

УДК: 004.896

## ВОЗНИКНОВЕНИЕ 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:* В данной работе дана всесторонний обзор истории обучения с подкреплением, начиная с его истоков в психологии и исследованиях поведения животных и заканчивая новейшими методами глубокого обучения с подкреплением.

*Ключевые слова:* искусственный интеллект, Машинное обучение, обучения с подкреплением

## THE EMISSION OF 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:* This paper provides a comprehensive overview of the history of reinforcement learning, from its origins in psychology and animal behavior research to the latest methods of deep reinforcement learning.

*Keywords:* artificial intelligence, machine learning, reinforcement learning

В последние годы искусственный интеллект (ИИ) становится все более важной областью с разработкой новых алгоритмов и технологий, которые позволили машинам обучаться и принимать решения на основе данных.

Машинное обучение — это раздел искусственного интеллекта, который фокусируется на разработке алгоритмов, позволяющих машинам учиться на основе данных без явного программирования. Обучение с подкреплением — это особый тип машинного обучения, который включает в себя обучение агентов принятию решений на основе поощрений и наказаний. Этот подход широко используется в робототехнике, играх и других приложениях, где агентам необходимо научиться выполнять сложные задачи методом проб и ошибок.

Было написано несколько обзоров и указателей по истории искусственного интеллекта и машинного обучения, но лишь немногие из них были сосредоточены конкретно на развитии обучения с подкреплением. В данной работе я стараюсь постараться дать всесторонний обзор истории обучения с подкреплением, начиная с его истоков в психологии и исследованиях поведения животных и заканчивая новейшими методами глубокого обучения с подкреплением.

Обучение с подкреплением уходит своими корнями в психологию и изучение поведения животных, где исследователи уже давно интересуются тем, как животные учатся методом проб и ошибок. Ранние исследования в этой области, такие как Закон эффекта Торндайка, предполагали, что животные учатся, связывая определенное поведение с положительными или отрицательными результатами. Эти результаты заложили основу для разработки алгоритмов обучения с подкреплением, которые стремятся воспроизвести этот процесс в искусственных агентах.

Происхождение обучения с подкреплением можно проследить до ранних исследований в области психологии и изучения поведения животных. Концепция обучения с подкреплением основана на идее, что организм может научиться совершать определенные действия, основываясь на последствиях этих действий.

В 1930-х годах американский психолог Б.Ф. Скиннер ввел концепцию оперантного обусловливания, представляющий собой тип обучения, при

котором поведение модифицируется его последствиями. Работа Скиннера была сосредоточена на том, как можно обучить организмы реагировать на раздражители, основываясь на последствиях их действий. Например, крыса могла бы научиться нажимать на рычаг, чтобы получить пищевое вознаграждение, или избегать нажатия на рычаг, если она получила удар электрическим током.[1]

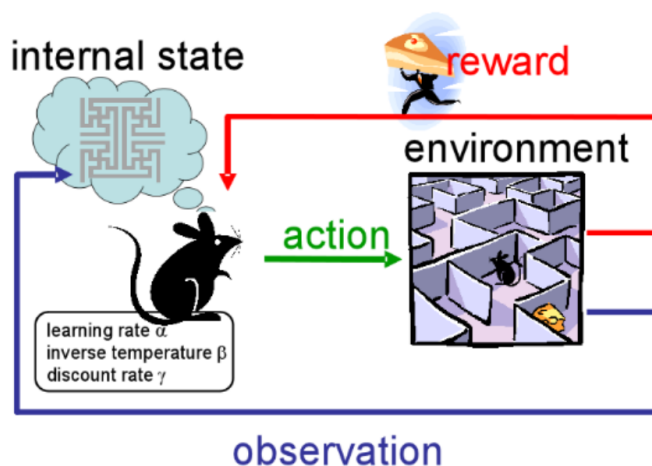


Рис.1 Схема обучения с подкреплением.[12]

Идеи Скиннера получили дальнейшее развитие у других психологов, включая Эдварда Торндайка и Кларка Халла. Закон эффекта Торндайка гласит, что поведение, за которым следует положительное последствие (награда), с большей вероятностью повторится в будущем, в то время как поведение, за которым следует отрицательное последствие (наказание), с меньшей вероятностью повторится. Теория подкрепления Халла, основанная на идеях Торндайка, ввела концепцию снижения влечения, которая относится к идее о том, что организмы мотивированы действовать на основе своих физиологических потребностей.[1]

Эти ранние идеи в психологии и исследованиях поведения животных заложили основу для развития обучения с подкреплением в искусственном интеллекте. Идея использования поощрений и наказаний для управления поведением была применена к машинам, и были разработаны алгоритмы, позволяющие машинам учиться методом проб и ошибок.

Одним из самых ранних примеров обучения с подкреплением в ИИ была работа Артура Сэмюэля в 1950-х годах. Сэмюэль разработал программу, которая могла бы играть в шашки на экспертном уровне, учась на собственном опыте. Программа использовала алгоритм обучения с подкреплением, который вознаграждал программу за победы в играх и наказывал за проигрыши. Со временем программа научилась делать лучшие ходы и смогла побеждать противников-людей.

С тех пор исследователи продолжали совершенствовать алгоритмы обучения с подкреплением, делая их более эффективными. Сегодня обучение с подкреплением является фундаментальной концепцией в области искусственного интеллекта, которая находит применение в широком спектре областей, от здравоохранения и финансов до робототехники и игр.

#### **Литература**

- [1]. Levine, S., Finn, C., Darrell, T., Abbeel, P. (2016). End-to-end training of deep visuomotor policies. *Journal of Machine Learning Research*, 17(1), 1334-1373.
- [2]. Gabriel Dulac-Arnold, Daniel Mankowitz, Todd Hester(2019) Challenges of Real-World Reinforcement Learning