

УДК: 004.896

ПРОБЛЕМЫ В ОБЛАСТИ ГЛУБОКОГО ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Шарибаев А Н, магистрант

Московский физико-технический институт

Шарибаев Р Н, магистрант

Наманганский инженерно-технологический институт

Абдулазизов Б. Т, доцент

Тохиржонова М. Р, студент

Наманганский государственный университет

Аннотация: исследуется разработка алгоритмов обучения с подкреплением, от простых методов проб и ошибок до более сложных методов глубокого обучения с подкреплением. Также обсуждаются проблемы и ограничения обучения с подкреплением и его потенциальные применения в различных областях, таких как робототехника, игры и здравоохранение. В заключение кратко излагаются основные выводы и предлагаются направления для будущих исследований.

Ключевые слова: обучения с подкреплением, метод проб и ошибок, робототехника, разработка алгоритмов

CHALLENGES IN DEEP REINFORCEMENT LEARNING

Sharibaev A N, undergraduate

Moscow Institute of Physics and Technology

Sharibaev R N, undergraduate

Namangan Institute of Engineering and Technology

Abdulazizov B. T, Associate Professor

Tokhirjonova M. R, student

Namangan State University

Abstract: The development of reinforcement learning algorithms is explored, from simple trial and error methods to more complex methods of deep reinforcement

learning. It also discusses the problems and limitations of reinforcement learning and its potential applications in various fields such as robotics, gaming, and healthcare. The conclusion summarizes the main findings and suggests directions for future research.

***Keywords:** reinforcement learning, trial and error, robotics, algorithm development*

Обучение с подкреплением (RL) за последние годы добилось значительного прогресса, но оно по-прежнему сталкивается с рядом проблем и ограничений. Некоторые из этих проблем включают:

Эффективность выборки: алгоритмы RL требуют большого количества взаимодействий с окружающей средой для изучения эффективных политик. Это может быть дорогостоящим с точки зрения вычислений и отнимать много времени и может ограничить применимость RL к реальным задачам;

Обобщение: алгоритмы RL часто обучаются для конкретной задачи или среды и могут испытывать трудности с обобщением для новых задач или сред. Это может ограничить масштабируемость RL для реальных задач, требующих обобщения;

Безопасность и этика: агенты RL могут изучать политику, которая может не соответствовать человеческим ценностям и предпочтениям, и могут вести себя неожиданным или небезопасным образом. Обеспечение безопасного и этичного поведения агентов RL является ключевой задачей при разработке систем RL;

Стабильность и робастность: алгоритмы RL могут быть чувствительны к гиперпараметрам и изменениям окружающей среды, что может затруднить их обучение и привести к переобучению. Обеспечение стабильности и робастности алгоритмов RL важно для их применимости к реальным задачам;

Объяснимость: агенты RL могут изучать сложные стратегии и представления, которые могут быть трудны для понимания и интерпретации

людьми. Обеспечение прозрачности и объяснимости агентов RL важно для их внедрения в областях, критически важных для безопасности;

Разреженные сигналы вознаграждения: алгоритмы RL полагаются на сигналы вознаграждения для изучения эффективной политики, но во многих реальных проблемах сигналы вознаграждения могут быть разреженными или их трудно определить. Это может ограничить применимость RL к реальным задачам, требующим сложных структур вознаграждения;

Чтобы устранить некоторые из этих проблем и ограничений, исследователи разрабатывают новые алгоритмы и методы, которые могут повысить эффективность, обобщенность, безопасность, стабильность и объяснимость систем RL. Некоторые из этих подходов включают:

RL на основе моделей: Алгоритмы RL на основе моделей изучают модель окружающей среды, которую можно использовать для планирования и обоснования будущих результатов. Это может повысить эффективность выборки и обобщение алгоритмов RL, а также позволить им обрабатывать сложные и непрерывные входные пространства.

Многозадачность и мета-RL: Многозадачные и мета-RL алгоритмы учатся решать множество взаимосвязанных задач или быстро и эффективно адаптироваться к новым задачам. Это может улучшить масштабируемость и обобщенность алгоритмов RL, а также позволить им работать в разнообразных и изменяющихся средах.

Обратный RL: алгоритмы обратного RL изучают функцию вознаграждения, которая генерирует заданный набор экспертных демонстраций. Это может позволить алгоритмам RL учиться на демонстрациях экспертов вместо взаимодействия методом проб и ошибок, что может повысить эффективность их выборки и обобщения.

Безопасный RL: Алгоритмы безопасного RL гарантируют, что агенты RL ведут себя безопасным и этичным образом, путем учета ограничений безопасности и мониторинга поведения агента во время обучения и развертывания.

Подводя итог, можно сказать, что, хотя RL сталкивается с рядом проблем и ограничений, исследователи разрабатывают новые алгоритмы и методы, которые могут повысить его эффективность, обобщенность, безопасность, стабильность и объяснимость. Дальнейшее развитие RL, вероятно, окажет значительное влияние на науку и технику и откроет новые возможности в таких областях, как робототехника, автономные системы и взаимодействие человека и робота.

Литература

- [1]. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.
- [2]. Kaelbling, L. P., Littman, M. L., Moore, A. W. (1996). Reinforcement learning: A survey. Journal of artificial intelligence research, 4, 237-285.