ISSN 2959-6157

Application of Machine Learning in Human-Computer Intelligent Confrontation

Liwen Xing

School of Mathematical Sciences, Tongji University, Shanghai, China
*Corresponding author: 2053307@tongji.edu.cn

Abstract:

Human-computer intelligent confrontation is a widely used technique in a lot of game scenarios, and the artificial intelligence makes a lot of success in different fields. The level of artificial intelligence improves quickly with the development of machine learning. Machine learning refers to enhancing the capabilities of artificial intelligence through the interaction of machines and data. Among all kinds of machine learning, the most famous one is deep reinforced learning. This technology is currently widely used in human-computer confrontations. This article introduces the technical principles of machine learning and reviews the development history of machine learning theory. Besides, this article introduces the application of machine learning in the two important competitive fields of Go and Texas Hold'em, as well as its development and achievements in these two areas. This article can enable other scholars to have a more comprehensive and systematic understanding of machine learning and its applications in Go and Texas Hold'em.

Keywords: Machine learning; AlphaGO; Texas hold'em; Human-computer intelligent confrontation.

1. Introduction

1.1 Background

Human-computer intelligent confrontation is an important area in computer science and artificial intelligence. It mainly studies the basic theories and methods of machine intelligence defeating human intelligence through human-machine and machine-machine confrontation in different game scenarios, such as Go and Texas Poke game [1]. Scientists want to know how to use intelligent technology to enable artificial intelligence to defeat humans in the confrontation between humans and computers in each game area. In 1943, McCulloch proposed the "Artificial Neural Network Model". From then on, the human-computer intelligence confrontation developed very quickly and made a groundbreaking achievement in a lot of game areas, such as AlphaGo in Go and Deep Blue Computer in chess. Nowadays, Artificial intelligence has beaten humans in a lot of fields, such as chess, card games, and even computer games. The most famous technique in human-computer intelligent confrontation is machine learning, and one of the most famous and useful machine learning is deep reinforcement learning. By giving artificial intelligence data, machines could enhance one's computing ability by learning these data. Ultimately, artificial intelligence will utilize the data provided by humans to surpass human capabilities.

1.2 Literature Review

In previous studies, scholars have conducted in-depth analyses of the specific principles and algorithms in the field of human-computer gaming. In the field of deep reinforcement learning, some scholars have studied double permutation table optimization algorithms. Permutation table based on Zobrist hash technology is a common cache acceleration method. The double-layer permutation table can solve the problem of updating records that are relatively shallow and cannot be stored in the table. At the same time, avoid storing too many expired nodes using a single depth-first replacement strategy [2]. This technology can enhance the computing power and efficiency of artificial intelligence. In the field of computer gaming, some scholars have studied the minimax algorithm. The minimax algorithm is common in computer game analysis. In practical applications, it can form a game tree by using the minimax values [3]. Optimizing this algorithm can yield the optimal search algorithm, which is a widely used algorithm in machine learning. In the field of Go, some scholars have studied how to use game trees to solve computer Go-capturing problems [4]. Common search algorithms include Alpha-Beta Search and pn Search. In Texas Hold'em, some researchers studied the Counterfactual Regret Minimization (CFR) algorithm and deep reinforcement study algorithm [5]. In summary, scholars have researched the basic theory in human-computer intelligent confrontation and its application in Go and Texas Hold'em.

1.3 Motivation

Many scholars have conducted research in the human-computer intelligence confrontation field and made achievements. These scholars' research often focuses on a theory, so it is hard for others to have a more comprehensive understanding of this field. This article summarizes the important theories and achievements in the field of human-computer intelligent confrontation, which is beneficial for others to have a preliminary understanding of this field.

2. Methodology

2.1 Machine Learning and Game Theory

The ability of artificial intelligence improves to such a big degree with the help of the development of machine learning. Machine learning mainly studies how to enhance capabilities by enabling machines to interact with data. For example, when researchers train AlphaGo, they give AlphaGo many chess manuals of the top Go masters, and after AlphaGo studies these chess manuals, its ability to Go will improve. In the early times, traditional machine learning paid more attention to single people and their interaction with the environment. But this kind of machine learning will ignore other machines in the environment and cause damage to the stationary of the environment as other machines are also intelligent and their action will change with time passing [6].

To study the multi-agent situation occurring in the human-computer intelligence confrontation, some researchers use game theory as the game theory studies strategic interactions between multiple self-interested individuals. Nowadays, with the development of game theory and computer science, scientists have had a lot of success in a lot of areas, such as chess, Poke, and even computer games. Traditional machine games are divided into perfect information games, such as Go and chess, and imperfect information games, such as Texas Poke [7]. Perfect information games mean that each participant accurately grasps the actions of other participants when taking actions, as well as the actions of other participants before taking actions. For example, when two players play Go, each player knows all the information about the game. Imperfect information games mean that no participant was able to obtain action information from other participants. For example, a player could not know information about the cards in another player's hands in a Texas poker game. The current development of intelligent game adversarial technology based on algorithmic game theory, deep reinforcement learning, and online convex optimization is fully capable of solving zero-sum perfect information games between two players such as Go, and research on imperfect information games has also made significant progress. A lot of artificial intelligence has beaten humans in both perfect information games and Imperfect information games.

2.2 Deep Reinforce Learning

One of the most well-known machine learning methods is deep reinforcement learning. Reinforcement learning includes four parts: environment, state, reward, and action. Under a certain state, the artificial intelligence will choose and execute a certain action according to a certain strategy. Then this state will transition to the next state with a certain probability. At the same time, the environment feeds back a reward signal to the artificial intelligence. In multiple interactions with the environment, artificial intelligence accumulates rewards by adjusting its strategies to reach their maximum value and evaluate the effectiveness of the strategies. Reinforcement learning has a long history of almost 90 years and plays an important role in human-computer intelligent confrontation.

3. Literature References

3.1 The Solution of the Computer Go Capturing Problems

Go is one of the oldest and most complex board games in the world. Go uses a 19 * 19 rectangular grid board and black and white circular pieces to play, with each side taking turns placing one piece. As there are 361 feasible places on the board, every move of a chess player has many choices. The artificial intelligence that runs Go requires powerful computing and judgment abilities. The first thing that needs to be solved is the problem of eating in Go. The purpose of eating calculation is to determine whether a piece of chess can be eaten and how to eat it. Eating subroutines is fundamental and crucial for computer Go programs. Almost all computer Go programs contain eating submodules. Before 2007, scientists have already solved this problem by using the Game Tree Search Algorithm. Fig.1 is an example. The goal of black chess is to eat white chess with triangular symbols. The seven numbers are all search solutions of artificial intelligence, and the number 2 is the correct answer. The two different search algorithms, Alpha-beta search, and proof-number search could select multiple candidates' moves and calculate the correct move from them. These two methods each have their advantages and disadvantages. But in most situations, the results of the proof-number search are better than the Alpha-beta search.

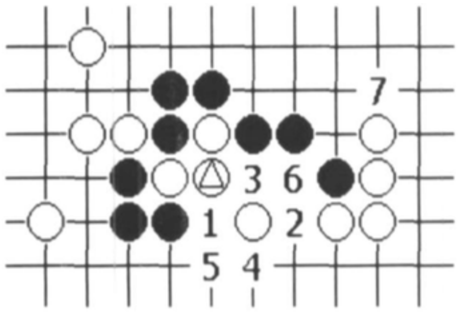

Fig. 1 One example of capturing a problem [4]

3.2 Artificial Intelligence Surpasses Humans in Go

Although scientists finished this question before 2007, the calculation of computer Go programs still lagged far behind compared to top human players in 2007. In the 2015 World Computer Go Championship, South Korean Go software challenges Chinese professional Go players, but the Go software still loses despite being given four or five pieces [8]. Significant achievements occurred in 2016. The artificial intelligence AlphaGo made by Google beat Lee Se-Dol, the best Go player at that time, with a total score of 4-1, which means that artificial intelligence has defeated humans in Go. Unlike other AI, AlphaGo integrates deep learning, reinforcement learning, and game theory very well. At first, it could study a lot of chess manuals by using deep study. Then AlphaGo applies reinforcement learning to gain more chess games by playing against oneself. Next, AlphaGo will use deep learning techniques to evaluate the win-loss ratio of each pattern. At last, the optimal location is determined through a Monte Carlo tree search [9].

3.3 Development in Nowaday

From then on, artificial intelligence develops very quickly and leads humans a lot. After that, Alpha Zero, an improved version of AlphaGo, was created. Compared with AlphaGo, Alpha Zero does not need to learn from chess manuals from human players [10]. It used a completely independent reinforcement learning algorithm, which means that Alpha Zero could improve itself just after knowing the basic rules of Go without the help of the Human Go Game Manual Training. This is a big access because even every decision made by top players may not necessarily have the highest winning rate. Alpha Zero also replaces the policy network and value network with a neural network. It abandons the Monte Carlo method, using only a simplified version of the tree search to estimate the position and win rate of drops. Nowadays, the most widely used AI go player is Fine Art made by Tencent. Fine Art won the championship of the 10th UEC Cup World Computer Go Conference in 2017 and the World Intelligent Go Open Championship in 2019. Fine Art is often used in researching and forecasting a real-time competition.

4. Applications in Texas hold'em

4.1 The Development of Artificial Intelligence in the Field of Texas Hold'em

Texas hold'em has a history of over 100 years and is one of the most famous poker games in the world. Each player will be given two cards and five common cards will be displayed one by one in the following stages. These public cards can be used by players to form the best combination of 5 playing cards with their hands. In the game, players compete for chips by comparing the combination size of their opponent's cards with the public cards. The ultimate winner will win all the chips in the bottom pool. As players could not know the cards in other players' hands, this game is an imperfect information game. The world's first Texas Hold'em intelligent AI was created by the University of Alberta in Canada. This program could take a certain action under a given situation. After that, the ability of artificial intelligence improved quickly. In 2015, the AI poker player Cepheus made by Bowling first cracked the winning method of two-person limited bet Texas Hold'em [11]. The AI poker player Libratus beat the best human player in a two-person limited bet Texas Hold'em in 2017. After that, the artificial also beat humans in multi-person Texas Hold'em: the AI Texas player Pluribus defeated 5 top human players in a six-person table game and achieved victory in 2019.

4.2 The Most Famous Artificial Intelligence in Texas Hold'em

The most widely used technique in Texas Hold'em AI is the counterfactual regret minimization (CFR) algorithm. This is an iterative algorithm, and it could decompose the overall minimum regret value into independent information sets to calculate the local minimum regret value. Using this method to converge to a Nash equilibrium solution has theoretical guarantees in a large-scale two-player zero-sum game. The most well-known AIs are Cepheus, Libratus, and Pluribus, and all of them use the CFR algorithm. Cepheus solves the HUL problem, which means Two-person limited bet Texas Hold'em, for the first time. Cepheus uses fixed-point computation and compression to solve storage problems and combines pruning and omission to calculate the average strategy, breaking through the limitations of the CFR algorithm in terms of computational scale. Although it could lose in a single game, it could win in a long-term game. Libratus is the first AI poker player which beat the humans. It has three modules. The first module is used to abstract actions and hand abstraction in game theory. The second module is an embedded sub-game solving algorithm. The last module is a self-improve module and Libratus could improve itself as time goes by. Pluribus is improved from Libratus, but it has some differences. The goal of Pluribus is not a specific game theory concept, but to solve multiplayer poker problems from an experimental perspective. The fact shows that it could make a good strategy and beat humans. An important point is that Pluribus adopts a new online search method and discovers better strategies in the competition through real-time search. These new techniques helped Pluribus beat 5 top human players in 2019.

5. Limitations and Future Overlook

Nowadays, Scientists have made many groundbreaking achievements in the human-computer intelligent confrontation field. At the same time, the computing power level of artificial intelligence is also increasing from generation to generation. Artificial intelligence has beaten humans in all intellectual-competition games. In Go chess, artificial intelligence has never been beaten by the human player after Lee Se-Dol, and the best artificial intelligence now could let the best human player take two to three moves in a row at the beginning and then beat the human player. When it comes to Texas Hold'em, the artificial intelligence has beat the human player in almost all kinds of rules. However, there are still some limitations in this field. Some algorithms still have vulnerabilities. For example, modern Go artificial intelligence may still make some misjudgments in certain special chess manuals. Sometimes they even mistake a lost game for a winning game. Moreover, when the situation is at a disadvantage, artificial intelligence cannot make a threatening counterattack. In Texas Hold'em, the consumption of computing and storage resources in artificial intelligence is enormous. For example, the operating cost of Libratus is approximately $7000 per day, with a storage space of 2.6PB. Besides, the CFR algorithm adopts the same strategy when facing different opponents and cannot adjust its strategy to restrain the opponent. In future research, scientists could optimize algorithms and invent better search and game strategies, as this can reduce computation time and workload, and save storage space. Moreover, researchers also need to improve programs and algorithms and address potential vulnerabilities. In theoretical research, scientists can study new theories and solve problems that current theories cannot solve. This will help to invent faster and smarter artificial intelligence.

6. Conclusion

This article introduces the theory of human-computer intelligent confrontation and its application in different areas. Machine learning, especially deep reinforcement learning, plays an important role in the development of artificial intelligence and human-computer intelligent confrontation. In Go chess, the artificial intelligence AlphaGo beat the best human player. In Texas Hold'em, different artificial intelligences have beaten the best human players in different rules. Although there are still some problems in artificial intelligence, they will be solved in the future with the fast development of the technique. This article helps people to briefly understand the important theories and achievements in this field.

References

[1] Dong shi, Lu Xiao Bin. Research on Key Issues in Human-Machine Intelligence Game Confrontation. Proceedings of the 12th China Command and Control Conference (Volume I), 2024, 322-326.

[2] Wang Dong Nian, Wang Jun Wei, Xue Shi Chao, Wang Chao, Xu Chang Ming. Research on double permutation table optimization algorithm based on deep reinforcement learning. Journal of Chongqing University of Technology (Natural Sciences), 2024, 05: 145-153.

[3] Miao Sha. Research and improvement of search algorithms in computer games. China New Communications, 2023, 10: 61-63.

[4] Zhang Pei Gang, Chen Ke Xun. Using different search algorithms to solve computer Go capturing problems. Journal of Intelligent Systems, 2007, 03: 84-90.

[5] Han Sheng Yu. Research on the strategy model of Texas hold'em game. Master 2023

[6] Zhou Lie, Yin Qi Yue, Huang Kai Qi. Game-theoretic learning in human-computer gaming. Journal of Computer Science, 2022, 09: 1859-1876.

[7] Luo Jun Ren, Zhang Wan Peng, Su Jiong Ming, Wei Ting Ting, Chen Jing. Research progress on sequential imperfect information game solving in computer games, Control and Decision making, 2023, 10.

[8] Xie Chang Hong. The Application of artificial intelligence in Go research. Information Systems Engineering, 2019, 01: 105.

[9] An Bo. Artificial intelligence and game theory: Starting from AlphaGo, China Development Observation, 2016, 06: 13+17.

[10] Du Kang Hao, Song Rui Zhuo, Wei Qing Lai. Review of reinforcement learning applications in machine games, Control Engineering, 2021, 10: 1998-2004.

[11] Yuan Wei Lin, Liao Zhi Yong, Gao Wei, Wei Ting Ting, Luo Jun Ren, Zhang Wan Peng, Chen Jing. Survey on intelligent game of computer poker, Chinese Journal of Network and Information Security, 2021, 05: 57-76.