DOI: 10.53469/jtpes.2024.04(12).02

Game AI Training Based on Reinforcement Learning and Deep Reinforcement Learning

Chenyu Shen

School of Computer and Software, Chengdu Jincheng University, Chengdu 611731, Sichuan, China

Abstract: ***This article explores the importance of game AI training as an important area of cross fusion between computer science and artificial intelligence, particularly in the core position of reinforcement learning research. Game AI training is not only a hot topic in technical practice, but also a key carrier environment for promoting innovation in artificial intelligence theory and methods. Currently, in the practical process of game AI training, we are facing many challenges, including the dual pressure of ethical considerations and technological innovation. These challenges require us to delve into key technical issues in game AI training, such as precise analysis of coefficients and delayed feedback, effective exploration of high-dimensional states and complex action spaces, and improving the robustness of strategy learning in unstable environments. In response to these challenges, this article proposes an innovative solution based on the latest developments in deep reinforcement learning. By integrating the advantages of reinforcement learning and deep learning, we have constructed a basic framework for deep reinforcement learning based on attention mechanism. This framework aims to solve the problem of cluster intelligence in complex environments, by intelligently allocating attention resources to improve the decision-making efficiency and accuracy of AI systems in processing massive amounts of information and dynamic environments. This study not only provides a new technological path for game AI training, but also provides theoretical support and practical guidance for the application of artificial intelligence in a wider range of fields.***

Keywords: Game AI training; Intensive learning; Deep reinforcement learning.

1. INTRODUCTION

The definition of game artificial intelligence (AI) training has shown a relatively broad and evolving trend in both academia and industry. This field aims to endow the gaming environment with an appropriate level of intelligence through advanced computer science and artificial intelligence technology, thereby making the gaming experience more realistic, fun, and significantly enhancing its challenge[1]. In short, game AI is a comprehensive reflection of this series of technologies and applications, which profoundly affects every aspect of game design, development, and player interaction. Within the overall framework of game AI, there are two distinct yet complementary implementation paths. Firstly, we are exploring AI based on Finite State Machine (FSM) or Behavior Tree (BT) features[2][3]. The design concept of this type of AI is relatively intuitive and easy to control, with its core being to decompose complex game behavior into a series of predictable state transitions or decision nodes. Finite state machines achieve the basic logical judgment and action selection of AI roles by defining different states and their transition conditions; The behavior tree is further refined on this basis, organizing and managing more complex decision logic and behavior sequences through a tree like structure[4]. Both methods allow developers to accurately anticipate and analyze AI behavior, thereby achieving a high degree of controllability and consistency in game design. However, with the increasing complexity of the gaming world and the growing demands of players, AI relying solely on finite state machines or behavior trees is no longer sufficient to meet the needs of all scenarios[5]. So, another type of AI based on neural networks emerged, becoming another important direction for the development of gaming AI. This type of AI simulates the working mechanism of human brain neurons, optimizes the neural network structure using genetic algorithms, and combines a large amount of data calculation and analysis to generate highly adaptive and learning capable non qualitative AI[6]. Compared to the former, AI based on neural networks is more difficult to predict and analyze in terms of behavioral performance, as they can adjust strategies in real-time according to changes in the gaming environment, and even demonstrate a certain degree of "creative" thinking. This uncertainty not only adds more surprises and challenges to the game, but also greatly enriches the player's gaming experience. It is worth noting that these two types of AI do not exist in isolation, but can be flexibly combined in game design to achieve optimal results. For example, in certain scenarios that require highly precise control, AI based on finite state machines or behavior trees can be used to ensure the stability and predictability of the game; In situations where it is necessary to showcase the complex emotions, personalities, or strategic depth of AI characters, it is more inclined to use neural network-based AI to enhance the realism and immersion of the game. In summary, game artificial intelligence training, as a bridge connecting technology and art, is constantly driving innovation and development in the gaming industry[7][8][9]. Through in-depth exploration of the integration and application of different technological paths such as finite state machines, behavior trees, and neural networks, we have reason to believe that the future gaming world will be more diverse and engaging, bringing players unprecedented gaming experiences.

2. ANALYSIS OF THE CURRENT SITUATION OF GAME AI TRAINING BASED ON REINFORCEMENT LEARNING AND DEEP LEARNING

In the development of computer learning technology, we can use reinforcement learning theory to design game AI that can transform the process of presenting intelligent characters in the entire game into a simple Markov model structure[10]. Intelligent characters use their intelligent sensing system or regional environment and their own work form, fully combining their own experience, to effectively select and analyze the operational behavior to be executed. The main implementation of this behavior affects the game system and realizes the continuous progress of the game process[11]. At the same time, feedback analysis and evaluation of the intelligent character's execution behavior are often conducted under environmental conditions. For intelligent characters, the main task objective is to obtain the optimal feedback state of the environment by obtaining the best behavioral measures. Intelligent characters executing this form of operation in the game environment can be structured into a basic Markov chain[12]. The best feedback effect obtained after a persistent loop is the main task objective of the current work. The traditional machine learning methods are mainly based on low dimensional inputs, and the actual convergence effect in this environment is relatively satisfactory[13]. However, with the gradual enhancement of the actual performance of different hardware facilities and the innovation of neural network technology itself, there are relatively more problems faced by game intelligence, and technical issues are gradually becoming more prominent. At present, they mainly focus on the following aspects.

One is that the spatial and motion space technology in high-dimensional modes faces certain bottleneck problems. At present, behavioral intelligence technology can only perform simple walking and cannot achieve effective prediction, judgment, decision-making, and other complex behavioral patterns[14][15][16][17][18]. Secondly, the feedback in the game is relatively sparse, and there is also a certain degree of delay effect. The entire game mode and process are difficult to have a direct impact on the environment in special circumstances, making the training and analysis of intelligent objects even more difficult[20][21][22]. In response to the emergence of such problems, neural network pattern optimization has become the main measure for adjusting high-dimensional action space management technology, and reinforcement learning provides a potential response path for feedback coefficients and delays. The intelligent operation mode can establish a communication efficient management mode or share relevant parameters for unstable environments, thereby achieving the best response.

3. BASIC ANALYSIS OF GAME AI TRAINING EXPERIMENTS USING REINFORCEMENT LEARNING AND DEEP REINFORCEMENT LEARNING

Reinforcement learning mainly regards the interaction between the entire intelligent agent and the environment as the overall core problem, and constructs the maximum cumulative reward method through selective behavior analysis, so as to continuously implement learning optimization measures and form the best work sequence from it. It is possible to use reinforcement learning in a game environment to obtain a strategic approach for AI behavior in the entire complex environment based on reward based training. This introduces the reinforcement learning mode of the primary system.

The construction of deep reinforcement learning mainly combines the reinforcement learning ability foundation formed by deep learning, which can demonstrate the best working effect in many environmental situations. However, in the process of facing length decision problems, it often presents poor performance. Based on the theories of deep reinforcement learning, multi-agent systems, reinforcement layering, and attention mechanisms, we aim to innovate the working mode. Reinforcement learning is the process of constructing a Kolmogorov chain for the decision definition of a single agent based on a comprehensive observation system[23]. The implementation of this process is mainly formed by constructing a tuple structure. In a special practical environment, the intelligent agent is always in a stable state environment, and through the time limit of policy execution actions, it obtains environmental feedback rewards and gains. At the same time, according to the structural form of the transition equation, it enters the next state mode. For Markov chains built around intelligent individuals, define reward feedback effects with loss coefficients in the environment and define action values. By maximizing the numerical response, we can obtain an optimal decision strategy result, so the optimal strategy function value used is constant. In the process of reinforcement learning, there is no fixed Markov process, and the agent needs to learn the optimal strategy through interaction with the environment. The widely used measure in various current environments is the combination of deep learning, which estimates the core numerical function through backup iteration[24].

The research and development in the field of functions has been relatively rapid, thanks to the research, analysis, and breakthrough presentation of deep neural network data. Combined with deep neural networks, some high-dimensional data can be directly transmitted and processed. The operation method of approximating equations with these constructed functions is included in it. At present, the Deep Q-Learning (DQN) we mentioned is a widely accepted processing method and approach, which has been widely used in different domain environments, including Go, Atari games, and so on[11][12][14]. The most specific presentation method is to extract certain empirical elements from memory in the first iteration environment to update and transmit their parameter information. During the update process, the minimum loss function equation is used to ensure the best working effect. The experience memory adopts the most advanced queue representation form, in which the intelligent exploration strategy and experience tuple content data stored can support the stable progress of later operations. The speed and frequency of updating the relevant network parameters of the target in this form are relatively low, and the experience replay learning mechanism combined with it is constructed into a stable Deep Q-Learning relationship.

In order to better address the problem of not achieving optimal work results in feedback reward coefficients or delayed environments, a reinforcement learning mode is constructed by using a specialized framework structure such as a system. This framework is defined in each time environment, and the intelligent experience selects a primitive action or measures covering multiple steps for analysis. The implementation of each strategy involves different original actions or other strategic means, while completing the progress of work tasks based on random function information. Therefore, we will extend the traditional Markov policy process into a semi Markov decision-making process to address sparse feedback and delayed feedback to various problems.

By combining the training analysis of multiple agents, an independent network baseline training method can be constructed in the existing environment, and a network environment can be built separately for each intelligent object, placing multiple intelligent objects together to complete the training operation. In this context, by analyzing the communication protocol of intelligent objects, it is possible to enable them to share some of the observation information data, achieving the goal of accurate decision-making. Perform feature analysis and extraction in convolutional neural networks, then analyze and judge intelligent objects, achieve intelligence data sharing through communication center analysis and judgment, and achieve efficient processing of attention mechanisms. At present, in order to analyze the research and analysis of intelligent objects in action games under deep reinforcement learning methods, we can integrate the working characteristics of small input in the upper layer structure, use it as the visual field of intelligent objects for image transmission feedback, and then perform feature extraction analysis of convolutional networks. The extracted feature data information is shared and analyzed towards intelligent objects through the construction of communication centers, and then the entire data information is transmitted to the main operation structure for weighting and tracking analysis. After transferring to network training mode, multiple practical planning target structures will appear. The input of the lower level structure is mostly based on the perspective of intelligent objects and the planning objectives formed by the upper level structure. It mainly involves training and analyzing the network to input relevant execution actions, which are then handed over to the intelligent objects of the game for execution, achieving integration with the environment and ultimately building task objectives. This process continuously completes training operations, with the ultimate goal of achieving optimal collaborative processing and optimal working mode strategies under policy parameters through analysis.

4. IMPLEMENTATION OF GAME AI TRAINING BASED ON REINFORCEMENT LEARNING AND DEEP REINFORCEMENT LEARNING

In order to achieve the work objectives of each subtask, it is necessary to accelerate the analysis and processing of different subtasks, so as to truly complete the relevant task content. In the reinforcement learning task objective based on random neural networks, a general skill approach in the environment is pre learned using random neural networks. At the same time, different skill approaches are adjusted for the implementation and promotion of each task training, as well as individual strategy measures. In this proposed reinforcement learning model, a dual layer structure of workers and managers is designed. Workers use the relevant requirements and transfer results of each step as the transfer result task objective of managers. Due to the prominent practical nature of managers, the exploration direction and ability of agents are comprehensively improved and optimized.

Reinforcement learning has significant advantages in complex game design environments. Due to the influence of environment specific long sequences, reward coefficients, and other characteristics in the construction of different platform environments, traditional reinforcement learning methods themselves cannot achieve the best results among game subjects. In addition, some games such as mazes and ant search for items also require special design analysis to achieve the best results due to the high complexity of different scene environments. However, the method of implementing reinforcement learning through hierarchy requires the construction of separate sub objectives in the context of task decomposition, and based on a hierarchical reward analysis. This can greatly optimize and improve learning effectiveness, implement control measures, and achieve the best task level. Therefore, reinforcement learning method is to design complex games through learning, thereby forming potential optimization response methods.

Although reinforcement learning can effectively address the core problem of decision-making in the current working mode, this structural form mostly requires strategy training methods, with relatively low sample utilization efficiency and high practical difficulty in training. In this context, combining reinforcement learning with alternative strategy training and implementing alternative strategy correction methods to replace sub target structures in the sample through maximum estimation analysis. The analysis of the experimental results shows that the basic algorithm performs better than traditional algorithm structures. However, traditional algorithms use random sampling methods to seek an approximate replacement for the calculation of sub targets. This approach cannot effectively restrict and affect the sub target space, and sub targets can still choose a meaningless or non directly implementable method, resulting in unstable working effects due to the problems existing in lower level learning.

With the development and optimization of industrial structure, the working mechanism and behavioral environment of games have become more realistic and diverse. In many environments, complex behaviors need to be designed to achieve interactive analysis between AI games. This level of technology greatly affects the quality of game products themselves and the gaming experience of users. The design of such AI games is often time-consuming and labor-intensive, posing a huge challenge for game developers. This invention is based on reinforcement learning methods and decomposes the existing behavioral means of artificial intelligence and environmental interaction into a series of small task target themes. In the process of constructing the Q-table, accelerate the innovation of task goal construction to ensure the orderly completion of various tasks.

5. CONCLUSION

In order to achieve analysis of complex environments and effective judgment of complex behavioral decision-making problems, it is necessary to utilize relevant device engines and plugin information to develop and construct AI adversarial intelligent systems. In this context, the framework structure constructed in this article itself has distinct advantages. Only by doing a good job in game AI training with reinforcement learning and deep reinforcement learning can the best work processing effect be achieved to fully promote the game, reduce the direct impact of different hidden problems on game implementation, and improve the game experience.

REFERENCES

[1] Chen, H., Shen, Z., Wang, Y., & Xu, J. (2024). Threat Detection Driven by Artificial Intelligence: Enhancing Cybersecurity with Machine Learning Algorithms.

[2] Jurafsky, D., & Martin, J. H.. (2007). Speech and language processing: an introduction to speech recognition, computational linguistics and natural language processing. Prentice Hall PTR.

[3] Li, L., Gan, Y., Bi, S., & Fu, H. (2024). Substantive or strategic? Unveiling the green innovation effects of pilot policy promoting the integration of technology and finance. International Review of Financial Analysis, 103781.

[4] Bethard, S., Jurafsky, D., & Martin, J. H.. (2008). Instructor's Solution Manual for Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (Second Edition).

[5] Yao, J. (2024). The Impact of Large Interest Rate Differentials between China and the US bn the Role of Chinese Monetary Policy -- Based on Data Model Analysis. Frontiers in Economics and Management, 5(8), 243-251.

[6] Z. Ren, "Enhancing Seq2Seq Models for Role-Oriented Dialogue Summary Generation Through Adaptive Feature Weighting and Dynamic Statistical Conditioninge," 2024 6th International Conference on

Communications, Information System and Computer Engineering (CISCE), Guangzhou, China, 2024, pp. 497-501, doi: 10.1109/CISCE62493.2024.10653360.

[7] Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences, 622, 178-210.

[8] Wang, Z., Zhu, Y., He, S., Yan, H., & Zhu, Z. (2024). LLM for Sentiment Analysis in E-commerce: A Deep Dive into Customer Feedback. Applied Science and Engineering Journal for Advanced Research, 3(4), 8-13.

[9] Jurafsky, D., & Martin, J. H.. (2007). Speech and language processing: an introduction to speech recognition, computational linguistics and natural language processing. Prentice Hall PTR.

[10] Tian, Q., Wang, Z., Cui, X. Improved Unet brain tumor image segmentation based on GSConv module and ECA attention mechanism. arXiv preprint arXiv:2409.13626.

[11] Zhu, Z., Wang, Z., Wu, Z., Zhang, Y., & Bo, S. (2024). Adversarial for Sequential Recommendation Walking in the Multi-Latent Space. Applied Science and Biotechnology Journal for Advanced Research, 3(4), 1-9.

[12] Bethard, S., Jurafsky, D., & Martin, J. H.. (2008). Instructor's Solution Manual for Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (Second Edition).

[13] Wyard, P. J.. (1992). Connectionist natural language processing: an introduction. Springer Netherlands.

[14] Xu Y, Shan X, Guo M, Gao W, Lin Y-S. Design and Application of Experience Management Tools from the Perspective of Customer Perceived Value: A Study on the Electric Vehicle Market. World Electric Vehicle Journal. 2024; 15(8):378. https://doi.org/10.3390/wevj15080378

[15] Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W.. (2011). Natural language processing: an introduction. Journal of the American Medical Informatics Association Jamia, 18(5), 544.

[16] Teller, & Virginia. (2000). Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition daniel jurafsky and james h. martin (university of colorado, boulder) upper saddle river, nj: prentice hall (prentice hall ser. Computational Linguistics, 26(4), 638-641.

[17] Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W.. (2011). Natural language processing: an introduction. Journal of the American Medical Informatics Association Jamia, 18(5), 544.

[18] Xie, Y., Li, Z., Yin, Y., Wei, Z., Xu, G., & Luo, Y. (2024). Advancing Legal Citation Text Classification A Conv1D-Based Approach for Multi-Class Classification. Journal of Theory and Practice of Engineering Science, 4(02), 15–22. https://doi.org/10.53469/jtpes.2024.04(02).03

[19] Teller, & Virginia. (2000). Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition daniel jurafsky and james h. martin (university of colorado, boulder) upper saddle river, nj: prentice hall (prentice hall ser. Computational Linguistics, 26(4), 638-641.

[20] Shen, Z., Ma, Y., & Shen, J. (2024). A Dynamic Resource Allocation Strategy for Cloud-Native Applications Leveraging Markov Properties. International Journal of Advance in Applied Science Research, 3, 99-107.

[21] Z. Ren, "A Novel Feature Fusion-Based and Complex Contextual Model for Smoking Detection," 2024 6th International Conference on Communications, Information System and Computer Engineering (CISCE), Guangzhou, China, 2024, pp. 1181-1185, doi: 10.1109/CISCE62493.2024.10653351.

[22] Belhaouari, S. B., Ahmed, S., & Mansour, S. (2014). Optimized K‐Means Algorithm. Mathematical Problems in Engineering, 2014(1), 506480.

[23] Santhi, V., & Jose, R. (2018). Performance analysis of parallel k-means with optimization algorithms for clustering on spark. In Distributed Computing and Internet Technology: 14th International Conference, ICDCIT 2018, Bhubaneswar, India, January 11–13, 2018, Proceedings 14 (pp. 158-162). Springer International Publishing.

[24] Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. Electronics, 9(8), 1295.