

УДК 004.8

## ВЛИЯНИЕ ШУМА И НЕПОЛНОТЫ ДАННЫХ ГРАФА ЗНАНИЙ В ЗАДАЧЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Смоляков И.Ю. (ИТМО)

Научный руководитель – кандидат технических наук, доцент Муромцев Д.И. (ИТМО)

**Введение.** В работе рассматривается гибридный подход к задаче обучения с подкреплением (RL) с использованием графов знаний. Данный подход видится более перспективным и гибким по сравнению с подходами основанными только на классических и неросетевых подходах в задачах, где есть большое количество структурированной информации о задаче. В частности анализируется поведение данного гибридного подхода с использованием зашумлённых или не полных данных в исходном графе знаний. Неполные данные могут возникать в различных прикладных областях, где доступ к полной информации о состояниях и действиях ограничен или недоступен. Исследование подобных ситуаций призвано раскрыть устойчивость и эффективность представленного гибридного RL алгоритма в условиях ограниченной информации.

**Основная часть.** В данном исследовании представлена гибридная модель которая по мимо наблюдения RL среды на вход получает данные из графа знаний. Для передачи данных из графа, на нём обучается модель предоставляющая вектора для объектов этого графа. Данные вектора объединяются с наблюдением среды, что позволяет использовать не только ситуативные данные, но и доступную внешнюю информацию. Так как часто в реальном мире данные могут быть не точными, устаревшими, неполными и тд, то может быть полезным знать как данные эффекты могут влиять на предложенный подход. Для проведения исследования была выбрана следующая методика:

1. **Формирование неполных данных:** для каждого эксперимента случайным образом была отсечена и/или модифицирована часть графа знаний, используемого в RL алгоритме.
2. **Обучение RL политики:** на основе сформированных неполных данных была произведена обучение RL политики. Для сравнения результатов, также была обучена RL политика на полных данных без отсечения.
3. **Анализ результатов:** После завершения обучения RL политик, были проанализированы полученные результаты. Сравнивались показатели эффективности и стабильности работы RL алгоритма при различных уровнях неполноты данных. Оценивались показатели, такие как средняя награда, скорость сходимости и устойчивость обучения.

Для формирования неполных данных использовались следующие методы[1]:

1. Маскировка одиночных триплетов. Удаление триплетов из графа по вероятности или по какой-либо метрике (например, энтропии)
2. Маскировка всех триплетов некоторыми триплетами с определенными отношениями.
3. Шум. Для некоторых случайных триплетов случайным образом изменяем некоторую сущность или изменяем отношение.

В качестве тестовой среды использовалась среда Minigrid[2].

**Выводы.** Были проведены эксперименты по анализу поведения предложенного RL подхода в среде Minigrid. В результате было получено незначительное влияние маскирования графа знаний. В дальнейшем планируется рассмотреть влияние изменений в более семантически обширном графе знаний или с использованием в среде с динамически изменяемой средой и частей графа.

**Список использованных источников:**

1. Pujara J. Sparsity and Noise: Where Knowledge Graph Embeddings Fall Short / J. Pujara, E. Augustine, L. Getoor // Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing EMNLP 2017 / Citation Key: pujaraSparsityNoiseWhere2017. – Copenhagen, Denmark: Association for Computational Linguistics, 2017. – Sparsity and Noise. – С. 1751-1756.
2. Minigrid & miniworld: Modular & customizable reinforcement learning environments for goal-oriented tasks / M. Chevalier-Boisvert [и др.] // CoRR. – 2023. – Т. abs/2306.13831.