

ВИКОРИСТАННЯ MODEL-BASED ПІДХОДУ ДЛЯ МУЛЬТИАГЕНТНОГО НАВЧАННЯ З ПІДКРІПЛЕННЯМ

Доценко Д. В., Бурлаченко І. С.

Чорноморський національний університет імені Петра Могили, Миколаїв

В останні роки навчання з підкріпленням, особливо мультиагентне навчання з підкріпленням, знайшло широке застосування у багатьох областях, від робототехніки до ігрової індустрії. Величезна кількість ресурсів, що виділяється на взаємодію агентів з середовищем, тоді як сама властивість «sample efficiency» передбачає навчання агентів на якомога меншій кількості даних. Ідея model-based підходу формулюється наступним чином: замість того, щоб постійно взаємодіяти з середовищем для тестування нових дій під час навчання агента, потрібно збирати дані з середовища, а саме в який стан агент переходить, роблячи дії в поточному стані. Враховуючи зазначені дані можна запустити окремий процес навчання з передбаченням за станом та дією агента для підвищення ефективності навчання з допомогою нейронної мережі. Підходи [1], які використовують цю техніку, зазвичай називають методами комунікації, оскільки повна автономія агентів скомпрометована для кращої продуктивності.

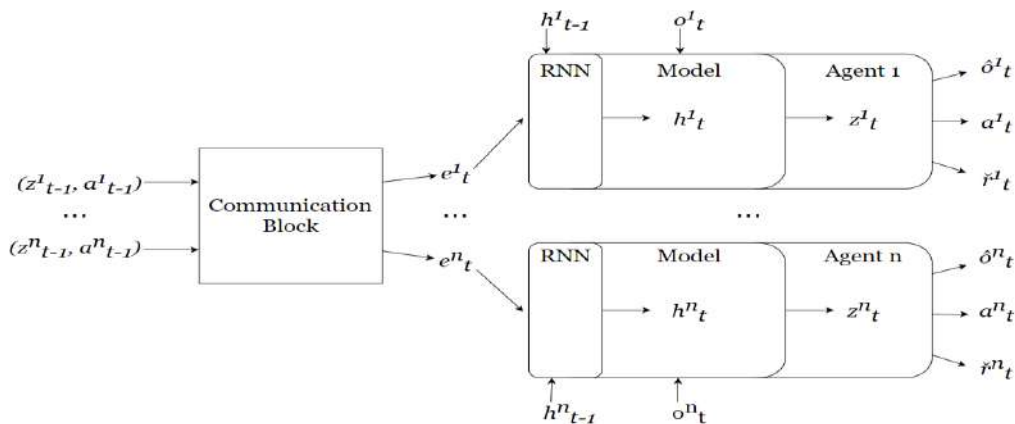


Рисунок 1 – Етапи мультиагентного навчання НМ

На рис.1 ви можете побачити, що у нас є латентні стани z та дії агентів a . Ці вектори ми передаємо до Communication Block, який є звичайним стеком Attention, який у свою чергу, видає вектори e для кожного агента, яким він може оновити свою модель середовища. Для цього може використовуватись рекурентна нейронна мережа, що переводить прихований стан h на поточний крок, а вже з цього стану та спостереження агента o , передбачається поточний латентний стан z у кожного агента. Було розглянуто особливості підходу навчання з підкріпленням на основі моделей в мультиагентних середовищах. Реалізація підходу викликала низку проблем, як з практичної, так і з теоретичної точки зору. Виходячи з отриманих результатів, можна зробити висновок, що застосування цього підходу має значний потенціал, він може зробити дослідження в галузі навчання нейронних мереж більш доступним і ефективним.

Література:

1. Baker B., Markov T., McGrew B. - Emergent tool use from multi-agent interaction - [Online resource] - Access mode: <https://openai.com/research/emergent-tool-use>