

О Т З Ы В

официального оппонента на диссертационную работу

Козлова Даниила Александровича

«Интеграция иерархических ансамблей и трансформерных архитектур в алгоритмы обучения с подкреплением», представленную на соискание ученой степени кандидата технических наук по специальности 1.2.1 Искусственный интеллект и машинное обучение

Актуальность темы диссертационной работы

Актуальность исследования обусловлена быстрым развитием робототехники и необходимостью адаптивного управления роботами в условиях высокой сложности и неопределенности. В современном мире растет потребность в решениях, которые позволяют роботам не только выполнять предопределенные задачи, но и быстро адаптироваться к изменениям окружающей среды, принимать автономные решения и оптимизировать свое поведение в реальном времени. Это особенно важно для задач, где детальное моделирование среды оказывается затруднительным или невозможным.

Обучение с подкреплением (RL) становится одним из ключевых методов, предоставляющих роботу способность к самообучению, адаптации и оптимизации действий без необходимости полного знания модели среды. Применение RL в робототехнике уже продемонстрировало свою эффективность на примере автономного транспорта, роботизированных рук для промышленного производства и роботов для поисково-спасательных операций. Эти примеры показывают, что RL позволяет роботам успешно действовать в условиях высокой изменчивости среды, решая задачи, требующие гибкости и автономности.

Дополнительные преимущества методов обучения с подкреплением заключаются в возможности управления различными типами роботизированных устройств — от беспилотных летательных аппаратов до манипуляторов и шагающих роботов. Поскольку RL не требует точной модели среды, он предоставляет уникальные возможности для разработки алгоритмов, способных обучаться на основе взаимодействия с реальной средой, что повышает их применимость в самых разных условиях и областях.

Новизна проведенных исследований и полученных результатов

Диссертационная работа вносит вклад в область алгоритмов обучения с подкреплением для задач управления роботами, функционирующих в трехмерных средах, посредством разработки и исследования новых подходов к интеграции RL и трансформеров, а также иерархического ансамблирования алгоритмов. Новизна работы заключается в следующем:



1. Разработана методика для оценки полезности наблюдений окружающей среды, позволяющая оптимизировать набор входных данных для агентов RL. Это дает возможность упорядочить наблюдения по их значимости, что непосредственно влияет на эффективность обучения агента.
2. Предложена и реализована модель интеграции RL и кодировщика трансформера, которая обеспечивает эффективное кодирование последовательностей состояний. Этот подход позволяет улучшить качество принимаемых агентом решений за счет более точного представления информации об окружающей среде, что особенно актуально для задач в сложных и динамично меняющихся условиях.
3. Создан новый алгоритм, совмещающий трансформерный кодировщик с алгоритмом RL Soft Actor-Critic (SAC), что обеспечивает улучшенное представление состояний и повышает эффективность обучающего процесса. Данный алгоритм превосходит классический SAC, демонстрируя улучшенные показатели по качеству решений.
4. Предложен метод иерархического ансамблирования для алгоритмов RL, который позволяет объединить несколько алгоритмов в иерархическую структуру без дополнительных обращений к среде, что сокращает требования к вычислительным ресурсам. Разработанный метод предоставляет возможность повышать качество обучения и стабилизировать процесс за счет взаимодополняющих свойств алгоритмов.
5. Создан алгоритм на основе иерархического ансамблирования, в котором распределяются роли между алгоритмом DQN, выполняющим управляющую функцию, и алгоритмами SAC и REDQ, действующими как управляемые. Такой подход повышает общую производительность и качество обучения, эффективно использует сильные стороны каждого алгоритма в ансамбле.

Обоснованность и достоверность положений, выводов и заключений, содержащихся в диссертации

Обоснованность и достоверность результатов диссертационной работы подтверждаются применением современных методов статистического анализа, тщательной экспериментальной проверкой предложенных алгоритмов и их сравнением с известными решениями. В процессе работы были проанализированы различные алгоритмы обучения с подкреплением, выявлены их ограничения и проведена серия тестов, демонстрирующих преимущество предложенных подходов.

Результаты работы были апробированы на международных научных конференциях и опубликованы в рецензируемых изданиях, что подтверждает их научную ценность. Кроме

того, полученные результаты были использованы в научных и образовательных проектах, что свидетельствует о их практической применимости и эффективности.

Значимость результатов, полученных в диссертации, для науки и практики

Научная значимость работы состоит в развитии и совершенствовании методов обучения с подкреплением для управления роботами в сложных и динамических трехмерных средах. Предложенные модели и алгоритмы, интегрирующие кодировщик трансформера и иерархические структуры RL, расширяют существующие подходы к самообучающимся агентам, улучшая их способность к адаптации и повышая качество принимаемых решений. Результаты работы делают вклад в теорию машинного обучения, предлагая новые стратегии для кодирования последовательностей состояний и управления агентами, что может быть применимо к широкому спектру задач, связанных с робототехникой и интеллектуальными системами.

Практическая значимость диссертации подтверждается возможностью применения разработанных методов и алгоритмов в реальных робототехнических системах, требующих высокой автономности и способности действовать в условиях неопределенности. Результаты исследования имеют потенциал для внедрения в такие области, как автономный транспорт, промышленная робототехника и системы поиска и спасения. Разработанные подходы могут повысить надежность и эффективность роботизированных устройств, улучшить их адаптацию к внешней среде и расширить возможности использования в различных сценариях, включая высокоточные задачи в изменяющихся условиях.

Замечания по содержанию и оформлению диссертационной работы

1. Как было верно сказано в заключении параграфа 1.4, проблема компромисса между исследованием и эксплуатацией является одним из ключевых аспектов в обучении с подкреплением. При этом, несмотря на важность описанной проблемы, в диссертационной работе Козлова Даниила Александровича не рассмотрены на достаточном уровне методы достижения компромисса (вместо рассмотрения представлена лишь ссылка на работы касательно задачи многорукого бандита, где компромисс достигается в отдельно взятом частном случае).
2. В разделе 1 при упоминании классического Q-learning (см. разделы 1.6.1 и 1.6.3) приводится формулировка, учитывающая ϵ -жадную стратегию, позволяющую достигать компромисс между исследованием среды и использованием уже известных действий. При в формальной записи вместо использования классического ϵ -коэффициента вводится

понятие скорости обучения, физический смысл которой никак не раскрывается в рамках рассматриваемого раздела.

3. В параграфе 3.4 раздела 3 рассматриваются результаты экспериментального исследования по интеграции архитектуры трансформера с методом Soft Actor-Critic. В параграфе отмечается, что такая интеграция приводит к значительному улучшению производительности исследуемого агента, особенно в сложных трехмерных случаях. Однако ниже приводится таблица изменения суммарной награды при использовании разработанного алгоритма, из которой следует, что улучшение производительности наблюдалось не во всех тестовых средах. Хочу отметить, что в диссертационной работе Козлова Даниила Александровича не приводится анализ причин такого ухудшения.
4. Аналогично предыдущему замечанию, анализ экспериментальных результатов разработанного метода иерархического ансамблирования алгоритмов обучения с подкреплением, описанного в разделе 4, показывает, что, хотя предложенный подход в среднем обеспечивает более высокий уровень вознаграждения, в некоторых кейсах (задача ходьбы) суммарная награда уменьшается (хотя и незначительно). Ввиду этого было бы желательно провести более детальный анализ причин и масштабов полученных изменений.

Указанные замечания, впрочем, не являются критическими, не нарушают целостность работы и не сказываются заметным образом на корректности полученных результатов.

Заключение

Указанные выше замечания носят, в основном, рекомендательный и редакционный характер и не оказывают существенного влияния на общее положительное восприятие работы. Диссертационная работа Козлова Д.А. «Интеграция иерархических ансамблей и трансформерных архитектур в алгоритмы обучения с подкреплением» является завершённым научным исследованием, результаты опубликованы в десяти работах. Из них одна работа в изданиях, рекомендуемых ВАК, четыре работы опубликованы в изданиях, индексируемых в БД Scopus. Диссертация отвечает требованиям Положения о присуждении учёных степеней утверждённого постановлением Правительства Российской Федерации №842 от 24 сентября 2013 года, а ее автор заслуживает присуждения ученой степени кандидата технических наук по специальности 1.2.1. Искусственный интеллект и машинное обучение.

Заведующий отделом федерального государственного
учреждения «Федеральный исследовательский центр
«Информатика и управление» Российской академии наук»
(ФИЦ ИУ РАН), доктор технических наук по специальности
2.3.1 – Системный анализ, управление и обработка
информации, статистика, доцент
email: vva@smartengines.com
тел.: +7 906 753 9377

04.12.2024

 Арлазаров В.В.

Подпись Арлазаров В.В. заверяю:

Начальник отдела кадров

И.А. Петрова
2024 г.

