

УДК 159.9

DOI: 10.34670/AR.2024.64.46.036

Обучение с подкреплением и его применение в менеджменте**Саранцева Светлана Геннадьевна**

Аспирант,
МИРЭА – Российский технологический университет,
119454, Российская Федерация, Москва, пр. Вернадского, 78;
e-mail: svesil23@mail.ru

Аннотация

В данной статье рассматривается метод обучения с подкреплением, его история, корреляция из теоретического научения Скиннера в метод машинного обучения в робототехнике и далее применение в стратегическом управлении организационной деятельностью. Проанализированы психологические факторы принятия решений руководителем компании и раскрыты барьеры, встречающиеся на пути реализации стратегии. Разработанная в статье модель позволит профессиональной деятельности быть более эффективной и продуктивной за счет сокращения затрат и времени на выработку управленческих решений. Подводя итог, стоит отметить, что, конечно, мозг человека превосходит машины в задачах на не алгоритмические решения, благодаря духовным ценностям и внутренним психологическим факторам человека, чему мы ранее могли убедиться, но машины все же превосходят в скорости и памяти. Лучшее взаимодействие на данный момент в сфере управления и принятия решений между мозгом и искусственным интеллектом может быть достигнуто через расшифровку того, что делает мозг, и передачу этой информации в компьютер.

Для цитирования в научных исследованиях

Саранцева С.Г. Обучение с подкреплением и его применение в менеджменте // Психология. Историко-критические обзоры и современные исследования. 2024. Т. 13. № 1А. С. 120-128. DOI: 10.34670/AR.2024.64.46.036

Ключевые слова

Обучение с подкреплением, reinforcement learning, психология менеджмента, психология управления, психология принятий решений, психология профессиональной деятельности, инженерная психология, психология труда.

Введение

Теория обучения с подкреплением или программированного обучения была предложена в начале 50-х гг. 20 в. американским психологом Б.Ф. Скиннером для обеспечения двух основных характеристик обучения:

- формирования у учащихся способности самостоятельно добывать знания;
- повышения управляемости процесса учения со стороны педагога.

Для Скиннера характерна биологизация процесса обучения. В этом случае в качестве основной психологической категории рассматривается понятие «научения» [Мних и др., 2015, 529-533].

Основная часть

Научение – процесс и результат приобретения индивидуального опыта.



Рисунок 1 - Схема обучения с подкреплением по Скиннеру (сегодня)

Итак, обучение с подкреплением – это психологический процесс, в основе которого лежит механизм обучения на основе последствий действий. Этот подход основан на идее, что поведение может быть усилено или подавлено благодаря последующим последствиям этого поведения [Пинто и др., 2017, 3-26].

Основой обучения с подкреплением является принцип оперантного условного рефлекса. Согласно ему, поведение усиливается или угнетается через подкрепление, которое может быть положительным или отрицательным. Б.Ф. Скиннер, проводя эксперименты на животных, показал, что они могут научиться различным поведенческим моделям, если получают подкрепление в виде награды или избегания наказания [Глешер и др., 2010, 585-595].

- Положительное подкрепление – усиление поведения путем предоставления стимула после выполнения действия.
- Отрицательное подкрепление – усиление поведения путем устранения нежелательного стимула после выполнения действия.

Области применения в психологии

- Терапевтический контекст: Обучение с подкреплением используется в психотерапии, например, при лечении фобий. Терапевт может применять подкрепление для поощрения пациента противостоять своим страхам или избегать нежелательного поведения.
- Образование и развитие: в образовательных контекстах принципы подкрепления могут применяться для стимулирования желаемого поведения студентов, повышения их мотивации и улучшения учебных результатов.
- Развитие навыков: в сфере развития профессиональных навыков, таких как лидерство или коммуникация, применение обучения с подкреплением может помочь индивидам усвоить и укрепить желаемые компетенции.

Подводя итог, стоит отметить, что, несмотря на популярность и применение, подход имеет свои критерии. Некоторые эксперты считают, что фокус на подкреплении может привести к поверхностному пониманию или механическому выполнению задачи без глубокого понимания материала или задачи. Безусловно, обучение с подкреплением представляет собой важный и широко применяемый подход в психологии, который основан на механизмах усиления или угнетения поведения через последствия. Однако, как и любой метод, он требует грамотного применения и адаптации к конкретным ситуациям и индивидам.

Обучение с подкреплением в машиностроении и робототехнике

Фрагменты теории обучения с подкреплением разрабатывались, как мы могли увидеть в предыдущей главе, с середины 1950-х годов [Хессел и др., 2019, 207-320] Но идеи, которые сформировали современное обучение с подкреплением, были собраны воедино лишь в 1989 году, когда Крис Уоткинс формализовал математический аппарат Q-обучения (Q-learning), используемый для обучения RL-агентов. Но о современном обучении с подкреплением, о начале его распространения, можно было говорить лишь после прорывной работы по AlexNet, вышедшей в 2012 году [Йех и др., 2017, 3407-3414]. А появление проекта ImageNet ознаменовало собой тот момент, когда глубокая нейронная сеть впервые значительно превзошла ранее используемые подходы к классификации объектов на сложных изображениях реального мира.

Таким образом, обучение с подкреплением (Reinforcement Learning) – это метод машинного обучения, в котором система (агент) обучается методом проб и ошибок. Идея заключается в том, что агент взаимодействует со средой, параллельно обучаясь, и получает вознаграждение за выполнение действий [Левин и др., 2016, 1-40]

В обучении с подкреплением используется способ положительной награды за правильное действие и отрицательной за неправильное [Кобер и др., 2013, 1238-1274]. В результате метод присваивает положительные значения желаемым действиям, чтобы побудить агента, и отрицательные значения – к нежелательным. Это программирует агента на поиск долгосрочного и максимального общего вознаграждения для достижения оптимального решения. Такие долгосрочные цели не дают агенту возможности останавливаться на достигнутом [Ленц, 2015, 4-5]. Со временем система учится избегать негативных действий и совершает только

ПОЗИТИВНЫЕ.



Рисунок 2 - Развернутая схема обучения с подкреплением (сегодня 2)

Области применения обучение с подкреплением в машиностроении:

- Робототехника для промышленной автоматизации
- Планирование бизнес-стратегии.
- Автоматизация внутри самого машинного обучения.
- Продвинутое система рекомендации.
- Управление движением робота, автопилот.

В заключении следует отметить, что обучение с подкреплением требует больших вычислительных ресурсов и времени, особенно когда пространство для действий у модели велико [Джейкобсон, 2018, 232-248].

Обучение с подкреплением в менеджменте

Ранее мы рассмотрели историю появления обучения с подкреплением или программированного обучения, его зарождение и применение в робототехнике. Теперь, хотелось бы провести связь и корреляцию данного обучения с теорией принятия решения в менеджменте.

Как мы помним, процесс обучения с подкреплением базируется на принципе того, что результат действия агента зависит не только от состояния, в котором это действие выполнялось, но и от действия другого игрока.

Вследствие чего метод программированного обучение целесообразен в ситуациях, когда агент сталкивается со случайным результатом от своих действий или с результатом, который зависит от действий других агентов.

Теперь обратимся к специфике функции принятия решения в деятельности

предпринимателя.

Итак, в структуре управленческой деятельности руководителя можно выделить несколько важнейших элементов – функции целеполагания, прогнозирования, планирования, организации, мотивирования, контроля. Все они по-своему важны для успешной деятельности предпринимателя, однако, многие специалисты в области менеджмента полагают, что наиболее существенным элементом среди этих аспектов работы руководителя является функция принятия управленческих решений [Абдалла и др., 2019, 801-812]. Это связано еще и с тем, что большинство других функций руководитель может делегировать своим подчиненным, но если он делегирует функцию принятия решений, то сразу теряет бразды правления и возможность управлять ситуацией, автоматически переставая быть руководителем. Если говорить об определениях, то решением называют выбор одной из ряда альтернатив в процессе достижения поставленных целей.

В своей реальной работе предпринимателю каждый день приходится иметь дело с информацией и принимать множество решений, которые касаются самых различных сфер его деятельности. Иногда ошибки в этом процессе могут быть легко исправлены, но иногда они могут оказаться фатальными для его бизнеса [Юн и др., 2019, 464-485]. Однако, хотя все понимают, что эффективная работа руководителя немислима без умения принимать качественные решения, этому умению предприниматели, как правило, не обучены, и вынуждены обучаться этому умению самостоятельно через процедуру проб и ошибок.

Отличительной особенностью функции принятия решения является то, что ее гораздо труднее стандартизировать и свести к ограниченному набору алгоритмов, в связи с этим в ней очень велика роль субъективных собственно психологических факторов [Сильвер и др., 2017, 354-359]. Существуют, конечно, многочисленные правила, процедуры и методы принятия решения, которые облегчают этот процесс, однако каждый бизнесмен по своему личному опыту знает, как велика роль субъективных и интуитивных факторов в этом процессе. В силу этого функция принятия решения является предметом изучения двух наук – теории управления, и психологии.

Классификация управленческих решений

Длительное время – вплоть до возникновения поведенческого подхода – теория управления базировалась на постулате рациональности поведения в целом и принятия решения в частности. Он состоит в том, что руководитель должен и может строить свое поведение и принимать решения, ориентируясь на максимальный учет всех факторов ситуации. Это обусловило развитие так называемых жестких схем управления, исходящих из представлений о «человеке рациональном». Однако в дальнейшем было показано, что объективно присущие человеку психофизиологические ограничения делают невозможным строго рациональное поведение и принятие решения, а полный учет всех объектных факторов при этом также в принципе невозможен. В результате была разработана концепция «ограниченной рациональности», одним из главных тезисов которой является то, что субъективные, психологические особенности – это объективные, ограничивающие факторы поведения. Они оказывают важное, а часто – определяющее влияние и на процессы принятия решения, и на управление.

Существует различные подходы к типологии управленческих решений. Одна из классификаций разделяет их интуитивные, основанные на суждениях и рациональные решения. Интуитивные решения отличаются тем, что в них слабо представлены те этапы, которые были рассмотрены выше и, в частности, этапа осознанной оценки – «взвешивания» альтернатив. Решение, основанное на суждениях, – это выбор, обусловленный знаниями и прошлым опытом.

Состояние агента (руководителя) определяется семи звеньями стратегии принятия решения

(рис. 3)

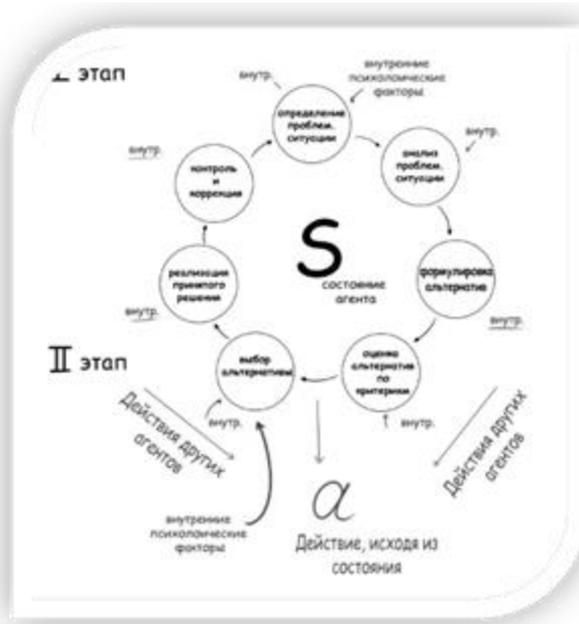


Рисунок 3 - Этапы принятия управленческого решения или внутреннее состояние агента

В отличие от них рациональные решения не имеют непосредственной опоры только на прошлый опыт как «здравый смысл», а принимаются на основе последовательности аналитических процедур [Чен, 2019, 106-119]. На каждом из этапов стратегического планирования его сопровождают внутренние психологические факторы (рис. 4).



Рисунок 4 - Психологические факторы принятия управленческих решений

Далее, как видно на рис. 3, агент производит действие «а», исходя из состояния «S» и на его действие уже оказывают влияние и внешние психологические факторы – действия других контрагентов. В результате агент получает, либо вознаграждение – удовлетворяющую оценку своих действий, фиксируя это состояние положительным опытом, либо отрицательную, получая негативный опыт.

Но стоит также отметить, как пишет В.А. Смирнов, практически каждое управленческое решение отражает индивидуальность его инициатора и систему его ценностей. Роль человеческого фактора очень велика и проявляется в двух аспектах: влияние личностных характеристик на процесс разработки управленческого решения и личностные оценки уже существующего решения [Бидер, 2019, 479-493].

Заключение

В заключение обозначим дальнейшие пути исследования, возникающие на данном этапе цифровизации общества и требующие психологического анализа и изучения:

- Итак, индустрия 4.0 пытается отдать часть интеллектуального труда машинам, но человек все еще нужен для контроля и принятия решений;
- Индустрия 5.0 охватывает все более широкие пространства, включая процесс обучения и подготовку кадров. Эти модели могут внедряться во все аспекты жизни, включая психологические портреты людей [Доу и др., 2011, 1204-1215];
- В сфере управленческой деятельности и принятия решений искусственный интеллект может быть быстрее и мощнее, но не может решать задачи на не алгоритмические стратегии, которые требуют интуиции и догадок, а, по утверждению Ричарда Фейман, великие открытия не делаются алгоритмическим путем, а скорее интуитивно или с помощью догадок [Конидарис и др., 2018, 215-281].

Подводя итог, стоит отметить, что, конечно, мозг человека превосходит машины в задачах на не алгоритмические решения, благодаря духовным ценностям и внутренним психологическим факторам человека, чему мы ранее могли убедиться, но машины все же превосходят в скорости и памяти. Лучшее взаимодействие на данный момент в сфере управления и принятия решений между мозгом и искусственным интеллектом может быть достигнуто через расшифровку того, что делает мозг, и передачу этой информации в компьютер.

Библиография

1. Абдалла С. и др. Обучение с подкреплением в бизнесе и менеджменте: прошлое, настоящее и перспективы на будущее // *European Journal of Operational Research*. 2019. V. 278 (3).
2. Бидер С., Ниссен В. Применение обучения с подкреплением в процессе принятия оперативных решений // *Европейский журнал операционных исследований*. 2019. № 274 (2).
3. Глешер Дж. и др. Состояния в сравнении с вознаграждениями: сигналы об ошибках диссоциативного нейронного прогнозирования, лежащие в основе обучения с подкреплением на основе моделей и без моделей // *Neuron*. 2010. 66 (4).
4. Джейкобсон С.Х., Алемзаде Х. Обучение с подкреплением для составления расписания: обзор литературы // *Омега*. 2018. 77.
5. Доу Н. и др. Влияние моделей на выбор людей и ошибки предсказания стриатума // *Neuron*. 2011. 69 (6).
6. Йех С. и др. Изучение визуомоторного контроллера для реального роботизированного захвата с использованием имитируемых изображений глубины // *Материалы Международной конференции IEEE по робототехнике и*

- автоматизации. 2017.
7. Кобер Дж. и др. Обучение с подкреплением в робототехнике: обзор // Международный журнал исследований робототехники. 2013. 32 (11).
 8. Конидарис Г.Д. и др. От навыков к планированию задач: основа для изучения моделей, основанных на правилах // Журнал исследований искусственного интеллекта. 2018. 61.
 9. Левин С. и др. Комплексное обучение глубоким зрительно-моторным стратегиям // Журнал исследований машинного обучения. 2016. 17 (39).
 10. Ленц И. и др. Глубокое обучение для обнаружения роботизированных захватов // Международный журнал исследований робототехники. 2015. 34.
 11. Ли М. Обучение с подкреплением для систем управления: обзор // Automatica. 2019. 111, 108633.
 12. Мних В. и др. Контроль на уровне человека посредством глубокого обучения с подкреплением // Nature. 2015. 518 (7540).
 13. Пинто Л., Асфур Т., Крагич Д. Асимметричное исследование-эксплуатация в эпизодическом обучении с подкреплением // Международный журнал исследований робототехники. 2017. 36 (1).
 14. Саттон Р.С., Барто А.Г. Обучение с подкреплением: введение. Издательство Массачусетского технологического института. 1998.
 15. Сильвер Д. и др. Овладение шахматами и сегои путем самостоятельной игры с использованием общего алгоритма обучения с подкреплением // Nature. 2017. 551.
 16. Фу Дж. и др. Подход, основанный на обучении с подкреплением, для контроля риска переутомления при вождении в сценарии следования за автомобилем // IEEE. 2019. 7.
 17. Хессел М. и др. Обзор обучения с подкреплением в робототехнике. Основы и тенденции в робототехнике. 2019.
 18. Чен Б., Сюй С. Изучение влияния обучения с подкреплением на принятие организационных решений // Журнал организационного поведения. 2019. 40 (1).
 19. Чжан У. и др. Поиск нейронной архитектуры для городских энергетических систем. 2018.
 20. Чирзин И.А., Трегубов А.А. Обучение с подкреплением в управлении цепочками поставок: обзор и направления будущих исследований // Журнал промышленной инженерии и менеджмента. 2019. 12 (3).
 21. Шульман Дж. и др. Оптимизация политики доверительного региона. 2015.
 22. Юн Дж. и др. Приложения для обучения с подкреплением в управлении доходами: обновленный обзор // Журнал управления доходами и ценообразованием. 2019. 18 (6).

Reinforcement learning and its application in management

Svetlana G. Sarantseva

Postgraduate,
Russian Technological University,
119454, 78, Vernadskogo ave., Moscow, Russian Federation;
e-mail: svesil23@mail.ru

Abstract

The theory of reinforcement learning or programmed learning was proposed in the early 50s. 20th century American psychologist B.F. Skinner to ensure two main characteristics of learning: developing students' ability to independently acquire knowledge; increasing the controllability of the learning process on the part of the teacher. This article discusses the reinforcement learning method, its history, correlation from Skinner's theoretical learning to the machine learning method in robotics and further application in strategic management of organizational activities. The psychological factors of decision-making by the head of the company are analyzed and the barriers encountered in the implementation of the strategy are revealed. The model developed in the article will allow professional activity to be more effective and productive by reducing the cost and time for making management decisions. To summarize, it is worth noting that, of course, the human brain is superior to machines in tasks involving non-algorithmic decisions, thanks to spiritual values and

internal psychological factors of a person, which we could previously verify, but machines are still superior in speed and memory. The best interaction to date in control and decision-making between the brain and artificial intelligence can be achieved through deciphering what the brain is doing and transmitting that information to a computer.

For citation

Sarantseva S.G. (2024) Obuchenie s podkrepleniem i ego primenenie v menedzhmente [Reinforcement learning and its application in management]. *Psikhologiya. Istoriko-kriticheskie obzory i sovremennye issledovaniya* [Psychology. Historical-critical Reviews and Current Researches], 13 (1A), pp. 120-128. DOI: 10.34670/AR.2024.64.46.036

Keywords

Reinforcement learning, reinforcement learning, management psychology, management psychology, psychology of decision-making, psychology of professional activity, engineering psychology, labor psychology.

References

1. Abdalla S. et al. (2019) Reinforcement learning in business and management: past, present and future prospects. *European Journal of Operational Research*, 278 (3).
2. Bider S., Nissen V. (2019) Application of reinforcement learning in the process of making operational decisions. *European Journal of Operational Research*, 274 (2).
3. Chen B., Xu X. (2019) Studying the influence of reinforcement learning on organizational decision making. *Journal of Organizational Behavior*, 40 (1).
4. Chirzin I.A., Tregubov A.A. (2019) Reinforcement learning in supply chain management: a review and directions for future research. *Journal of Industrial Engineering and Management*, 12 (3).
5. Dow N. et al. (2011) The influence of models on people's choice and prediction errors in the striatum. *Neuron*, 69 (6).
6. Fu J. et al. (2019) A reinforcement learning approach to control the risk of fatigue while driving in a car following scenario. *IEEE*, 7.
7. Glasher J. et al. (2010) States versus rewards: Dissociative neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66 (4).
8. Hessel M. et al. (2019) *A review of reinforcement learning in robotics. Fundamentals and trends in robotics.*
9. Jacobson S.H., Alenzadeh H. (2018) Reinforcement learning for scheduling: a literature review. *Omega*, 77.
10. Kober J. et al. (2013) Reinforcement learning in robotics: a review. *International Journal of Robotics Research*, 32 (11).
11. Konidaris G.D. et al. (2018) From skills to task planning: a framework for learning rule-based models. *Journal of Artificial Intelligence Research*, 61.
12. Lee M. (2019) Reinforcement learning for control systems: a review. *Automatica*, 111, 108633.
13. Lenz I. et al. (2015) Deep learning for detection of robotic grippers. *International Journal of Robotics Research*, 34.
14. Levin S. et al. (2016) Complex training of deep visuomotor strategies. *Journal of Machine Learning Research*, 17 (39).
15. Mnich V. et al. (2015) Human-level control through deep reinforcement learning. *Nature*, 518 (7540).
16. Pinto L., Asfour T., Kragic D. (2017) Asymmetric exploration-exploitation in episodic reinforcement learning. *International Journal of Robotics Research*, 36 (1).
17. Shulman J. et al. (2015) *Optimization of trust region policy.*
18. Silver D. et al. (2017) Mastering chess and shogi through independent play using a general reinforcement learning algorithm. *Nature*, 551.
19. Sutton R.S., Barto A.G. (1998) *Reinforcement learning: an introduction.* MIT Press.
20. Yeh S. et al. (2017) Study of a visuomotor controller for real robotic grasping using simulated depth images. In: *Proceedings of the IEEE International Conference on Robotics and Automation.*
21. Yoon J. et al. (2019) Reinforcement learning applications in revenue management: An updated review. *Journal of Revenue and Pricing Management*, 18 (6).
22. Zhang W. et al. (2018) *Searching for neural architectures for urban energy systems.*