всем множестве действий в случае однократного их выполнения.

Поскольку время пребывания автомата на конкретной площадке зависит от совокупного влияния конструкции автомата, текущей глубины памяти и последовательности штрафов и поощрений, обусловленной значениями генератора случайных чисел, то было принято решение построить серию из фиксированного числа независимых прохождений.

Это позволило:

1. Получить средние величины математических ожиданий рассматриваемых типов автоматов в соответствии со значениями глубины памяти.
2. Получить средние, статистически достоверные величины интервалов времени, требующиеся для выполнения автоматами всех доступных действий при условии, что в качестве минимальной единицы времени было выбрано значение одного хода, выполняемого автоматом.
3. Проверить корректность обработки результатов экспериментов при помощи соответствующей величины коэффициента вариации значений средней эффективности по всей серии экспериментов для каждого из четырех типов автоматов, которая не только не превышала 10 процентов, но с ростом глубины памяти вне зависимости от типа автомата неуклонно уменьшалась.

Для проведения серии вычислительных экспериментов потребовалось создание специализированного программного обеспечения, реализующего алгоритмы поведения исследуемых типов автоматов и окружающей среды по правилам игры в размещения, позволяющего задавать соответствующие типы автоматов, их характеристики (значение глубины памяти) и различные параметры окружающей среды, коллективы автоматов, как совокупности независимых агентов.

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.

Табл. 2. Математические ожидания поощрений и штрафов

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	_____

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.

Табл. 3. Среднее время работы автомата

Автомат	Память 1	Память 2	Память 3	Память 4	Память 5	Память 6	Память 7	Г
АЛТ	74	695	3761	251394	3648854	19393688	927170104	224
АКРН	74	695	4527	264297	3795658		48562927	927
АРББ	74	759	5179	264297	3795658	53461387	1029193271	264
АКРЛ	116	2152	50814	2426164	89700530	3627473535	126294097025	—

Для анализа поведения автоматов в игре в размещения были выбраны коллективы, состоящие из 4 автоматов, а в качестве критерия оптимальности поведения была выбрана, как наиболее гибкая, сумма поощрений и штрафов по всему коллективу.

Дополнительно, такая величина оптимальности показывает среднестатистическое расположение автоматов по площадкам, т. Е. с ростом памяти автоматы стремятся разместиться на наилучших площадках, максимизируя суммы поощрений и минимизируя фактор конкуренции в коллективе, как отрицательный фактор.

Табл. 4. Интегральные оценки эффективности функционирования в игре в размещения.

Память	АЛТ	АКРН	АРББ	АКРЛ
1	1,4865	1,4865	1,4865	1,5135
2	2,2266	2,2266	2,1267	2,4860
3	2,6450	2,6662	2,5837	2,6814
4	2,8363	2,8012	2,7534	2,8846
5	2,9152	2,8955	2,8707	2,9365
6	2,9541	2,9331	2,9195	2,9631
7	2,9724	2,9561	2,9500	—

Из предыдущей таблицы мы можем видеть среднее время (количество ходов), требуемое для получения указанного результата коллективом, так как в качестве времени функционирования коллектива выступает однотипный эталонный автомат.

Автомат Крылова [\[10\]](#) располагает ярко выраженным механизмом инерционности поведения, обеспечивающим, как длительное (по сравнению с другими автоматами) пребывание в наилучшем состоянии, так и наиболее медленное приближение к наилучшему размещению коллектива в целом.

Заключение

Рассмотренные выше системы (коллективы), 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 способах обработки входных сигналов от среды.

Нами были рассмотрены 4 типа автоматов.

Выявлены две автоматные стратегии:

1. За счет высокой скорости переключения посетить все доступные площадки (автомат с линейной тактикой);

2. за счет высокой инерционности длительное время оставаться в наилучшем состоянии (автомат Крылова).

Однако, второй результат достигается существенно длительным периодом времени для перехода коллектива в оптимальное статистически устойчивое состояние, что при сравнительно небольших значениях глубины памяти приводит к значительным вычислительным затратам, что иллюстрирует таблица средних периодов времени функционирования эталонных автоматов.

Суммы поощрений и штрафов численно равны суммам математических ожиданий наилучших площадок, занимаемых автоматами, следовательно, дальнейшее усложнение конструкций автоматов, т. е. рост глубины памяти не может привести к значительному увеличению эффективности функционирования коллектива в целом, так как разница между возможным наилучшим значением и достигнутым результатом составляет в среднем не более 10- 20 процентов от возможного при значительном увеличении времени (модельного и реального) функционирования.

Полученные результаты могут быть положены в основу формирования оптимальных по сложности и вычислительным затратам коллективов автоматов для решения более сложных оптимизационных задач.

К таким задачам относятся:

1. Организация различных форм информационного взаимодействия между отдельными автоматами с целью повышения эффективности функционирования коллектива в целом. Такое влияние характеризуется достижением заданного (уже известного) уровня эффективности функционирования коллектива автоматов с меньшим требуемым значением глубины памяти, приводящим к более низким вычислительным и временным затратам при проведении вычислительных экспериментов, так как коллективы синхронно функционирующих автоматов плохо поддаются процессу распараллеливания общего потока вычислений. Фактор временных затрат особенно ощущается, например, при исследовании поведения автомата Крылова уже при сравнительно небольших значениях глубины памяти;
2. внесение в каждый из рассмотренных типов автоматов структурных изменений, приводящих коллектив автоматов к формированию процедуры общей кассы, как внутреннего свойства коллектива без введения такой внешней централизованной процедуры;
3. оценка влияния вносимых изменений на эффективность функционирования каждого из рассмотренных типов автоматов;
4. формирование автоматной реализации роевых (муравьиных) алгоритмов, характеризующихся согласованным движением всего коллектива к нахождению экстремума целевой функции в условиях отсутствия централизованного управления согласованными действиями отдельных агентов (автоматов).

Библиография

1. Стефанюк В.Л. Локальная организация интеллектуальных систем. М.: Физматлит. 2004. С. 328.
2. Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота. Модели поведения. М.: Ленанд. 2019. С. 304.
3. Журавлев А.Л., Савченко Т.Н., Головина Г.М. Математическая психология: школа В. Ю. Крылова-Сер. Научные школы Института психологии РАН. М.: Изд-во ИП РАН. 2010. С.

512.

4. Жданов А.А. Автономный искусственный интеллект. М.: Лаборатория знаний. 2024. С. 362.

5. 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 EDN: FEIPTC URL: https://e-notabene.ru/itmag/article_69011.html

6. Карпов В.Э., Карпова И.П., Кулинич А.А. Социальные сообщества роботов. М.: УРСС. 2019. С. 352.

7. Карпов В.Э., Королева М.Н. К вопросу о формализации этики поведения коллаборативного робота // Информационные и математические технологии в науке и управлении. 2022. № 4 (28). С. 223-233.

8. Поспелов Д.А. Игры и автоматы. М.: Энергия. 1966. С. 136.

9. Цетлин М.Л. Исследования по теории автоматов и моделированию биологических систем. М.: Наука. 1969. С. 316.

10. Поспелов Д.А. Вероятностные автоматы. М.: Энергия. 1970. С. 88.

11. Варшавский В.И. Коллективное поведение автоматов. М.: Наука. 1973. С. 408.

12. Варшавский В.И., Поспелов Д.А. Оркестр играет без дирижера: размышления об эволюции некоторых технических систем и управление ими. М.: Наука. 1984. С. 208

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Рецензируемая статья посвящена сравнительному анализу целесообразного коллективного поведения четырех типов автоматов в условиях игры в размещения.

Методология выполненной работы базируется на использовании подходов и методов из теорий игр, массового обслуживания и машинного обучения, сопряжена с применением автоматных моделей и моделированием коллективного поведения при помощи автоматов, которые получили название игры в размещения, а также на описании проведения вычислительного эксперимента.

Актуальность работы определяется тем, что организация сложного поведения децентрализованных систем находит свое отражение в таком разделе машинного обучения, как «Коллективное поведение автоматов», имеющем прикладное значение в решении задач автоматной оптимизации, интеллектуального управления, построении робототехнических систем, а также в математической психологии.

Научная новизна рецензируемого исследования, к сожалению, авторами не сформулирована.

Структурно в статье выделены следующие разделы: Введение, Автоматная реализация; Общая постановка задачи, Биологические предпосылки, Типы конструкций автоматов, Динамические среды, Игры двух автоматов, Игра в размещения, Оценка скорости переключения, Проведение вычислительного эксперимента, Заключение и Библиография.

Авторами приведены описания автоматов Цетлина, Кринского, Роббинса, Крылова. Эти 4 типа автоматов рассмотрены и при проведении вычислительного эксперимента. Были выявлены две автоматные стратегии: во-первых, за счет высокой скорости переключения посетить все доступные площадки (автомат с линейной тактикой); во-вторых, за счет высокой инерционности длительное время оставаться в наилучшем состоянии (автомат Крылова). В публикации по итогам вычислительного эксперимента

сделан вывод о том, что автомат Крылова располагает ярко выраженным механизмом инерционности поведения, обеспечивающим, как длительное (по сравнению с другими автоматами) пребывание в наилучшем состоянии, так и наиболее медленное приближение к наилучшему размещению коллектива в целом.

Библиографический список включает 10 источников – научные публикации отечественных авторов по рассматриваемой теме на русском языке, опубликованные за период с 1966 по 2023 годы. В тексте публикации имеются адресные отсылки к списку литературы, подтверждающие наличие апелляции к оппонентам.

Из резервов улучшения работы, следует указать следующие. Во-первых, в вводной части публикации уместно в более доступной для широкого круга читателей форме сформулировать актуальность рассматриваемых вопросов и сферу их применения. Во-вторых, после ознакомления со статьей не ясно, каким образом, с использованием каких инструментальных средств были получены значения времени работы автоматов, на основе которых определены средние величины. В-третьих, в публикации не сформулирована цель работы и не отражены полученные авторами элементы приращения научного знания, научная новизна работы не раскрыта. В-четвертых, при описании практического применения полученных результатов исследования не стоит ограничиваться общей фразой: «Полученные результаты могут быть положены в основу формирования оптимальных по сложности и вычислительным затратам коллективов автоматов для решения более сложных оптимизационных задач», а перечислить хотя бы некоторые такие задачи.

Рецензируемый материал соответствует направлению журнала «Программные системы и вычислительные методы», отражает результаты проведенного авторского исследования, может вызвать интерес у читателей, но материал нуждается в доработке в соответствии с высказанными замечаниями.

Результаты процедуры повторного рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Представленная статья на тему «Анализ целесообразного поведения различных типов автоматов в условиях игры в размещения» соответствует тематике журнала «ПРОГРАММНЫЕ СИСТЕМЫ И ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ» и посвящена актуальному вопросу, связанного с ростом вычислительных возможностей многопользовательских и многопоточных систем возникает объективная потребность в анализе и управлении совокупностью агентов действующих децентрализованно, в соответствии с собственным целеполаганием и взаимодействующих друг с другом и заданной окружающей средой.

В статье представлен широкий анализ литературных российских источников по вопросам применения автоматного подхода, предоставляющего в распоряжение исследователя формализованного (в рамках дискретной математики) метода, позволяющего выполнить формальную постановку задачи и проанализировать поведение агентов и самой заданной среды обитания в терминах автоматной модели (коллектива автоматов), входных, выходных и внутренних алфавитов, а также правил перехода и результатов выбора.

Авторами статьи самостоятельно проведен комплексный анализ систем (коллективов), состоящих из достаточно простых автоматов, обладающих свойством целесообразного поведения. Авторами были рассмотрены 4 типа автоматов и выявлены две автоматные стратегии: за счет высокой скорости переключения посетить все доступные площадки (автомат с линейной тактикой); за счет высокой инерционности длительное время

оставаться в наилучшем состоянии (автомат Крылова).

Стиль и язык изложения материала является достаточно доступным для широкого круга читателей. Статья по объему соответствует рекомендуемому объему от 12 000 знаков.

Статья достаточно структурирована - в наличии введение, заключение, внутреннее членение основной части (авторами рассмотрены вопросы: Автоматная реализация, Биологические предпосылки, Стационарная окружающая среда, Типы конструкций автоматов, Динамические среды и др.).

Авторами проведена серия вычислительных экспериментов для которых потребовалось создание специализированного программного обеспечения, реализующего алгоритмы поведения исследуемых типов автоматов и окружающей среды по правилам игры в размещения, позволяющего задавать соответствующие типы автоматов, их характеристики (значение глубины памяти) и различные параметры окружающей среды, коллективы автоматов, как совокупности независимых агентов. Для каждого типа автомата авторами была проведена серия из десяти 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.

Практическая значимость четко обоснована.

К недостаткам можно отнести следующие моменты: отсутствует четкое выделение предмета, объекта.

Рекомендуется четко сформулировать предмет, объект. Также будет целесообразным добавить о перспективах дальнейшего исследования.

Статья «Анализ целесообразного поведения различных типов автоматов в условиях игры в размещения» требует доработки по указанным выше замечаниям. После внесения поправок рекомендуется к повторному рассмотрению редакцией рецензируемого научного журнала.