Рецензия на выпускную квалификационную работу

Васильева Дениса Валерьевича

на тему:

'Методы внутренней мотивации в обучении с подкреплением на основе модели'

В диссертации Васильев Д.В. анализирует поведение функций расстояния для целенаправленного обучения с подкреплением, основанного на модели мира.

Обзор современного состояния содержит описание различных подходов, используемых для обучения функции расстояний, а также краткий обзор применения данных функций для решения задач, отличных от целенаправленного обучения с подкреплением. Также описывается процедура обучения агента целенаправленной стратегии в модели мира Dreamer и приводится краткое введение в преемственные представления. В постановочном разделе вводятся основные определения области обучения с подкреплением и обозначения, необходимые для описания процедур обучения функций расстояния. В теоретической части представлен анализ проблем, возникающих при обучении функций расстояния с помощью решения задачи регрессии в воображении; дается объяснение важности использования признаковых описаний состояний и локального вознаграждения; выполнено описание процедуры обучения расстояния с использованием уравнения Беллмана, которое лишено недостатков предыдущих подходов. Раздел, посвященный экспериментам, содержит описание лабиринтных сред, на которых оценивалась эффективность использования признаковых описаний состояний и измененной процедуры обучения функции расстояния, а также приведен краткий анализ более простых методов обучения целенаправленных стратегий, не использующих функции расстояния. Тестируемый метод позволяет более точно достигать удаленных (и поэтому сложных) состояний на среде точечного лабиринта.

Преимуществом данной работы является актуальность решения задачи целенаправленного обучения с подкреплением в отсутствии внешнезаданной функции вознаграждения. Автор демонстрирует умение поиска слабых мест алгоритмов как с помощью теоретического анализа, так и проведением экспериментов. Среди недостатков стоит отметить, что выявленные проблемы, судя по всему, не были единственной причиной плохого поведения агента (что следует из проведенных экспериментов).

Учитывая описанные недостатки, работа заслуживает оценки "Отлично" (9 баллов). Рекомендую присвоение студенту квалификации магистра.

Хачумов Михаил Вячеславович, к.ф.-м.н.,
старший научный сотрудник
Федерального исследовательского центра
«Информатика и управление»
Российской академии наук (ФИЦ ИУ РАН)

Хачумов М.В.

Подпись М.В. Хачумова заверяю
Ученый секретарь ФИЦ ИУ РАН
д.т.н. В.Н. Захаров
19 » 06 2024 г.