

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

УДК 004.85

Зязюлькин
Сергей Павлович

Модели и алгоритмы на основе машинного обучения для разработки
игровых интеллектуальных агентов

АВТОРЕФЕРАТ

на соискание степени магистра

по специальности 1-40 80 04 – Информатика и технологии программирования

Научный руководитель
Нестеренков С.Н.
к.т.н., доцент

Минск 2021

КРАТКОЕ ВВЕДЕНИЕ

Под агентом понимается сущность, которая взаимодействует с окружающей средой, выполняя на основе наблюдений определённые действия и получая за них вознаграждение. Под эту абстрактную модель подходит бесчисленное множество процессов в современном мире. Агентом, например, может являться робот, которому необходимо переместиться из точки А в точку В, управляя сервоприводами в конечностях. Вознаграждение выдаётся за приближение к точке назначения и избегание препятствий на пути. Агентом также может являться программа, торгующая акциями на бирже. Действиями являются продажа и покупка акций, а в качестве вознаграждения, например, выступает разница между имеющейся денежной суммой в начале и конце торгового дня.

Проблема принятия оптимальных решений изучается в различных областях научного знания. Одной из таких областей является машинное обучение с подкреплением, напрямую адресующее проблему выбора действий агента с целью максимизации суммарного вознаграждения. Машинное обучение с подкреплением является активно развивающейся областью, многие проблемы ещё не решены, чуть ли не каждый день совершаются новые прорывы и поднимаются неразрешённые вопросы. Тем не менее, машинное обучение с подкреплением уже позволяет превзойти человека в различного рода задачах. Так, машины уже давно значительно превзошли человека в шахматах, игре Го, посягают на первенство в современных компьютерных играх, таких как Starcraft 2, Dota 2.

Процесс обучения агента является трудоёмким даже в случае решения простой проблемы. Зачастую требуется несколько дней (или даже недель) обучения на десятках GPU. Поэтому для тестирования алгоритмов и моделей машинного обучения с подкреплением используются относительно простые задачи и среды, не требующие больших затрат на решение и эмуляцию. Одним из популярных бенчмарков для алгоритмов и моделей машинного обучения являются игры для Atari 2600.

В рамках магистерской диссертации рассмотрены различные алгоритмы машинного обучения с подкреплением, способы повысить эффективность исследования игрового пространства агентами, разработан и реализован алгоритм машинного обучения с подкреплением Randomly Motivated DQN, выполнено сравнение разработанного алгоритма с алгоритмом Rainbow DQN, а также выполнен анализ обученных игровых интеллектуальных агентов.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью магистерской диссертации является разработка нового алгоритма на основе машинного обучения с подкреплением для обучения игровых интеллектуальных агентов.

Для достижения поставленной цели необходимо решить *следующие задачи*:

1. проанализировать современные алгоритмы на основе машинного обучения с подкреплением для разработки игровых интеллектуальных агентов, а также различные технические улучшения процесса обучения игровых интеллектуальных агентов;

2. проанализировать существующие механизмы внутренней мотивации игровых интеллектуальных агентов для повышения эффективности исследования игровой среды;

3. разработать и реализовать новый алгоритм на основе машинного обучения с подкреплением для обучения игровых интеллектуальных агентов;

4. выполнить обучение игровых интеллектуальных агентов разработанным алгоритмом;

5. провести анализ итогов обучения, а также выполнить сравнение разработанного алгоритма с существующими алгоритмами.

Объектом исследования являются игровые интеллектуальные агенты.

Предметом исследования являются модели и алгоритмы на основе машинного обучения с подкреплением для обучения игровых интеллектуальных агентов.

Основной *гипотезой*, положенной в основу магистерской диссертации, является возможность разработки эффективных игровых интеллектуальных агентов для различных игр при помощи моделей и алгоритмов на основе машинного обучения с подкреплением.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Автоматизация в современном мире охватывает практически все сферы человеческой деятельности. Разработка интеллектуальных агентов позволяет заменить человека в ранее немых областях. Удобной виртуальной средой для разработки, исследования и тестирования интеллектуальных агентов являются компьютерные игры.

Личный вклад соискателя

Результаты и методы, описанные в диссертации, получены соискателем лично. Вклад научного руководителя Нестеренкова С.Н. заключается в формулировке целей и задач исследования.

Апробация результатов диссертации

Основные положения магистерской диссертации были представлены на II международной научно-практической конференции «Синергия теоретического и практического подхода в научных исследованиях и разработках 21 века» (Кемерово, Россия, 2019); VI международной научно-практической

конференции «BIG DATA and Advanced Analytics» (Минск, Беларусь, 2020); международной научной конференции «Информационные технологии и системы 2020» (Минск, Беларусь, 2020); VII международной научно-практической конференции «BIG DATA and Advanced Analytics» (Минск, Беларусь, 2021).

Публикация результатов диссертации

По теме диссертации опубликовано 4 печатные работы в рецензируемых сборниках трудов и материалов международных конференций.

Структура и объём диссертации

Диссертация состоит из введения, пяти глав, заключения, списка использованных источников и приложений. Общий объём работы составляет 77 страниц, из которых основного текста – 59 страниц, 33 рисунка на 28 страницах, 6 таблиц на 5 страницах, 64 формулы на 25 страницах, список использованных источников из 24 именованных на 2 страницах и 2 приложения на 16 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

В **первой** главе диссертации рассмотрены современные алгоритмы на основе машинного обучения с подкреплением для обучения игровых интеллектуальных агентов. В первом разделе главы приведены основные определения из области машинного обучения с подкреплением:

- агент;
- среда;
- действие;
- вознаграждение;
- наблюдение.

Во втором разделе рассмотрен классический алгоритм машинного обучения с подкреплением DQN, его различные улучшения, а также следующие модификации:

- N-step DQN;
- Double DQN;
- Noisy Networks;
- Prioritized Experience Replay;
- Dueling DQN;
- Categorical DQN.

В третьем разделе первой главы кратко рассмотрены следующие Policy Gradient алгоритмы машинного обучения с подкреплением:

- Reinforce;
- A2C;
- A3C;

- PPO;
- TRPO;
- ACKTR.

В четвёртом разделе описана игровая консоль Atari 2600, а также различные улучшения процесса обучения игровых интеллектуальных агентов для игр на Atari 2600.

Во **второй** главе рассмотрены различные механизмы внутренней мотивации игровых интеллектуальных агентов для повышения эффективности исследования игровой среды:

- Counter-based – первый раздел.
- Intrinsic Curiosity Module – второй раздел.
- Random Network Distillation – третий раздел.
- Episodic Curiosity – четвёртый раздел.
- Fast and Slow exploration – пятый раздел.

В шестом разделе описаны особенности реализации алгоритмов на основе машинного обучения с подкреплением, использующих механизмы внутренней мотивации агентов.

В **третьей** главе представлен новый алгоритм Randomly Motivated DQN для обучения игровых интеллектуальных агентов. В первом разделе указаны модификации алгоритма DQN, используемые в алгоритме Randomly Motivated DQN. Во втором разделе представлен новый механизм внутренней мотивации агентов, получивший название «Случайная мотивация». В третьем разделе подробно описана используемая в Randomly Motivated DQN функция потерь, а также её основные компоненты. В четвёртом разделе рассмотрены дополнительные модификации алгоритма Randomly Motivated DQN.

Четвёртая глава посвящена деталям программной реализации алгоритма Randomly Motivated DQN:

- Используемые технологии – первый раздел.
- Описание архитектур используемых в алгоритме нейронных сетей – второй раздел.
- Выбор гиперпараметров алгоритма – третий раздел.

В **пятой** главе диссертации приведены результаты обучения игровых интеллектуальных агентов для игр на Atari 2600. В первом разделе указана конфигурация оборудования, используемого для обучения игровых интеллектуальных агентов. Второй раздел посвящён описанию алгоритма Rainbow DQN, с которым сравнивается разработанный алгоритм Randomly Motivated DQN. В следующих разделах приведены результаты обучения игровых интеллектуальных агентов при помощи алгоритмов Randomly Motivated DQN и Rainbow DQN:

- Игра Breakout – третий раздел.
- Игра Alien – четвёртый раздел.
- Игра Space Invaders – пятый раздел.
- Игра Private Eye – шестой раздел.
- Игра Gravitar – седьмой раздел.

Итоги обучения подведены в последнем (восьмом) разделе.

ЗАКЛЮЧЕНИЕ

Основные научные результаты диссертации

1. Проведён обзор современных алгоритмов на основе машинного обучения с подкреплением для обучения игровых интеллектуальных агентов, а также механизмов внутренней мотивации игровых интеллектуальных агентов для повышения эффективности исследования игровой среды.

2. Представлен новый механизм внутренней мотивации игровых интеллектуальных агентов. Разработан и программно реализован новый алгоритм Randomly Motivated DQN для обучения игровых интеллектуальных агентов.

3. Продемонстрирована эффективность алгоритма Randomly Motivated DQN при обучении игровых интеллектуальных агентов для игр на Atari 2600.

Рекомендации по практическому использованию результатов

1. Проведённый обзор формирует представление о современном состоянии области машинного обучения с подкреплением, формирует теоретическую базу для разработки новых и улучшения существующих алгоритмов для обучения интеллектуальных агентов.

2. Представленные в алгоритме Randomly Motivated DQN идеи и дополнительные модификации могут использоваться вместе или в отдельности как составные элементы при разработке новых алгоритмов для обучения интеллектуальных агентов.

3. Алгоритм Randomly Motivated DQN не ограничен использованием только для обучения игровых интеллектуальных агентов и может применяться для обучения интеллектуальных агентов другого назначения.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

1. Зязюлькин, С.П. Игровые интеллектуальные агенты для игр Atari 2600 на основе машинного обучения с подкреплением / С.П. Зязюлькин, С.Н. Нестеренков // Синергия теоретического и практического подхода в научных исследованиях и разработках 21 века: материалы II Междунар. науч.-практ. конф., Кемерово, 15 ноября 2019 г. / Западно-Сибирский науч. центр. - Кемерово, 2019. - С. 34-37.

2. Зязюлькин, С.П. Использование DQN для обучения агентов игр (Atari 2600) / С.П. Зязюлькин, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VI Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 20-21 мая 2020 года): в 3 ч. Ч. 2 / редкол. : В. А. Богуш [и др.]. - Минск : Бестпринт, 2020. - С. 274-280.

3. Зязюлькин, С.П. Использование actor-critic алгоритмов при обучении агентов для игр на ATARI 2600 / С.П. Зязюлькин, С.Н. Нестеренков // Информационные технологии и системы 2020 (ИТС 2020) = Information Tehnologies and Systems 2020 (ITS 2020) : материалы междунар. науч. конф., Минск, 18 ноября 2020 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск: БГУИР, 2020. - С. 74-75.

4. Зязюлькин, С.П. Использование внутренней мотивации при обучении агентов для игр на Atari 2600 / С.П. Зязюлькин, С.Н. Нестеренков // BIG DATA and Advanced Analytics = BIG DATA и анализ высокого уровня : сб. материалов VII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 19-20 мая 2021 года): / редкол. : В. А. Богуш [и др.]. - Минск : Бестпринт, 2021. - С. 194-202.