# Exploring the Role of Reinforcement Learning in Video Game Environments

Xingyu Hou

School of Computer Science, Central South University, Changsha, 410083, China
8208201128@csu.edu.cn

**Abstract.** With the progressive advancement of technology in the gaming industry and the growing maturity of Reinforcement Learning (RL) techniques, the utilization of RL in gaming has significantly evolved. However, there is a noticeable dearth of comprehensive overview and summary in this domain. Thus, this paper aims to address this gap. It begins by providing a concise explanation of the concept of RL and subsequently reviews its development within the gaming field. A comparative analysis between traditional artificial intelligence and artificial intelligence integrating RL in gaming is also conducted to highlight the advantages and disadvantages of RL. Furthermore, the study presents two application cases demonstrating the potential of RL in enhancing gaming experiences. Ultimately, the paper concludes with a summary and a future perspective on this subject matter. Through this work, we hope to provide valuable insights and contribute to the further exploration of RL in the gaming industry.

**Keywords** Reinforcement Learning, Artificial Intelligence, Video Game.

## 1 Introduction

In the past years, there have been notable advancements in the field of Artificial Intelligence (AI), specifically in the realm of reinforcement learning (RL). RL, a subdomain of machine learning, centers on teaching agents to make sequential choices by engaging with their environment. A fascinating use case of RL that has captured the attention of researchers is its integration into video games [1]. As game technology continues to advance and user experience requirements continually increase, traditional programming methods often face challenges when dealing with complex and diverse game scenarios. Rule-based AI agents, in particular, often lack flexibility and struggle to adapt to changes in the game environment and player behavior. Reinforcement learning, on the other hand, aims to mimic human learning processes, enabling game agents to continuously improve their performance and adaptability through interactive learning.

In RL, agents learn optimal action strategies through interaction with the game environment. Agents select actions based on the current environmental state and receive rewards or penalties from the game environment. Through repeated trials and


P. Kar et al. (eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Advances in Computer Science Research 108,
https://doi.org/10.2991/978-94-6463-300-9_20

learning, agents can optimize their decision-making processes, gradually mastering the rules and strategies of the game. This learning approach allows agents to not only handle situations that can be anticipated in advance but also deal with unknown game scenarios, showcasing strong robustness and adaptability. The application of RL in games has achieved remarkable results. Through reinforcement learning, game agents can learn efficient action selection strategies, demonstrating impressive capabilities in gameplay. For example, in complex mazes, game agents can find optimal paths, avoiding traps and enemies, through exploration and trial-and-error. Furthermore, agents in adversarial games can learn opponents' strategies through reinforcement learning and exploit their weaknesses to improve their own winning probabilities.

While RL has made some progress in the gaming domain, there is a lack of in-depth literature summarizing and exploring this field. Therefore, this article aims to explore the application of RL in video game, focuses on the development of RL in the field of electronic games, a comparison between RL AI and traditional game AI, examples of applying RL AI in the gaming industry, and the future of RL in the gaming field. This paper will discuss various RL algorithms used in video game playing, highlight notable achievements, and examine the impact of these developments on the gaming industry.

## 2 The Basic of RL

### 2.1 The Definition of RL

RL constitutes a self-regulated learning paradigm grounded in the principles of Markov Decision Processes (MDPs). Fundamentally, this domain revolves around the notion of agents continually refining their decision-making frameworks through persistent engagement with their environment, with the ultimate aim of accruing maximal cumulative rewards for specific tasks.

Within the context of RL, agents select distinct actions premised on current observations or states, prompting the environment to yield correlative rewards and novel observations or states in return. Reward signals act as indicators, reflecting the degree to which the agent's actions, within a particular state, influence task-specific objectives [3]. By learning from environmental feedback, agents can adapt their decision-making approaches to maximize cumulative rewards over an extended period. It is imperative to emphasize the significance of striking a balance between exploration and exploitation in RL. To achieve this equilibrium, agents must display the audacity to investigate previously unexplored state and action spaces while concurrently utilizing established knowledge effectively. When a typical RL algorithm solves the MDP problem, usually involves steps such as Policy Evaluation, Policy Improvement, and Policy Iteration. It is mainly divided into Value-based Methods, Policy-based Methods and Actor-Critic Methods.

Value function-based methods: This class of methods is represented by Value Iteration algorithms and Quality learning (Q-learning) algorithms. The value iteration algorithm updates the action value function with the given state value function by means of dynamic programming to obtain the optimal strategy. The Q-learning algorithm estimates the action value function independently of the model, and iteratively updates it with the help of Bellman Expectation Equation to converge to the optimal strategy.

Policy-based method: The algorithm is Policy Gradient and Reinforce algorithm. The optimal strategy is found through the gradient ascending strategy parameters, and the objective function is optimized, that is, the cumulative reward of the environment. This method searches directly in the policy space and does not involve the calculation of value functions. The Reinforce algorithm was a strategy gradient algorithm based on the Monte Carlo Method and used the sampled trajectories to estimate the gradient information to update the policy parameters.

Actor-critic approach: The Actor learns the strategy and the Critic evaluates the strategy. Such algorithms include Deterministic Policy Gradient algorithm, Soft Actor-Critic (SAC) algorithm, etc. Through the collaborative training of actors and critics within the scope of RL, it can simultaneously take into account direct policy optimization as well as utilize value functions to assess the efficacy of the employed policy [4]. This enables an effective balance between learning rate and search precision during the quest for the optimal strategy.

### 2.2 The Evolution of RL in Game Field

After Alpha Go made great achievements in Go, Deep mind turned its attention to more complex Real-Time Strategy games (RTS) in 2019, which is also considered the earliest application of RL in the game field (Fig.1). They trained AlphaStar to play StarCraft 2, pre trained through Supervised learning, and then used RL for subsequent training. AlphaStar not only defeated a large number of top game players on the ladder, reached the rank of master, but also achieved a high victory rate against professional players. As an AI training project of the same scale as almost the same period as Alphastar, OpenAI also achieved tremendous success in the Multiplayer Online Battle Arena (MOBA) type game Defense of the Ancients 2 (DOTA2) in 2019 April, with their AI not only achieving a 0.994 victory rate, but also defeating the world champion team. After summarizing the experience of the DOTA project, Open AI became curious about whether AI could learn to use props in the game [4]. The program they trained, Hide and Seek, achieved unexpected results, and in the game, the AI not only learned to use props to avoid finding, but also learned to exploit the loopholes in the game's mechanics to win.


**Fig. 1.** The picture shows some cases of RL in recent years [4]

# 3 Compare RL with the Traditional Game AI

## 3.1 The Traditional Artificial Intelligence Implementations in Games

The traditional methods of game AI are to build a Finite-state machine and a Behavior tree. These methods are very intuitive and simple. Even people who do not understand AI can design different responses for different trigger scenarios according to logic. The Behavior tree has a decision logic that is easy to script/visualize, and can quickly and conveniently organize behavior decisions for AI. However, the disadvantages are also obvious. AI responses are controlled by fixed code logic, The richer the response, the more complex the tree, and thus the greater the performance overhead. Moreover, even a complex tree cannot cope with all the situations that may arise, which means that traditional game AI can only control insignificant Non-Player Characters (NPCs) without the ability to make their own decisions through player behavior or analyze the game itself. For gamers, they do not want repetitive and monotonous gaming experiences, while traditional game AI cannot handle complex scenarios.

## 3.2 The RL Based Artificial Intelligence in Games

The current mainstream game AI is called learning-based AI. There are two main technologies.

The first one is Supervised learning, but the problem of Supervised learning is very dependent on the quantity and quality of data. AI needs enough data to learn good behavior, and needs data of all kinds of quality, at all levels, with enough quality data. The second one is RL. Compared with imitation learning and Supervised learning, RL has a core advantage that it does not require human data. AI plays with itself through games and generates its own strategies through constant iteration. Just like the process of human trial and error and improvement. RL needs to constantly try to make AI make decisions, such as economy and medical care. In the field of transportation, a mistake may mean a huge loss of economic property or even life, but in the field of games, mistakes almost cost nothing, which also makes the game field the best testing ground for RL. Through the constantly improved algorithm and training architecture, the AI using RL not only accelerates the training speed, but also improves the training results. In many game fields, AI has reached the level of professional players. Through RL, AI has the ability to understand the game mechanism and self-generate decisions. AI based on self-gaming is increasingly applied to the game field. They not only have made outstanding achievements, but also stimulate AI developers and game developers to constantly improve and perfect AI based on RL and games using RL AI.

However, there are still many difficulties in the application of AI based on RL in the game field. The huge performance cost and complexity of the project required for training make many smaller game companies have no ability to apply. It is also difficult for RL to train AI with universality. Sometimes the best strategy selected by AI does not conform to the design idea of the game. For example, in the game of Snake, after the length of the snake controlled by AI reaches a certain degree, it will choose to constantly circle to avoid hitting itself and completely give up eating the next food. Designers need to continuously improve reward and punishment systems and algorithmic strategies to make AI more in line with human expectations. Table 1. shows the advantage and disadvantage of traditional AI and AI with RL.

**Table 1.** Advantage and disadvantage of traditional AI and AI with RL

| | Advantage | disadvantage |
|---|---|---|
| Traditional AI | Simple design,<br>clear logic,<br>low performance overhead<br>easy scripting | Single feedback<br>Design difficulty increases with behavior<br>Inability to make self-decisions |
| AI with RL | Can analyze and make decisions<br>Professional-level gaming skills<br>Rich responses and unpredictable | Training AI is difficult and costly<br>AI decision-making with imperfect reward and punishment mechanisms is disappointing |

# 4 The Applications Cases of RL in Game Field

## 4.1 Application of AI technology in football game

The Google team launched Google Research Football (GRF), which provides a physics engine-based 3D soccer simulation to further advance the RL algorithm (Fig.2). Each GRF game consists of 3,000 steps and is divided into the first and second halves, players need to control one or more players against the enemy, the rules are the same as in a general football game, each step decision players can choose one of a total of 19 actions such as moving, passing, shooting, dribbling, tackles and sprinting to control the player.

TiKick uses the offline RL method to learn from the data of tens of thousands of WeKick self-games, learns the single agent strategy from the data of WeKick through behavioral cloning, and applies it to all players. Subsequently, add build in to the action set to control inactive players in the backcourt. Next, the Multi-Agent Behavioral Cloning (MABC) algorithm is used to train the model, selecting matches with a high number of goals from the dataset, and using these high-quality data to accelerate convergence and improve performance [5]. And train a critical network to score all actions, using its results to calculate.

The AI trained by TiKick is the first team that can complete the whole game. Their offline RL results can also be used as a pre training model to accelerate the Rate of convergence of subsequent RL training and improve its performance.

**Fig. 2.** The picture shows the Tikick match moment [5]

### 4.2 Application of AI technology in MOBA game

Compared to board games, MOBA games are much more complex, resulting in a significant increase in the difficulty of AI action prediction and decision-making. Tencent's team trained a 1v1 AI model for Honor of Kings, and achieved excellent results in real people's confrontations (Fig.3).

Tencent's experimental team trained an AI model, they divided their deep learning model into four sub modules - Artistic Intelligence Server, Dispatch Module, Memory Pool, and RL Learner. In terms of algorithm design, an actor critical neural network was developed for modeling MOBA action decisions. The network was optimized using a multi label proximal strategy optimization (PPO) objective, and proposed decoupling methods for action dependencies, attention mechanisms for target selection, action masks for efficient exploration, Long Short-Term Memory (LSTM) for learning skill combination, and an improved version of PPO-dual clip PPO-to ensure training convergence [6-8].

During the training process, Tencent used the beating speed of a weak model to quantitatively evaluate the training results, and set an Actions Per Minute (APM) limit of every 133ms for AI. Their AI model achieved a winning rate of 0.998 in 2100 games against top amateur players and only lost one set against professional players.

With Elo rating system (ELO) rating system as the benchmark, AI's decision-making and operation level in MOBA games has reached the level of professional players.


**Fig. 3.** The picture shows Tencent's division of AI for gaming AI [6]

## 5 The Future of RL in the Game Field

In future video games, there will be more and more features such as multiplayer confrontation and team building to enhance their fun, or there will be a large number of small monsters and bosses in the game as levels [9]. These specific scenarios have always led to a demand for well-designed game AI in the gaming industry to quickly produce a large number of high-quality intelligent agents in games.

At present, the gaming level of AI can reach or even surpass that of amateur human players in various game types. In the future, how to make AI have more gaming styles to create different experiences for players in each game will become the design focus of game companies. For example, when AI is the opponent of a player, the training model will try to enrich the AI's personality as much as possible. Some NPCs will actively attack the player, while others may adopt a cautious style of hiding and ambushing the player [9]. When AI is a player's teammate, training the model will make AI more team oriented. NPC will cooperate with the player's actions, identify the player's intentions in advance, capture the player's signals, and provide different reward functions to the intelligent agent, combined with cooperation [10]. When AI is a scene NPC, it should respond differently to the player's different actions to be more like a real person, rather than using fixed actions to interact with the player.

It can be predicted that with the continuous improvement of RL in the game field, the more personified and characteristic game AI will become the selling point of new games to attract players, and the game field is also an excellent test scenario for AI. The needs of players will stimulate the development of AI and enable AI to achieve new breakthroughs.

## 6 Conclusion

This article reviews the application of RL in the field of games. Firstly, the demand for RL in the game field is summarized, the concept of RL is summarized, and then the author reviewed the application and development process of RL in the game field, and the traditional artificial intelligence and RL artificial intelligence are compared, next, the author collected examples of application of RL in the game industry to look forward to the future of RL. The author focuses on comparing and introducing examples of applied RL to let readers understand the advantages, achievements and shortcomings of RL in the field of games.

In conclusion, RL clearly holds great potential and is making significant progress in the field of gaming. However, the application of AI in games currently faces some shortcomings. First, existing AI models struggle with complex situations and long-term planning, which is especially evident in RTS games. In addition, training RL models typically requires significant computational resources and time, which may hinder their widespread adoption in real-world game development. In addition, in some AI models, over-reliance on predefined rules and algorithms can lead to a lack of creativity and adaptability when faced with unfamiliar environments.

When examining the training model, it is necessary to highlight certain deficiencies. For example, many RL algorithms are very sensitive to hyperparameters and require meticulous calibration. Failure to do so may result in slow convergence or even no convergence. In some cases, determining an appropriate reward function to drive the model toward effective learning remains a challenge.

Going forward, the author expects the development of RL in gaming to focus on the following areas:

1. Develop AI's ability to innovate and learn autonomously, thereby reducing reliance on pre-established rules and algorithms;
2. Improve and optimize training methods to minimize the burden of computing resources and time;
3. Enhancing the generalization ability of AI, enabling it to handle complex scenarios and adapt to various game types;
4. Improve swarm intelligence through multi-agent learning and collaboration.

As technology advances, the use of RL in games is likely to become more common, providing players with a richer and more diverse gaming experience.

## References

1. Kai Arulkumaran et al. “Deep reinforcement learning: A brief survey”. In: IEEE Signal Processing Magazine 34.6 (2017), pp. 26–38.
2. Bowen Baker et al. “Emergent Tool Use from Multi-Agent Autocurricula”. In: arXiv e-prints, arXiv:1909.07528.2019.

3. Tuomas Haarnoja et al. “Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor”. In: Proceedings of the 35th International Conference on Machine Learning. July 2018.
4. Peter Henderson et al. “Deep reinforcement learning that matters”. In: Proceedings of the AAAI conference on artificial intelligence. Vol. 32. 1. 2018.
5. Shiyu Huang et al. “TiKick: Towards Playing Multi-agent Football Full Games from Single-agent Demonstrations”. In: arXiv e-prints, arXiv:2110.04507, 2021.
6. OpenAI et al. “Dota 2 with Large Scale Deep Reinforcement Learning”. In: arXiv e-prints, arXiv:1912.06680, 2019.
7. Oriol Vinyals et al. “Grandmaster level in StarCraft II using multi-agent reinforcement learning”. In: Nature 575.7782 (2019), pp. 350–354.
8. Deheng Ye et al. “Mastering Complex Control in MOBA Games with Deep Reinforcement Learn-ing”. In: arXiv e-prints, arXiv:1912.09729 2019.
9. Liang W et al. “Deep Reinforcement Learning for Video Game AI: A Survey”. In: IEEE Transactions on Cybernetics, (2022), pp. 1-15.
10. Gao Y et al. “Learning from Simulation for Deep Reinforcement Learning in Video Games”. In: IEEE Transactions on Emerging Topics in Computational Intelligence, (2023), pp. 1-12.