

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ ОБУЧЕНИЯ 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 предложена простая и легко интерпретируемая одноуровневая архитектуры агента на основе модели HTM [6]. Архитектура структурно представляет собой три взаимодействующих модуля: (1) память агента, построенная на основе блоков пространственно-темпоральной памяти HTM, используемой для хранения сенсорно-моторного опыта агента; (2) модуль, выполняющий оценивание выполняемых агентом действий, а также формирующий обратные синаптические связи, отражающие выработанную агентом стратегию поведения и играющие ключевую роль при выборе следующего действия; (3) модуль, ответственный за выбор следующего действия агента (учитывая предпочтения, основанные на предыдущем опыте агента). Не смотря на то, что предложенная архитектура была успешно апробирована как на задаче о классическом, так контекстуальном многоруком бандите с мгновенным или отложенным вознаграждением, для некоторых постановок задачи анализ полученных результатов показал низкую эффективность процесса обучения. В первую очередь данная проблема затрагивает случаи, когда фактически получаемое вознаграждение зависит от темпорального контекста размера меньшего, чем максимальная длина рассматриваемых в рамках решаемой задачи последовательностей наблюдений и действия.

Целью данной работы является предложение нового дизайна блока пространственно-темпоральной памяти, позволяющего агенту учитывать темпоральный контекст различной длины, но с ограничением верхнего значения.

Основная часть. Блок памяти HTM структурно представляет собой набор мини-колонок или групп нейронов. Любые два нейрона из одной мини-колонки имеют идентичное рецептивное поле, т.е. они реагируют на идентичные паттерны во входном образе, но для нейронов из разных колонок такой гарантии нет и, вероятнее всего, такие нейроны будут иметь в меньшей или большей степени различные рецептивные поля. При этом наличие множества нейронов в каждой мини-колонке используется для предоставления возможности формирования представления некоторого образа в различных темпоральных контекстах.

Число нейронов в каждой мини-колонки одинаковое и оно определяет количество контекстов, в которых может быть представлен образ с гарантированной возможностью отличить представления одного и того же образа в различных контекстах. При этом в оригинальном алгоритме модели HTM отсутствует возможность ограничения длины используемого темпорального контекста, что и является её главным ограничением, приводящим к ранее описанной проблеме.

Для снятия ограничения предлагается концепция памяти произвольного, но фиксированного порядка, в которой максимальная длина используемого темпорального

контекста является гиперпараметром. Создание такой памяти представляется наиболее естественными за счёт объединения блоков памяти первого (один нейрон в мини-колонке) и высокого (более одного нейрона в мини-колонке) порядков. При этом представление из памяти первого порядка всегда служит темпоральным контекстом первого порядка (в последующий момент времени) для памяти высокого порядка, а в памяти высокого порядка в качестве темпорального контекста также может использоваться её выходное представление (в предшествующий момент времени). Выходной сигнал каждого нейрона более не является бинарным, а представляет собой суперпозицию, из которой ясно за счёт темпоральных контекстов какой длины он стал активен. Именно последнее и предоставляет возможность ограничения максимальной длины используемого контекста в памяти высокого порядка, начало которого задаётся представлением из памяти первого порядка.

Для апробации предлагаемого решения была выбрана задача о контекстуальном многоруком бандите с отложенным вознаграждением, в которой получаемое в конце каждого эпизода вознаграждение зависит от темпорального контекста ограниченной длины.

Выводы. На основе описанной ранее концепции, был построен блок памяти произвольного, но фиксированного порядка. Данный блок стал ключевым элементом при создании нового прототипа агента на основе модели НТМ. Созданный прототип был апробирован на задаче, в которой вознаграждение зависело от темпорального контекста ограниченной длины.

Экспериментальное исследование показало состоятельность предлагаемого решения, а результаты отразили возможность агентом достичь максимального ожидаемого дохода при решении проблемы с такой постановкой задачи. Более того правильно установленная максимальная длина используемого темпорального контекста позволила решать поставленную задачу новому прототипу агента быстрее, чем оригинальному.

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. (науч. рук. Фильченков А.А.) Одноуровневая архитектура агента на основе модели иерархической темпоральной памяти // Сборник тезисов докладов конгресса молодых ученых. Электронное издание. – СПб: Университет ИТМО, [2023]. URL: <https://kmu.itmo.ru/digests/article/9901>

Канонир Г. (автор)

Подпись

Фильченков А.А. (научный руководитель)

Подпись