

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК 004.852

Соболь  
Андрей Владимирович

Использование метода обучения с подкреплением для реализации  
искусственного интеллекта

**АВТОРЕФЕРАТ**

на соискание академической степени  
магистра технических наук

по специальности 1-40 80 05 – Математическое и программное обеспечение  
вычислительных машин, комплексов и компьютерных сетей

Научный руководитель  
Волорова Наталья Алексеевна  
канд. техн. наук, доцент

Минск 2020

## ВВЕДЕНИЕ

В последнее десятилетие наблюдается возросший интерес к нейронным сетям, которые успешно применяются в самых различных областях: медицина, физика, химия, системы безопасности, игровая индустрия.

Нейронные сети вошли в практику везде, где требуется решать задачи прогнозирования, классификации или управления. Такие направления использования во многом обусловлены несколькими причинами: богатыми возможностями и простота в использовании. Нейронные сети – исключительно мощный метод моделирования, позволяющий воспроизводить чрезвычайно сложные зависимости. В частности, нейронные сети являются нелинейными по своей природе. На протяжении длительного времени линейное программирование было основным методом моделирования в большинстве областей, потому что для него хорошо разработаны процедуры оптимизации. Следовательно, нейронные сети могут решать некоторые типы задач, с которыми возникают сложности при применении методов линейного программирования.

Нейронные сети учатся на примерах. Пользователь нейронной сети подбирает представительные данные, а затем запускает алгоритм обучения, который автоматически воспринимает структуру данных. При этом от пользователя, требуется набор эвристических знаний о том, как следует отбирать и подготавливать данные для обучения, подбирать архитектуру сети и правильно интерпретировать результаты.

Метод обучения с подкреплением – это самостоятельное и уже вполне сформировавшееся направление кибернетических исследований. Обучение с подкреплением используется в различных областях науки: нейронных сетях, психологии, искусственном интеллекте, управлении,

исследовании операций и т. д. Главное достоинство этого метода – его сравнительная простота, но не реализация: наблюдаются действия обучаемого объекта и в зависимости от результата поощряют, либо наказывают данный объект, не объясняя обучаемому объекту, как именно нужно действовать.

Исследование процесса создания искусственного интеллекта является важной частью развития компьютерных игр и позволяет создавать интересные проекты, которые будут давать исключительный игровой опыт. Хотя обучение с подкреплением и являются бурно развивающимся направлением, его применение к разработке компьютерных игр требует комплексных исследований применимости. Не существует однозначного подхода к использованию обучения с подкреплением для создания искусственного интеллекта. В частности, важен вопрос уровня абстракции предметной области. Также формализация поведения искусственного интеллекта представляет собой сложную задачу, от которой зависит поведение конечной системы.

## ОБЩАЯ ХАРАКТЕРИСТИКА

*Целью* данной работы является исследование возможностей использования нейронных сетей и метода обучения с подкреплением для реализации и обучения искусственного интеллекта в различных играх, выделение преимущественных стратегий для игр с двумя игроками и сбор статистики об эффективности выбранных методов на примере прототипа игры.

В результате должны быть описаны эффективные стратегии и гиперпараметры для достижения наилучшего результата в обучении с подкреплением для игр с двумя игроками.

Для достижения данной цели сформировался ряд задач, которые необходимо выполнить:

- проанализировать предметную область: нейронные сети и их применение в обучении с подкреплением;
- изучить модель задачи обучения с подкреплением;
- выделить фундаментальные методы решения задачи обучения с подкреплением, проанализировать их возможности и оценить недостатки;
- рассмотреть новые направления в решении поставленной задачи, провести анализ и дать собственные рекомендации на основании комбинации математически преимущественных стратегий;
- разработать и реализовать прототип для проверки предложенных алгоритмов, а также для возможности сбора разносторонней статистики с дальнейшим ее анализом.

## СОДЕРЖАНИЕ

Диссертация выполнена на 68 страницах (без приложения). Пояснительная записка включает: 6 глав, 16 рисунков, 28 библиографических источников.

Графическая часть наглядно представляет выполненную работы и полученные результаты. Весь порядок изложения в магистерской диссертации подчинен руководящей идее. Каждая из глав имеет определенное целевое назначение и является базой для последующей. Дробление материала диссертации на главы, разделы, подразделы, а также их последовательность отражаются этапы выполнения работы по теме диссертации.

Во введении освещаются степени разработанности темы и оценка современного состояния решаемой задачи, основание и исходные данные для разработки темы. Дается обоснование актуальности темы магистерской диссертации, изложение целевой установки, определяются задачи работы.

В первой главе произведён обзор предметной области задачи, решаемой в рамках магистерской диссертации, рассмотрены вопросы о сущности нейронных сетей и принципе их работ, изложены основные принципы метода обучения с подкреплением.

Во второй главе дается аналитический обзор предметной области. Дается обзор постановки задачи обучения с подкреплением, описываются основные понятия предметной области в рамках поставленной задачи. Проводится анализ собранной информации.

В третьей главе происходит описание фундаментальных методов, дается характеристика основных подходов к решению поставленных задач и обосновывается либо опровергается целесообразность их использования.

В четвертой главе изложены рекомендации по сочетанию и улучшению существующих методов и стратегий для достижения наилучшего результата в решении задачи обучения с подкреплением.

В пятой главе описана математическая модель реализованного прототипа, изложены примененные стратегии в обучении нейронных сетей для комплексной оценки эффективности различных методов обучения.

В шестой главе изложены результаты выполненных в работе теоретических и экспериментальных исследований, собрана статистика на основе реализованного прототипа.

В заключении подводится итог проведенной работы, приводятся возможные пути ее практического использования.

## ЗАКЛЮЧЕНИЕ

В данной магистерской диссертации была рассмотрена задача обучения искусственного интеллекта игры методами машинного обучения.

В ходе выполнения работы были изучены следующие темы:

- типы нейронных систем;
- многослойные перцептроны;
- применение нейронных систем в машинном обучении;
- методы обучения искусственного интеллекта;
- метод обучения с подкреплением.

Метод обучения с подкреплением был выбран предметом исследования магистерской диссертации, как один из методов, возможности применения которого не ограничены строгими рамками и не описаны полно. Алгоритм метода обучения с подкреплением позволяет подавать информацию для обучения без детализации среды, что является несомненным его преимуществом.

В результате выполнения работы была проанализирована задача обучения с подкреплением, рассмотрены исторические способы решения задачи, обозначены основные преимущества, области применения и методы реализации (как фундаментальные, так и прогрессивные). На основании изученных данных было создано программное средство – реализована игра «Реверси» с обученным по методу обучения с подкреплением компьютерным игроком. Также в ходе оптимизации базовых алгоритмов была рассмотрена и улучшена математическая модель задачи разрабатываемой игры.

На основании тестирования обученного компьютерного игрока можно утверждать, что алгоритм проводит 90% партий с выигрышем у случайного соперника. Реализованная игра является законченным

продуктом, но не исключается возможность дальнейшего исследования темы и улучшения алгоритма.

В процессе технико-экономического обоснования проекта были получены результаты, свидетельствующие о том, что данный проект является экономически целесообразным.

Библиотека БГУИР



## СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1-А. Соболев, А.В. Методы улучшения алгоритма Q-обучения / А.Ю. Волчек, А.В. Соболев, научный руководитель Н.А. Волорова // Доклады БГУИР. Компьютерные системы и сети: материалы 54-й научной конференции аспирантов, магистрантов и студентов, 2018 г. – с. 153–154.

2-А. Соболев, А.В. Преимущества решения задачи обучения с подкреплением методами Монте-Карло / А.В. Соболев, научный руководитель Н.А. Волорова // Научное сообщество студентов XXI столетия. Технические науки: сборник статей по материалам LXXXIII студенческой международной научно-практической конференции № 11(82), Россия, г. Новосибирск, ноябрь 2019 г. – с. 56–60.

3-А. Соболев, А.В. Решение задачи обучения с подкреплением методом обучения на основе временных различий / А.В. Соболев, научный руководитель Н.А. Волорова // Технические науки: проблемы и решения. Сборник статей по материалам XXIX международной научно-практической конференции № 11(27), Россия, г. Москва, Изд. «Интернаука», 2019 г. – с. 122-127.