

ИЗУЧЕНИЕ НАВЫКОВ ДВИЖЕНИЯ ФИЗИЧЕСКИ-СИМУЛИРУЕМЫХ ПЕРСОНАЖЕЙ ПРИ ПОМОЩИ OFFLINE ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Евдокимов О.В. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Карсаков А.С. (ИТМО)

Введение. В последние годы методы иерархического обучения с подкреплением (HRL) [1; 2] приобретают все большую популярность в задачах генерации движений физически симулируемых персонажей. В HRL агент разделяется на две стратегии — высокого и низкого уровня, которые обучаются раздельно. Ключевая задача обучения стратегии низкого уровня заключается в поиске отображения скрытой переменной (эмбединга) на конкретное поведение агента. Большинство существующих подходов используют для этого онлайн обучение с подкреплением, что требует значительных вычислительных ресурсов и времени. В данной работе рассматривается возможность применения офлайн обучения с подкреплением на этом этапе и предлагается соответствующий метод, позволяющий сократить вычислительные затраты.

Основная часть. Обучение с использованием офлайн RL требует предварительно собранного набора данных, содержащего траектории движения персонажа и соответствующие управляющие воздействия. Однако сбор этого набора в процессе поиска эмбедингов делает обучение столь же затратным по времени, как и традиционные онлайн методы. В данной работе предлагается альтернативный подход: использование суррогатного набора данных, сформированного в ходе обучения множества агентов задаче полнотелесного отслеживания движений (full-body motion tracking).

На первом этапе был собран указанный набор данных с использованием кода из репозитория [3], при этом предварительно проведен подбор оптимальных гиперпараметров. Затем алгоритм ASE [1] был модифицирован для работы в офлайн режиме, после чего обучение проводилось с применением метода IQL [4], показывающего лучшие результаты среди других алгоритмов офлайн RL. Для оценки эффективности предложенного подхода была проведена серия экспериментов, в которых сравнивалась скорость обучения стратегии низкого уровня при офлайн предобучении и при обучении с нуля.

Выводы. Результаты исследования подтвердили эффективность офлайн обучения с подкреплением для поиска отображения эмбедингов навыков в стратегиях низкого уровня. Сравнительный анализ показал, что модель, прошедшая офлайн предобучение, значительно ускоряет процесс обучения по сравнению с обучением с нуля. Это делает предложенный метод перспективным направлением для дальнейшего развития систем генерации движений физических агентов, сокращая требования к вычислительным ресурсам и времени обучения.

Список использованных источников:

1. *Peng X.B., Guo, Yunrong, Halper, Lina, Levine, Sergey, Fidler, Sanja.* ASE: Large-Scale Reusable Adversarial Skill Embeddings for Physically Simulated Characters // ACM Transactions on Graphics. — 2022. — May. — ARXIV_ID: 2205.01906 MAG ID: 4229044820 S2ID: 4ba973b38e448b2060bd6e2cbc0255d767ddaf98.
2. *Luo Z., Cao J., Merel J., [et al.].* Universal Humanoid Motion Representations for Physics-Based Control. — 10/2023. — URL: <http://arxiv.org/abs/2310.04582> (дата обращения 12/29/2023); arXiv:2310.04582 [cs].
3. *Chen T.* NVlabs/ProtoMotions. — 01/2025. — URL: <https://github.com/NVlabs/ProtoMotions> (дата обращения 01/20/2025); original-date: 2024-09-24T08:41:56Z.
4. *Kostrikov I., Nair A., Levine S.* Offline Reinforcement Learning with Implicit Q-Learning. — 10/2021. — URL: <http://arxiv.org/abs/2110.06169> (дата обращения 01/20/2025); arXiv:2110.06169 [cs].