REINFORCEMENT LEARNING: FRAMEWORK, APPLICATIONS AND CHALLENGES

Manbir Sandhu*, Gurpreet Kaur*, Purnima*, Anuradha Saini**
Department of Computer Science, SGGS College, Chandigarh, Panjab University, Chandigarh
Department of Computer Applications, CBSA College, Chandigarh Group of Colleges, Landran
manbirsandhu.sggs26@gmail.com
anuradhasaini.1984@gmail.com
gurpretkg@gmail.com
purnima.arvind@gmail.com

ABSTRACT

Artificial Intelligence is the science of developing intelligent machines that can simulate human intelligence and perform complex tasks. Reinforcement learning (RL) is the trending interdisciplinary learning paradigm of machine learning and optimal control that mirrors the hit and trial learning mechanism of humans. The self-learning agents take actions in a dynamic environment, get rewarded or punished and use this feedback to learn optimal decision-making to maximize cumulative rewards over time. This objective of this paper is to present an insight into promising future technology. The introduction to the technology is followed by the key concepts that constitute the reinforcement framework, the different models of this paradigm, the applications that span a wide arena and the challenges facing this ever-evolving strategy.

Keywords: Artificial intelligence, Reinforcement learning, decision-making, learning paradigm

1. INTRODUCTION

In the recent years, artificial intelligence has made rapid strides in the development of machines that have improved learning capability and are highly responsive. Reinforcement learning has dramatically transformed the domain of AI by enabling the development of autonomous systems that can interact with their dynamic environment imitating the human learning behavior. Following the hit and trial method of human learning, these machines learn optimal interaction over time to continually improve their performance. It is due to the power of RL that many complex and unsolvable problems have found a solution. RL is based on the mathematical tool -Markov Decision Process (MDP) that is used for formalized decision-making by structuring RL problems. The aim of this study is to briefly understand how reinforcement learning differs from the previous techniques of machine language, introduce the framework of reinforcement learning by presenting its basic concepts followed by the common RL algorithms, the applications of reinforcement learning and the issues that limit it. This study can enable further exploration and research for specific problem-solving.

1.1 TYPES OF MACHINE LEARNING PARADIGMS

In the field of artificial intelligence (AI), the machines are not programmed but made to learn. The learning paradigms refer to various approaches or strategies by which machines can acquire knowledge, improve performance, and make decisions. There are several learning paradigms in AI, and they can be broadly categorized into the following three domains:

1. Supervised Learning: models are trained or learn from labeled data
2. Unsupervised Learning: modes detect patterns from the data that is unlabeled.
3. Reinforcement Learning(RL): agents learnoptimal decision-making by hit and trial throughinteraction with an unknown dynamic environment without any prior training with data samples.

2. FRAMEWORK OF REINFORCEMENT LEARNING

The basic components that form part of reinforcement learning:

1. Agent: an entity that performs actions in the environment and subsequently makes decisions.
2. Environment: is the domain or a system with which the agent interacts like an education system or video game etc. It is defined or created by a clear set of rules, the permissible actions that an agent can take and the possible states for that particular environment. Generally the action and current state are the inputs to an environment which processes the input to produce the resulting state and reward as the output and also gives feedback to the agent.

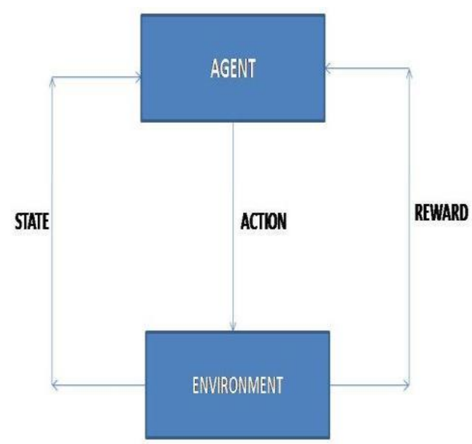


Fig. 1. Basic RL Framework

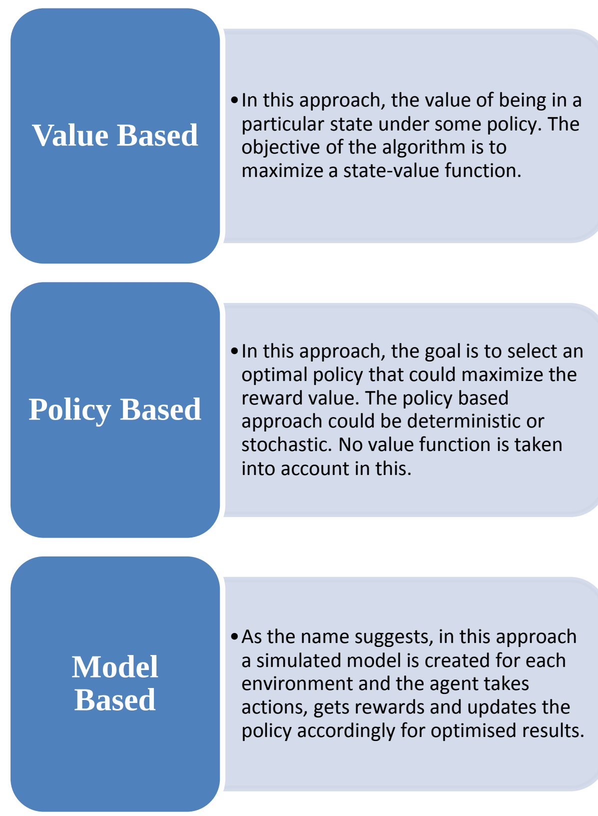
3. State: represented by coordinates (numbers or letters) refer to a

specificcondition or configuration of the environment. For any environment there is a start state and a terminal state and the intermediate states may differ in context and number from one environment to another. For a video game environment, the states could be menu state, loading state, playing state, pause state and victory/defeat state to name a few.

4. Action: can be a move or an activity that an agent is allowed to do in a particular state and environment respectively. In the context of gaming environment, action could be moving the cursor to the top, bottom, left and right.
5. Reward: is a numerical value provided by the environment in response to the agent's action in a particular state to indicate the immediate benefit or cost associated with the agent's action. Reward (positive or negative) is a sort of feedback to the agent's action in a particular state.
6. Policy: is a function that maps the set of states to the set of actions. An optimised policywill maximize the expected cumulative reward gradually. For this, the agent explores the diverse actions in different states, observes the outcomes, and adjusts its policy based on the received rewards.

3. REINFORCEMENT ALGORITHMS

Broadly, the machine learning algorithms that incorporate the learning mechanism based on reinforcement uses one of the below mentioned approach.



4. COMMON RL ALGORITHMS IN AI

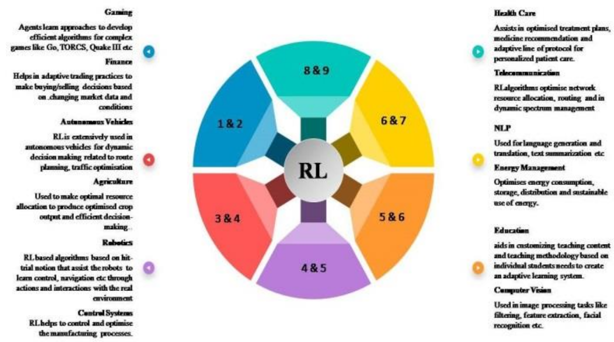
RL Algorithm	Features	Limitations
Q-Algorithm	• Model-free • Off-Policy • Based on Bellman Equation	• Large Memory Requirement • Fails to consider all actions and states in training • Unsuitable if the action and state space is large

State Action Reward State Action (SARSA)	• Model-Free • Single Policy based	• Large Memory Requirement • Fails to consider all actions and states in training • Unsuitable if the action and state space is large
Deep Q-Neural Network (DQN)	• Q-value function estimated using neural network • Permits direct approximation of optimal value using Experience Relay and Target network technique.	
Deep Deterministic Policy Gradient (DDPG)	• Actor-critic algorithm • Uses neural networks to determine policy and value functions • Suitable for continuous action space. • Permits direct approximation of optimal value using Experience Relay and Target network technique. • Uses another NN that finds the maximum approximation to determine the target.	
Trust Region Policy Optimisation (TRPO) and Proximal Policy Optimization(PPO)	• Single-policy algorithm • TRPO uses trust region to determine the optimal policy update. • PPO uses a "clipped" objective function to determine the optimal policy update. • Suitable for large high-dimensional state space and continuous action space.	• Complicated computation and implementation of TRPO

6. APPLICATIONS OF REINFORCEMENT LEARNING

Due to its versatility, reinforcement learning has found applications as an indispensable tool in diverse domains. Continuous research efforts are underway to determine newer and diverse realms that could gain from its applicability and at the same time it is being explored rapidly for the enhancement of the existing optimising problems in various areas. The prime areas of application include:

5. CHALLENGES OF REINFORCEMENT LEARNING



As the complexity of the real world problems increase, reinforcement learning (RL) is limited by a few challenges with respect to the key components of its framework.

1. Enormous state and action space: Certain complex problems/games have an exponentially high-dimensional state and action space that is beyond the processing capability of any computer. The solution to this issue

could be using some sort of approximation functions or techniques to reduce the state-action space while taking into account the problems like overfitting and errors arising due to approximation due to these functions.

2. Hit and Trial Based Learning: RL learns from real world interactions that have restricted, sparse and noisy data which affects the quality and reliability of the learning process and also makes it expensive and time-consuming. Incorporating prior knowledge and learning from previous experiences and introducing uncertainty information into the RL algorithms can increase the efficiency of data used in these algorithms.
3. Applicability of Learning: In order to improve their performance and adaptability, RL algorithms must be capable of applying the learned knowledge to new but similar tasks (generalisability) and also to new tasks/domains but with different features (transfer learning) which in itself is a task as the RL systems have no prior learning.
4. Try and Use:For adaptive and intelligent learning, RL algorithms face the challenge of striking a balance between trying new action and state space and the selection or use of the best learnt lessons to maximize rewards because both these activities could clash with each other adversely affecting the performance.

8. CONCLUSION

AI driven applications success lies in the machine's ability to learn and apply the learnt knowledge to diverse domains. Supervised, unsupervised and now reinforcement learning have exhibited promising results in their specific problem areas. The unification of these learning paradigms will be the future technological breakthrough. As of now, RL is still an evolving concept that requires focus on the development of directed strategies that could scale to high-dimension and increased complexity.

7. REFERENCES

[1] E. Akanksha, Jyoti, N. Sharma and K. Gulati, "Review on Reinforcement Learning, Research Evolution and Scope of Application," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 1416-1423, doi: 10.1109/ICCMC51019.2021.9418283.

[2] M. Naeem, S. T. H. Rizvi and A. Coronato, "A Gentle Introduction to

Reinforcement Learning and its Application in Different Fields," in IEEE Access, vol. 8, pp. 209320-209344, 2020, doi: 10.1109/ACCESS.2020.3038605.

[3] F. AlMahamid and K. Grolinger, "Reinforcement Learning Algorithms: An Overview and Classification," 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), ON, Canada, 2021, pp. 1-7, doi: 10.1109/CCECE53047.2021.9569056. **NA**

[4] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction," in IEEE Transactions on Neural Networks, vol. 9, no. 5, pp. 1054-1054, Sept. 1998, doi: 10.1109/TNN.1998.712192. **NA**

[5] K. Arulkumaran, M. P. Deisenroth, M. Brundage and A. A. Bharath, "Deep Reinforcement Learning: A Brief Survey," in IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 26-38, Nov. 2017, doi: 10.1109/MSP.2017.2743240.

[6] L. Lyu, Y. Shen and S. Zhang, "The Advance of Reinforcement Learning and Deep Reinforcement Learning," 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, 2022, pp. 644-648, doi: 10.1109/EEBDA53927.2022.9744760. NA

[7] I. Čík, J. Magyar, M. Mach and N. Ferenčík, "Reinforcement learning as a service," 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Herlany, Slovakia, 2020, pp. 45-50, doi: 10.1109/SAMI48414.2020.9108716. NA

[8] J. Jia and W. Wang, "Review of reinforcement learning research," 2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC), Zhanjiang, China, 2020, pp. 186-191, doi: 10.1109/YAC51587.2020.9337653.

[9] F. Sahba, H. R. Tizhoosh and M. M. A. Salama, "A Reinforcement Learning Framework for Medical Image Segmentation," The 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC, Canada, 2006, pp. 511-517, doi: 10.1109/IJCNN.2006.246725

[10] Hu M, Zhang J, Matkovic L, Liu T, Yang X. Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions. J ApplClin Med Phys. 2023 Feb;24(2): e13898. doi: 10.1002/acm2.13898. Epub 2023 Jan 10. PMID: 36626026; PMCID: PMC9924115.

[11] Kevin Zhou, Hoang Ngan Le, KhoaLuu, Hien V Nguyen, Nicholas Ayache, Deep reinforcement learning in medical imaging: A literature review,Medical Image Analysis,Volume 73,2021,102193,ISSN 1361-8415,https://doi.org/10.1016/j.media.2021.102193.

[12] Ayman El-Baz, GeorgyGimel'farb, Kenji Suzuki, "Machine Learning Applications in Medical Image Analysis", Computational and Mathematical Methods in Medicine, vol. 2017, Article ID 2361061, 2 pages, 2017. https://doi.org/10.1155/2017/2361061