*Федеральное государственное бюджетное образовательное учреждение* *высшего профессионального образования*

|  |  |
| --- | --- |
| **Gerb-BMSTU_01** | ***«Московский государственный технический университет имени Н.Э. Баумана***  ***(национальный исследовательский университет)»***  ***(МГТУ им. Н.Э. Баумана)*** |

ФАКУЛЬТЕТ Информатика и системы управления

КАФЕДРА Компьютерные системы и сети (ИУ6)

**РАСЧЁТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА**

**к научно-исследовательской работе на тему:**

«Анализ генетических алгоритмы с использованием цифровых автоматов»

Студент ИУ6-75  **\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ В.Д. Шульман**

(группа) (Подпись, дата) (И.О.Фамилия)

Руководитель работы **О.Ю. Еремин**

(Подпись, дата) (И.О.Фамилия)

Оглавление

[Реферат 3](#_Toc22198945)

[Определения, обозначения и сокращения 4](#_Toc22198946)

[Введение 6](#_Toc22198947)

[1. Актуальность НИР, цели и задачи 7](#_Toc22198948)

[1.1. Актуальность генетических алгоритмов 7](#_Toc22198949)

[1.2. Цели и решаемые задачи 8](#_Toc22198950)

[2. Сравнение генетических алгоритмов с детерминированными алгоритмами 8](#_Toc22198951)

[3. Классификация генетических алгоритмов 9](#_Toc22198952)

[3.1. Стационарные генетические алгоритмы 10](#_Toc22198953)

[3.2. Динамические генетические алгоритмы 10](#_Toc22198954)

[3.3. Поколенческие генетические алгоритмы 10](#_Toc22198955)

[3.4. Адаптивные генетические алгоритмы 11](#_Toc22198956)

[3.5. Многоуровневые генетические алгоритмы 11](#_Toc22198957)

[3.6. Параллельные генетические алгоритмы 11](#_Toc22198958)

[4. Анализ предметной области 12](#_Toc22198959)

[4.1. Основные объекты предметной области 13](#_Toc22198960)

[4.1.1. Агент 13](#_Toc22198961)

[4.1.2. Хранилище генов 13](#_Toc22198962)

[4.2. Основные информационные структуры предметной области 14](#_Toc22198963)

[4.2.1. Генетический код 14](#_Toc22198964)

[4.2.2. Генетическое древо 14](#_Toc22198965)

[4.2.3. Поле визуализации 14](#_Toc22198966)

[4.3. Основные процедуры предметной области 15](#_Toc22198967)

[4.3.1. Генетический алгоритм 15](#_Toc22198968)

[4.3.2. Метод отбора 15](#_Toc22198969)

[4.3.3. Сеанс 15](#_Toc22198970)

[5. Необходимые информационные технологии 16](#_Toc22198971)

[5.1. Централизованная обработка информации 16](#_Toc22198972)

[5.2. Децентрализованная обработка информации 16](#_Toc22198973)

[5.3. Информационная технология поддержки принятия решений 17](#_Toc22198974)

Реферат

Записка с. ?, рис. ?, табл. ?, источников ?, прил. А, Б, В, Г

ГЕНЕТИЧЕСКИЙ АЛГОРИТМ, СЕЛЕКЦИОННЫЙ МЕТОД, ОТБОР, АГЕНТ, ХРАНИЛИЩЕ ГЕНОВ.

Рассматриваемыми объектами в данной исследовательской работы являются генетические алгоритмы и методы отбора при решении прикладных задач, где в основе лежит использование цифровых автоматов.

Основной целью исследования в данной области является поиск и нахождения примеров эффективного применения цифровых автоматов в генетических алгоритмах, а также формулирование основных принципов при построении моделей и использованием цифровых автоматов. В работе рассматриваются различные способы представления цифрового автомата и также подходы к его модификации в процессе работы генетического алгоритма. Также в данном исследовании рассматриваются различные подходы к выработке метода отбора и выявлении оптимальны параметров для него.

Определения, обозначения и сокращения

Генетический алгоритм – алгоритм, осуществляющий поиск решения задачи оптимизации или моделирования путем случайного комбинирования параметров и отбора лучших решений.

Селекция – процесс, в рамках которого производится отбор по определенным критерия наиболее оптимального решения.

Селекционный метод – алгоритм, согласно которому производится оценка решений и формирование нового поколения

Агент – элементарная эволюционная единица, осуществляющая решение поставленной перед ней задачи

Популяция – множество агентов

Поколение – текущая популяция агентов, сформированная путем на основании предыдущих поколений, прошедших через селекционный отбор.

Жизненный цикл – период времени от появления агента в системе до его исключения из системы в результате работы селекционного алгоритма.

Генетический код – набор параметров и инструкций, которые выполнят агент в процессе своего жизненного цикла.

Мутация – случайное изменение генетического кода агента, приводящее к приближению или отдалению агента от оптимального решения

Оптимальное решение – решение, которое по набору признаков является наиболее предпочтительным

ГА – генетический алгоритм

ПГА – параллельный генетический алгоритм

Эвристический алгоритм – алгоритм решения задачи, который не является гарантированно точным или оптимальным

Детерминированный алгоритм – алгоритмический процесс, который выдает уникальный и предопределенный результат для заданных входных данных

Цифровой автомат – математическая модель дискретного устройства, которое принимает и выдает сигналы, принимая различные состояния

Введение

На данный момент достаточно широкое распространение получили методы решения задач, основанные на использовании генетических алгоритмов. Эти методы в основном используются для решения задач, решение для которых либо слишком сложно найти аналитически, составив алгоритм, либо они являются настолько абстрактными, что решение для них в виде детерминированного алгоритма вовсе не существует. Генетические алгоритмы позволяют решать такие задачи.

Сама идея генетических алгоритмов не нова. Еще в 60-ых годах прошлого века начались первые симуляции эволюционных процессов на тогдашних маломощных огромных ЭВМ. В 70-ых годах была разработана концепция эволюционного программирования с использованием конечных автоматов. Достаточно большую популярность генетические алгоритмы обрели после экспериментов Джона Холланда с так называемыми клеточными автоматами

Однако, к сожалению, из-за скромных вычислительных мощностей длительно время генетические алгоритмы оставались в основном в теоретической области и почти не использовались для решения прикладных задач. Лишь к концу 80-годов с ростом вычислительны мощностей стало внедряться оборудование, которое включало в себя генетические алгоритмы для решения поставленных перед ними задач. В 89-ом году компания Axcelis выпустила свой первый продукт для персональных компьютеров, в котором применялись генетические алгоритмы. Носил он символичное название “Evolver”.

Сейчас существует много вариантов реализации генетических алгоритмов, каждый из которых больше подходит для решения определенных задач. Например, нейронные сети хорошо решают задачу распознавания изображений, но решить задачу самообучающегося ИИ нейронные сети так и не смогли, несмотря на то, что прошло несколько витков их популярности.

# Актуальность НИР, цели и задачи

Генетические алгоритмы (ГА) являются подклассами более общей группы методов, называемыми эволюционными алгоритмами, которые объединяют различные варианты использования эволюционных принципов. В отличии от других технологий оптимизации ГА содержат популяцию пробных решений, которые конкурентно управляются с помощью определенных операторов. Генетическим алгоритмам присуще итеративное обучение популяции индивидов.

Генетические алгоритмы относится к классу эвристических алгоритмов поиска, т.е. они решают задачи, используя практические методы, которые не являются гарантированно точными и оптимальными, однако достаточно точны для решений поставленной перед ними задачи.

## Актуальность генетических алгоритмов

На волне популярности искусственного интеллекта генетические алгоритмы являются поводом для постоянных дискуссий и исследований. Несмотря на то, что ИИ представляет из себя целый комплекс различных систем, такие как системы распознавания, анализа, обучения, адаптации и т.п., генетические алгоритмы вполне могу решить часть задач, связанных с адаптацией и обучением.

Генетические алгоритмы актуальны на данный момент тем, что уже сейчас способы решать определенный спектр вполне реальных задач. Генетические алгоритмы могут использоваться для управления процессорами, балансируя нагрузку в многопроцессорной системе, они могут использоваться для решения задач по поиску наилучшего соотношения веса/прочности/размера/плотности, для решения задач по размещению объектов различных форм на определенной площади (трассировка плат или нарезка ткани с наименьшими потерями, как пример).

Один из недавних коммерческих примеров: израильская компания Schema разработала программный продукт Channeling для оптимизации работы сотовой связи путем выбора оптимальной частоты, на которой будет вестись разговор. В основе этого программного продукта и используются генетические алгоритмы.

Генетические алгоритмы приобрели большую популярность на фоне новой волны подъема интереса к нейронным сетям. Сейчас довольно быстро развиваются технологии по распознаванию изображения и звука. Для обучения нейронных сетей также используется генетический алгоритм, где после каждого раунда нейронные сети сравниваются по своей эффективности, после чего часть их них отсеивается, а на основании оставшихся генерируется новое поколение нейронных сетей и измененными весами переходов.

Одно из больших достоинств генетических алгоритмов – это возможность распараллеливания вычислений при поиске решений. Т.к. агенты, которые пытаются решить задачу, чаще всего не должны взаимодействовать с друг-другом, то обработку каждого из них можно вести параллельно. Раньше это не давало больших преимуществ из-за одноядерных процессоров, но сейчас с развитием многопроцессорных архитектур и асинхронных языков программирования поиск решения с использованием генетических алгоритмов может проходить в разы быстрее.

## Цели и решаемые задачи

Основной целью НИР является решение задач по поиску оптимальных подходов к реализации генетических алгоритмов с использованием цифровых автоматов и определению круга задач, для которых использование такого вида генетических алгоритмов было бы оправдано и эффективно

# Сравнение генетических алгоритмов с детерминированными алгоритмами

Генетические алгоритмы применяются чаще всего для поиска не оптимального решения, а более лучшего по сравнению с имеющимся. Это обосновывается тем, что для сложной задачи чаще всего требуется найти не идеальное решение, а то решение. Которое удовлетворяло бы имеющимся требованиям. При этом достижения идеального «оптимального» решения отходит на второй план. Однако при этом другие методы, ориентированные на нахождение оптимального решения, из-за их чрезвычайно высокой сложности становятся и вовсе нереализуемыми.

Основные отличия генетических алгоритмов от традиционных методов:

Генетические алгоритмы работают с кодами, в которых представлен набор параметров, напрямую зависящих от аргументов целевой функции. При этом интерпретация этих кодов происходит только перед началом работы алгоритма и после завершения для получения результата. В процессе работы редактирования кодов проходит независимо от их трактования, код рассматривается просто как битовая строка.

Для поиска генетический алгоритм использует несколько точек поискового пространства одновременно, а не переходит от точки к точке, как это делается в традиционных методах. Это позволяет преодолеть один из их недостатков – опасность попадания в локальный экстремум целевой функции, если она не является унимодальной, т.е. имеет несколько таких экстремумов.

Генетические алгоритмы в процессе работы не используют никакую дополнительную информацию. Это повышает скорость работы. Единственной используемой информацией область допустимых значений параметров и целевой функции в произвольной точке.

Генетический алгоритм использует как вероятностный правила для порождения новых точек, так и детерминированные правила для перехода от одних точек к другим. Одновременное использование элементов случайности и детерминированности дает значительно большей эффект, чем раздельное.

# Классификация генетических алгоритмов

Генетические алгоритмы являются частным случаем еще более обобщенной сферы – генетического моделирования, но генетические алгоритмы достаточно разнообразны и обладают собственной классификацией.

## Стационарные генетические алгоритмы

Стационарные генетические алгоритмы подразумевают использовать подход, в котором популяция обновляется частями, а не вся сразу, Идея заключается в итеративном создании одного или двух потомков и добавления их непосредственно в популяцию, с предварительным исключением некоторых существующих агентов, чтобы освободить необходимое пространство.

Данный генетический алгоритм, по сравнению с поколенческим, использует примерно в два раза меньше памяти, т.к. всегда рассматривается только одна популяция. Также в этом алгоритме агенты-родители остаются в популяции на более длительное время, т. к. длительность их жизненного цикла обусловлена не «жесткими параметрами моделирования», а параметрами самого агента, которые формируются под воздействием генетического алгоритма случайным образов.

Такой тип алгоритмов более характерен для реального мира, поколения не имеют между собой четких временных границ, а сменяемость поколений происходит не резко, «между раундами», а плавно, в процессе конкуренции.

## Динамические генетические алгоритмы

Противоположен по своим идеям стационарному генетическому алгоритму. Здесь подразумевается использования итеративного подхода, где после каждой итерации популяция проходит полное обновление, в рамках которого удаляются и добавляются агенты.

## Поколенческие генетические алгоритмы

Поколенческие генетические алгоритмы подразумевают итеративное следованию определенной последовательности действий:

Создание первоначальной популяции.

Вычисление функций приспособленности для агентов популяции (оценивание)

Выборка агентов из текущей популяции (селекция)

Мутации агентов

Вычисление функций приспособленности для всех агентов

Формирование нового поколения

Оценить результаты и, если они не удовлетворительны, то повторить начиная с 3 пункта.

## Адаптивные генетические алгоритмы

Основная идея адаптивного генетического алгоритма заключается в том, что такой алгоритм должен уметь изменять свои параметры в процессе работы. В качестве параметров здесь выступает размер популяции, вероятность мутации, количество удаляемых и добавляемых агентов, а также множество других параметров, которые зависят от специфики самой задачи. Целью создания адаптивного алгоритма является ускорение поиска решения или уменьшения затрат ресурсов при его поиске (например, память).

## Многоуровневые генетические алгоритмы

Такие алгоритмы представляют из себя многоуровневую систему, где каждый уровень генетического алгоритма оптимизирует работу вышестоящего генетического алгоритма и так по цепочке. Говоря более формально, многоуровневый генетический алгоритм – это система, в которой генетические алгоритмы нижнего уровня (оптимизируемые генетические алгоритмы) оптимизируют генетические алгоритмы верхнего уровня.

Этот принцип часто используется в нейронных сетях, где вершины нейронной сети разбиты по слоям, представляющие собой уровни генетического алгоритма. При этом возникает очень сложная задача оптимизации, где необходимо вычислить оптимальное количество слоев для наиболее быстрого поиска решения задачи, т.к. никакой универсальной формулы, дающая ответ на этот вопрос, не существует.

## Параллельные генетические алгоритмы

Такие алгоритмы основаны на разбиении популяции на несколько отдельных подпопуляций, каждая из которых будет, независимо от других подпопуляций, обрабатываться генетическим алгоритмом. Одновременно с этим, разнообразные миграции агентов порождают обмен генетическим кодом среди популяций, которые, как правило, улучшают точность и эффективность алгоритма.

Выделяют три типа параллельных генетическим алгоритмов:

* + Глобальные однопопуляционные ПГА (master-slave)
  + Однопопуляционные ПГА
  + Многопопуляционные ПГА

Модель «хозяин-раб» характеризуется тем, что в алгоритмах такого типа селекция принимает во внимание целую популяцию, в отличии от двух других моделей.

В алгоритмах второго класса существует главная популяция, но оценка целевой функции распределена среди нескольких процессоров. Хозяин хранит популяцию, выполняет операции ГА и распределяет агентов между подчиненными. Они же лишь оценивают агентов. Однопопуляционные ГА пригодны для массовых параллельных компьютеров и состоят из одной популяции. Данный класс ПГА может быть эффективно реализован на параллельных компьютерах.

Третий класс - многопопуляционные ГА более сложная модель, так как она состоит из нескольких подпопуляций, которые периодически, по установленным правилам, обмениваются агентами. Такой обмен агентами называется миграцией и управляется несколькими параметрами. Такие ГА очень популярны, но достаточны сложны как для понимания, так и для реализации, потому что последствия от эффекта миграции, на данный момент, не полностью исследованы. В то же время многопопуляционные ГА имеют сходство с «островной моделью» в популяционной генетике, которая рассматривает относительно изолированные общины; поэтому параллельные ГА в некоторых случаях называют «островными» параллельными ГА.

# Анализ предметной области

Предметная область информационной системы, реализующая генетические алгоритмы, включает в себя ряд объектов, данных и функций, которые её описывают. Они представлены на рис. 1.

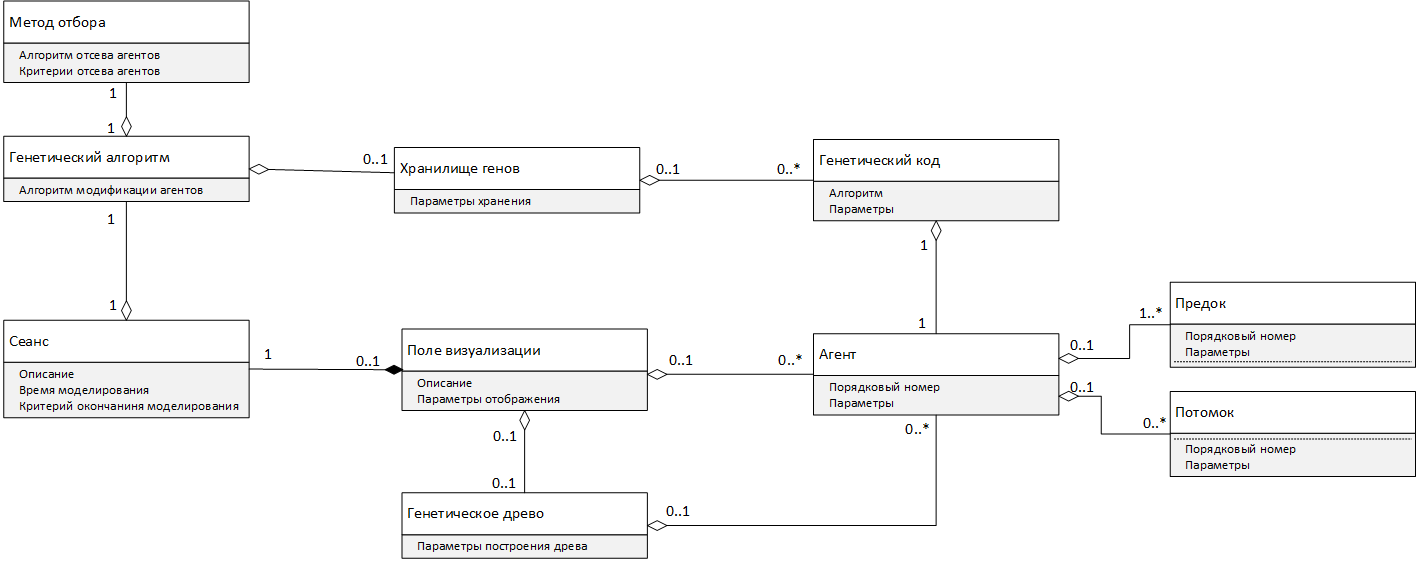


Рисунок – Концептуальная схема предметной области

## Основные объекты предметной области

Объекты представляют собой действующие лица, между которыми идет обмен информацией и которые осуществляют некоторые функции над друг-другом.

### Агент

Агент представляет из себя элементарную рабочую единицу ИС. Перед агентом ставятся задачи, для которой в системе идет поиск решения, как можно более приближенного к оптимальному. Существует множество агентов, параллельно решающих поставленную перед ними задачу. В начале процесса нахождения оптимального решения задачи все агенты находятся в схожих условиях и имеют идентичную конфигурацию. В течение некоторого отрезка времени в модели порождаются новые агенты с несколько отличной конфигурацией согласно генетическому алгоритму. Периодически с помощью метода отбора часть агентов удаляется из процесса поиска оптимального решения.

### Хранилище генов

Хранилище генов представляет из себя банк с данными. В этих данные содержится информацией о конфигурациях агентов. Хранилище генов содержит в себе информацию как о тех агентах, которые участвуют в текущем сеансе. Данные о конфигурации агентов из хранилища генов используются для формирования конфигурации новых агентов в процессе работы генетического алгоритма.

## Основные информационные структуры предметной области

Информационные структуры представляют из себя совокупность данных, которыми обмениваются объекты предметной области. Также данные используются функциями (алгоритмами) для получения результата, который может представлять из себя набор еще каких-то данных.

### Генетический код

Включает в себя набор параметров и инструкций, который использует агент в процессе своего жизненного цикла. Генетический код агента также можно определить, как его конфигурацию. Генетический код изменяется случайным образом в процессе работы генетического алгоритма. Генетический код наиболее агента, который наиболее эффективно решает поставленную перед ним задачу, является тем самым оптимальным решением, поиск которого производится в системе.

### Генетическое древо

Представляет из себя совокупность информации о родителях и потомках каждого агента, когда-либо появившегося в процессе сеанса поиска решения. Генетическое древо показывает какие ветви развития агентов имели какую численность в зависимости, и какая последовательность изменений привела к текущим генетическим кодам агентов

### Поле визуализации

Поле визуализации показывает текущее состояние сеанса работы генетического алгоритма.

Оно может в себя включать различные графические элементы:

Графики

Диаграммы

Графы состояний и переходов

Поля пространств, где размещаются и перемещаются агенты

Визуализации наполнения банков данных

Информация в поле призвана давать возможность объективно оценить текущие состояние сеанса и определить, был ли достигнут требуемый результат работы генетического алгоритма.

## Основные процедуры предметной области

Процедуры (алгоритмы), реализуемые в системе, использующие генетические алгоритмы, необходимы для обработки имеющихся данных и формирования новых.

### Генетический алгоритм

Наиболее важная часть ИС, определяющая всю логику ее работы и множество задач, которые эта система может решать.

Генетический алгоритм определяет каким образом будет происходить наследование генетического кода от агентов одного поколения к агентам последующего, а также какие случайные изменения будут происходить в генетическом коде новых агентов и с какой вероятностью.

### Метод отбора

Это алгоритм, который осуществляет отсев агентов в процессе моделирования, решения поставленной задачи которых оказалось наименее эффективно согласно критериям оптимальности решения этой задачи.

### Сеанс

Сеанс представляет из себя весь временной отрезок, в рамках которого агенты решают представленную перед ними задачу. Сеанс в зависимости от условий и специфики задач, стоящих перед ними, может проходить в различных формах:

Агенты решают поставленную задачу независимо, никак не контактируя и не взаимодействуя друг с другом

Агенты решают поставленную задачу в одной среде, весь сеанс поделен на периоды времени, представляющие из себя циклы, после каждого из которых происходит удаление и добавление агентов в зависимости от эффективности решения каждого.

Агенты решают задачу одновременно в одной среде соревнуясь друг с другом. При таком виде сеанса отбор осуществляется естественным образов в ходе конкуренции между агентами. При этом наиболее продвинутые агенты порождают других агентов, согласно генетическому алгоритму, поддерживая тем самым их общее количество. Этот тип сеанса больше всего напоминает естественные процессы, проходящие в живой природе.

# Необходимые информационные технологии

Для успешного построения системы, которая будет обрабатывать данные с использованием генетического алгоритма, необходимо использование различных информационных технологий

## Централизованная обработка информации

Централизованная обработка подразумевает, что система работа системы будет осуществляться на удаленном высокопроизводительном сервере.

Такой подход к обработке эффективен, если решаемые задачи требуют затрат больших вычислительным ресурсов или специализированного программного обеспечения. В случае с генетическими алгоритмами использование такого подхода будет преследовать цель нахождения оптимального решения в максимально сжатые сроки.

Ожидаемый эффект использования технологии – быстрое моделирование работы генетического алгоритма.

## Децентрализованная обработка информации

Децентрализованная обработка информации подразумевает использование распределенной системы, где в роли вычислительных мощностей выступает компьютеры абонентов.

Этот подход подходит, если скорость работы генетического алгоритма не столь важна, сколько возможность проводить параллельно множество таким процессов моделирования. При этом это дает каждому абоненту более широкие возможности по настройки параметров моделирования.

Ожидаемый эффект использования технологии – возможность вести множество моделирований работы генетического алгоритма, не ограничивая себя вычислительным мощностями центрального сервера системы.

## Информационная технология поддержки принятия решений

Эта технология организует взаимодействие пользователя и вычислительной системы, где в результате обработки входных данные от пользователя, выдается рекомендация.

Технология поддержки принятия решения ориентирована на решение плохо структурированных задач. Использование генетического алгоритма позволяет решать такие задачи, где не найдено аналитическое решение, но при этом можно составить математическую модель, которая описывала бы предметную область и критерии оптимальности решения.

Ожидаемый эффект использования технологии – возможность принимать от пользователя набор данных, описывающий задачу, и находить максимально приближенное к оптимальному решение этой задачи с использованием генетического алгоритма.

Особенность применения – необходимо обеспечивать дополнительный анализ информации от пользователя для составления математической модели решаемой задачи.