

**МЕТОДЫ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ ДЛЯ
УПРАВЛЕНИЯ ГРУППОЙ МОБИЛЬНЫХ ОБЪЕКТОВ НА
ПРИМЕРЕ ВИРТУАЛЬНОГО ФУТБОЛА**

Никитина Мария Алексеевна

Студент

Факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

E-mail: s02220473@mc.msu.ru

Научный руководитель — Фомичев Василий Владимирович

В настоящее время большое внимание уделяется развитию научных исследований в области искусственного интеллекта. В связи с этим стали популярны международные соревнования RoboCup, основной целью которых является создание автономных роботов-футболистов. В данной работе рассматривается задача управления группой мобильных объектов на примере игры в виртуальный футбол. Виртуальный футбол представляет собой сложную динамическую среду с объектами управления и различными стратегиями взаимодействия между ними.

В связи с развитием технологий, появляются новые возможности для управления группой мобильных объектов. Применение методов машинного обучения позволяет оптимизировать сложные стратегии, повысить эффективность алгоритмов. Наиболее хороших результатов помогает достичь применение обучения с подкреплением (Reinforcement Learning) [1].

В работе [1] рассматривается задача обучения футболистов-защитников, целью которых является удерживать мяч на некотором расстоянии от ворот. Применяется один из методов обучения с подкреплением - Batch Reinforcement Learning. Используется централизованная стратегия управления, при которой функция награды рассчитывается для всей системы, а не для каждого отдельного игрока.

Существуют и подходы, в которых учитывается состояние каждого агента. В работе [2] рассматривается мультиагентная система (MAS), приведены результаты экспериментов, показывающие, что такой подход позволяет добиться хорошего результата быстрее, благодаря взаимодействию объектов. Помимо использования метода обучения с подкреплением для выбора политики, в данной статье предлагается использовать систему назначения задач, отвечающую за назначение агенту определенной роли в команде, в зависимости от его расстояния до мяча, до ворот, и от возможного направления

удара мяча.

В рамках данной работы реализованы некоторые алгоритмы обучения с подкреплением, описанные в работах [2],[3], на языке программирования Python. В частности, реализован алгоритм Proximal Policy Optimization (PPO) с использованием библиотеки Keras для обучения агентов игре в футбол. Обучение и моделирование проведено в виртуальной среде Google Research Football [3], которая предоставляет возможности для обучения как одного агента, так и мультиагентной системы.

Литература

1. Ollino F., Solis M. A., Allende H. Batch reinforcement learning on a robocup small size league keepaway strategy learning problem //Proceedings of the 4th Congress on Robotics and Neuroscience, 2018, ser. CEUR Workshop Proceedings. – 2018. – С. 11-17.
2. Hu C., Xu M., Hwang K. S. An adaptive cooperation with reinforcement learning for robot soccer games //International Journal of Advanced Robotic Systems. – 2020. – Т. 17. – №. 3. – С. 1729881420921324.
3. Kurach K. et al. Google research football: A novel reinforcement learning environment //Proceedings of the AAAI conference on artificial intelligence. – 2020. – Т. 34. – №. 04. – С. 4501-4510.