

**ОДНОУРОВНЕВАЯ АРХИТЕКТУРА АГЕНТА НА ОСНОВЕ МОДЕЛИ
ИЕРАРХИЧЕСКОЙ ТЕМПОРАЛЬНОЙ ПАМЯТИ****Канонир Г. (Университет ИТМО)****Научный руководитель – инженер, к.ф.-м.н., Фильченков А.А. (Университет ИТМО)**

Введение. Современные методы обучения с подкреплением имеют ряд ограничений, наложенных используемой парадигмой искусственных нейронных сетей с точечной моделью нейрона, включая слабую устойчивость к шуму во входных данных [1], низкую эффективность хранения информации в модели, приводящей к появлению проблемы катастрофического забывания и невозможности непрерывного обучения [2], а также низкую эффективность процесса обучения [3]. Использование последних достижений нейронаук в рамках новой теории интеллекта — «теории тысячи мозгов» (The Thousand Brains Theory of Intelligence) [4], а также применение модели «иерархической темпоральной памяти» (Hierarchical Temporal Memory, HTM) [5], частично реализующей данную теорию в виде модели машинного обучения, имеют потенциал как для развития уже устоявшихся методов обучения с подкреплением, так и для создания новых подходов решения этой задачи.

Ранее автором данной работы была рассмотрена потенциальная возможность применения модели HTM в обучении с подкреплением, а также были проведены поиск и анализ научных исследований, посвящённых применению модели HTM в данной области [6]. В большинстве рассмотренных работ осуществлялась попытка использования только отдельных элементов данной модели в комбинации с более традиционными подходами для обретения полезных свойств, но также присутствовали работы, авторы которых стремились создать самостоятельную модель для обучения с подкреплением на основе модели HTM. Последние представляют наибольший интерес, но даже они оставляют вопрос построения архитектуры агента открытым. Причиной тому является незавершенность модели HTM, которая в настоящее время предлагает структурные элементы и алгоритмы для формирования блока пространственно-темпоральной памяти на основе входных данных, но не предлагает готового решения для выполнения сенсорно-моторной интеграции и формирования памяти для хранения сенсорно-моторного опыта агента. Кроме того, выбор наиболее предпочтительного действия на основе хранящегося в модели сенсорно-моторного опыта мало изучен.

Целью данной работы является разработка простой и легко интерпретируемой одноуровневой архитектуры агента на основе модели HTM, а также её апробация на задаче о контекстуальном многоруком бандите с мгновенным или отложенным вознаграждением.

Основная часть. Модель HTM предоставляет возможности для формирования пространственно-темпоральной памяти на основе поступающего на её вход потока данных. Очевидно, что она подходит для формирования сенсорно-моторной памяти агента с целью хранения опыта взаимодействия агента со средой, поскольку имеет более общее назначение.

Не смотря на то, что требуется более детальное определение архитектуры такой памяти, значительно более значимым препятствием является то, что HTM не является моделью обучения с подкреплением. В ней отсутствуют механизмы, позволяющие определять стратегию поведения и выбирать действия для выполнения с целью максимизации ожидаемого дохода агента. Идея для реализации таких механизмов будет предложена далее в этой работе.

В каждый момент времени в пространственно-темпоральной памяти HTM последовательно выполняются две следующие задачи: (1) происходит формирование представления поступившего входного образа в латеральном контексте; (2) выполняется прогнозирование представления входного образа в следующий момент времени. При этом на втором этапе происходит деполяризация (предварительная активация) нейронов, которые представляют собой объединение представлений множества потенциальных входных

образов в следующий момент времени. Такое «прогнозирующее смещение» позволяет сформировать на основе входных данных в следующий момент времени представление, учитывающее латеральный контекст.

Предположим, что у нас уже имеется архитектура сенсорно-моторной памяти агента на основе пространственно-темпоральной памяти НТМ, в рамках которой происходит сенсорно-моторная интеграция и состояние которой определяет состояние агента. Учитывая данное предположение, можно формировать память действий агента взяв за основу блок пространственно-темпоральной памяти НТМ и подавая на его основной вход представление текущего действия, а на латеральный вход — представление состояния агента. Таким образом становится возможно предсказывать множество потенциальных действий в следующий момент времени, основываясь на текущем состоянии агента.

Теперь предположим, что у нас имеется модуль оценивания, который получает на вход представление действия в контексте предыдущего состояния агента, получаемое от блока памяти действий, и выполняет его оценивание, основываясь на сигнале вознаграждения. Помимо этого становится возможным использовать данный модуль для создания дополнительного «прогнозирующего смещения», позволяющего из множества всех потенциальных действий выбрать наиболее предпочтительное действие, приводящее к максимизации ожидаемого дохода агента.

Выводы. На основе описанной ранее концепции, была предложена одноуровневая, т. е. без наличия глубокой иерархии, архитектура агента. Предложенная архитектура была апробирована на задаче о контекстуальном многоруком бандите с мгновенным и отложенным вознаграждением.

Экспериментальное исследование показало состоятельность предлагаемого решения, а результаты отразили возможность агентом достичь максимального ожидаемого дохода при решении проблемы с такой постановкой задачи.

В дальнейшем планируется продолжение работы над архитектурой агента для формирования полноценной иерархии, а также более основательная апробация предлагаемой архитектуры на более сложных задачах.

Список использованных источников:

1. Liu M. et al. Analyzing the noise robustness of deep neural networks // Proceedings of 2018 IEEE Conference on Visual Analytics Science and Technology (VAST). – p. 60 – 71.
2. Goodfellow I. J. et al. An empirical investigation of catastrophic forgetting in gradient-based neural networks. – 2013. URL: <https://arxiv.org/abs/1312.6211> (дата обращения: 21.11.2022).
3. Thompson N. C. et al. The computational limits of deep learning. – 2020. URL: <https://arxiv.org/abs/2007.05558>. (дата обращения: 21.11.2022).
4. Hawkins J. A thousand brains: A new theory of intelligence. Монография. – 2021. – 288 с.
5. Hawkins, J. et al. Biological and Machine Intelligence. – 2016 – 2020. URL: <https://numenta.com/resources/biological-and-machine-intelligence/> (дата обращения: 21.11.2022).
6. Канонир Г. Обучение с подкреплением на основе модели иерархической темпоральной памяти // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки. – 2022. – № 10. – С. 80-83.

Канонир Г. (автор)

Подпись

Фильченков А.А. (научный руководитель)

Подпись