DOI: 10.54254/2755-2721/83/2024GLG0061

Leveraging LSTM and NEAT for enhanced performance in multi-agent evolutionary games

Ruixiang Gao

International School, Beijing University of Posts and Telecommunications, Beijing, 100876, China

2021212730@bupt.edu.cn

Abstract. Nowadays, the application of deep learning in the field of evolutionary games has become a very popular topic. Humans use artificial intelligence as a powerful tool to predict the decision-making of multiple agents and to analyze it thoroughly, which can significantly reduce people's workload. Two of the most typical situations in game theory are the Iterated Prisoner's Dilemma (IPD) and the Iterated Snowdrift (ISD) games. In this paper, the Neuro Evolution of Augmenting Topologies (NEAT) algorithm is employed to perform population evolution in these two scenarios, and the Long Short-Term Memory (LSTM) model is utilized to predict the behavior of the population. The unique structure of the LSTM model contributes to its excellent predictive performance in forecasting the behavior of populations. Furthermore, this paper also investigates the changes in population intelligence and the frequency of cooperative behaviors during the process of population evolution, in order to explore the trends and specific proportions of different strategies as the population evolves.

Keywords: Long short-term memory, Iterated Prisoner's dilemma, Iterated snowdrift, Neuro evolution of augmenting topologies.

1. Introduction

In the field of artificial intelligence, the study of multi-agent systems (MAS) has been widely extended to a variety of game theory and complex strategy contexts, aiming to explore how agents make decisions in various interaction and competitive situations [1,2]. The Iterated Prisoner's Dilemma (IPD) and the Iterated Snowdrift (ISD) game are classic frameworks for studying cooperative and adversarial behavior, which provide an opportunity to understand how agents adjust their strategies in repetitive games. The purpose of this paper is to explore and realize the evolution of agent behavior in these two game scenarios by combining the advanced evolutionary algorithm Neuro Evolution of Augmenting Topologies (NEAT) and the powerful sequence prediction model Long Short-Term Memory (LSTM) [3,4].

As an effective neural network evolution mechanism, the NEAT algorithm allows the optimization and adjustment of network structure through natural selection and genetic algorithm principles. This algorithm is characterized by the ability to automatically increase or decrease the number of nodes and connections in the network during the evolutionary process, so as to continuously adapt and explore new policy spaces. Compared with traditional fixed-topology neural networks, NEAT provides greater flexibility and adaptability, making it an ideal tool for studying multi-agent games [1,5].

DOI: 10.54254/2755-2721/83/2024GLG0061

On the other hand, LSTM models are widely used in sequence prediction tasks because of their excellent time series data processing capabilities, especially when dealing with complex sequences with long-term dependencies [6]. In iterative games, agents must predict their future behavior based on the historical behavior of their opponents, which is a direct fit for the strength of the LSTM model. By combining NEAT and LSTM, this study aims to develop a new type of game agent that can not only learn the opponent's strategy, but also predict and adapt to the opponent's behavioral changes in iterative games.

This paper will first review the theoretical background of IPD and ISD and their applications in multi-agent systems [7,8]. Subsequently, the working principle of the NEAT algorithm and its application in multi-agent games will be introduced in detail, as well as how the LSTM model can be used to enhance the prediction ability of agents. In the methodology section, this paper will show how to combine the NEAT algorithm with the LSTM model, and how to set up experiments to evaluate the effectiveness and efficiency of the proposed method. Through the experimental results, this work will show the performance of the proposed method in different game environments, and compare it with the existing methods to verify its superiority and practicability. Finally, the significance of the findings will be discussed, as well as possible future research directions.

By deeply exploring the application of NEAT and LSTM in multi-agent games, this study not only expands the boundaries of these two technologies, but also provides new perspectives and tools for solving complex game problems. It is hoped that the research in this paper can provide valuable reference and enlightenment for the study of strategy learning and decision-making process in the field of artificial intelligence.

2. Method

2.1. Principle of IPD and ISD

In the exploration of multi-agent systems (MAS), the Iterated Prisoner's Dilemma (IPD) and Iterated Snowdrift Game (ISD) serve as critical frameworks for understanding the dynamics of cooperation and competition. These games are central to studying strategic interactions and the behavioral evolution among autonomous agents, offering insights into the complex decision-making processes in these environments.

The IPD is an extension of the classic Prisoner's Dilemma, where two players decide independently to cooperate or defect, repeatedly interacting which allows them to adapt strategies based on the outcomes of previous rounds. This repetition introduces a dynamic component, enabling the evolution of strategies that maximize personal gains while considering the benefits of mutual cooperation or the risks of mutual defection. On the other hand, the ISD differs in its payoff structure; cooperating when the other defects still yield some benefit, unlike in IPD where mutual defection results in the lowest payoff. This setup models scenarios where the cost of helping is less than the benefit to the recipient, making cooperation potentially more attractive even in the presence of defection.

Both IPD and ISD are instrumental in studying how agents adapt their strategies within MAS, crucial for developing AI systems that dynamically adjust to changing conditions or strategies of other agents. These games also help in understanding the evolution of cooperation among self-interested agents. By simulating various strategies and altering game conditions, researchers can observe the emergence and stabilization of cooperative behaviors. Furthermore, these theoretical tools assist in predicting agent behaviors in more complex scenarios, thereby improving the efficiency and effectiveness of MAS in real-world applications. The interconnected nature of agent interactions in MAS analyzed through these games also offers insights into how local interactions can influence global behaviors, providing a deeper understanding of network dynamics within MAS.

2.2. Neuro evolution of augmenting topologies

The NEAT algorithm starts with a simple initial population of minimal neural networks, evolving these structures over time through genetic mutations and crossover operations. Its structure is demonstrated

DOI: 10.54254/2755-2721/83/2024GLG0061

in Figure 1. The genetic mutations include adding new nodes by splitting existing connections, introducing new connections between previously unconnected nodes, and modifying connection weights either by perturbation or complete reassignment. During crossover, NEAT aligns genomes by their innovation numbers and inherits genes randomly from both parents, which ensures effective mixing of genetic material.

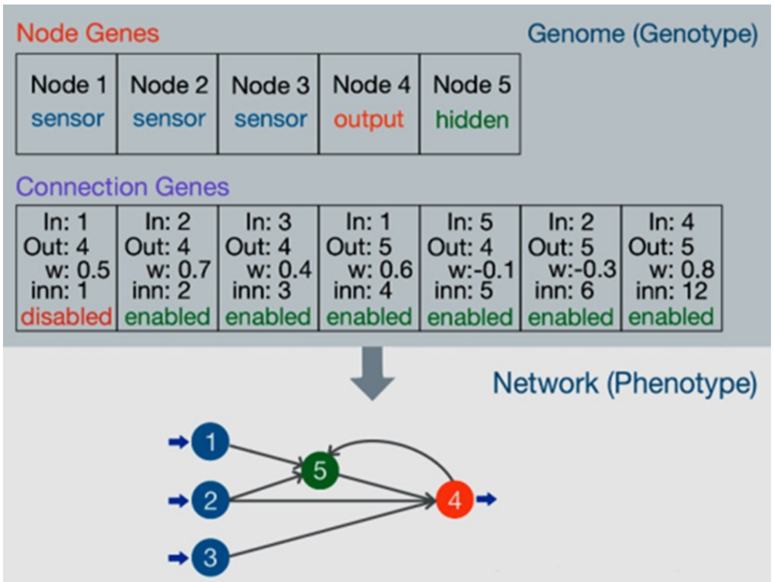

Figure 1. Basic structure of NEAT [9].

To protect and nurture novel configurations, genomes are grouped into species based on genetic similarity. This speciation allows innovations, such as new nodes or connections, to stabilize and optimize within their niche before competing across the broader population. This evolutionary mechanism helps NEAT maintain diversity within the population, crucial for exploring and exploiting new network architectures.

The diagram illustrates these principles, showing both the genotype representation of the network, which includes node and connection genes, and the phenotype, which is the actual operational neural network. Enabled connections define active pathways from input nodes to output nodes, including hidden layers that evolve over time, showing NEAT's dynamic ability to adapt and optimize network structures for varying tasks in uncertain environments.

2.3. Long short-term memory

LSTM is a type of recurrent neural network (RNN) designed to address the vanishing gradient problem inherent in standard RNNs, thereby enhancing the network's ability to learn long-range dependencies in sequential data [4,10]. An LSTM unit features a unique structure comprising a cell, an input gate, a forget gate, and an output gate, which work together to regulate the flow of information.

The LSTM cell acts as a conveyor belt, carrying relevant information throughout the sequence processing, allowing it to be accessed later. The input gate controls the extent of new information added to the cell state, while the forget gate decides what part of the existing information is kept or discarded, helping the network to focus on relevant data by removing less useful information. The output gate then controls the output flow from the cell state to the next hidden state, determining what is to be carried forward. This intricate gating mechanism enables the LSTM to effectively manage data through sequences, maintaining a balance between retaining old information and adding new, relevant data.

In the context of evolutionary game theory, where understanding the evolution of strategies within populations is critical, LSTMs offer significant advantages. Their ability to remember and accurately model long sequences of past interactions allows them to predict future shifts in strategies and

DOI: 10.54254/2755-2721/83/2024GLG0061

interactions more effectively than traditional models. This makes LSTMs highly suitable for analyzing and predicting complex, dynamic interactions in evolutionary games, optimizing strategies over time by learning from sequences of past outcomes. The diagram associated with this explanation illustrates the dynamic flow and updates within an LSTM unit, highlighting how it processes inputs and updates its states through its multiple gate mechanisms, essential for tasks that require a deep understanding of temporal dependencies.

3. Result

In this study, the LSTM model was employed to train and predict population behavior during the evolutionary process. Referring to Figure 2, it is evident that the prediction accuracy of the model progressively improves and stabilizes around 0.8 after 40-50 epochs. This performance metric indicates that the LSTM model is particularly effective in making reliable predictions. Furthermore, this research delved into the model's effectiveness in the Iterated Prisoner's Dilemma (IPD) and Iterated Snowdrift (ISD) scenarios as depicted in Table 1. The results show that the model's predictions for the IPD scenario are consistently accurate, demonstrating minor variations among different runs. In contrast, predictions for the ISD scenario display more significant fluctuations, reflecting the inherently unpredictable nature of this game scenario. The variability and trends in these predictions across different generations are detailed in Figure 3, Figure 4, and Figure 5, illustrating how the model adapts and refines its strategy over time. This analysis highlights the LSTM's capability to handle complex, dynamic systems and adapt its predictions based on evolving game dynamics and player strategies.

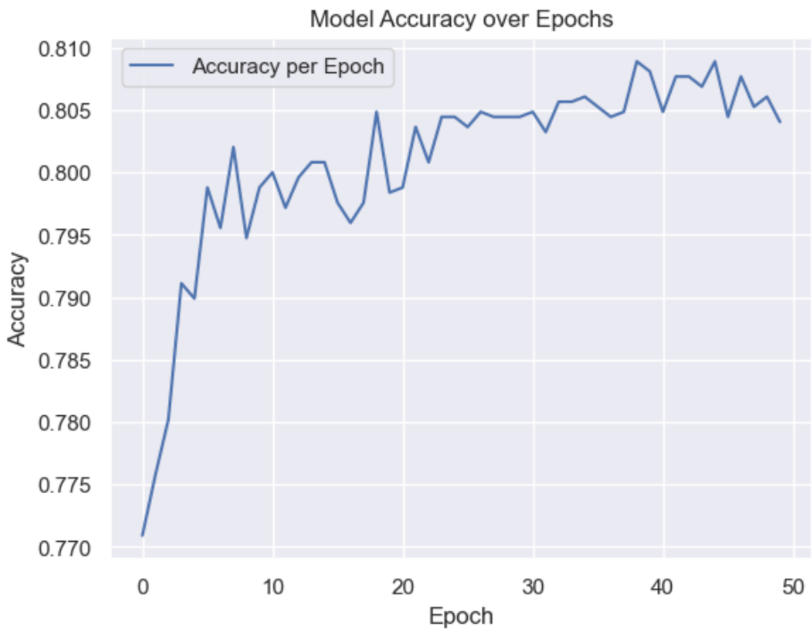

Figure 2. Accuracy over epochs (Figure Credits: Original).

Table 1. Performance measured by IPD and ISD.

	IPD	ISD
Actual Payoffs	10978	8503
Predicted Payoffs	10805	9192
Difference	173	-689

DOI: 10.54254/2755-2721/83/2024GLG0061

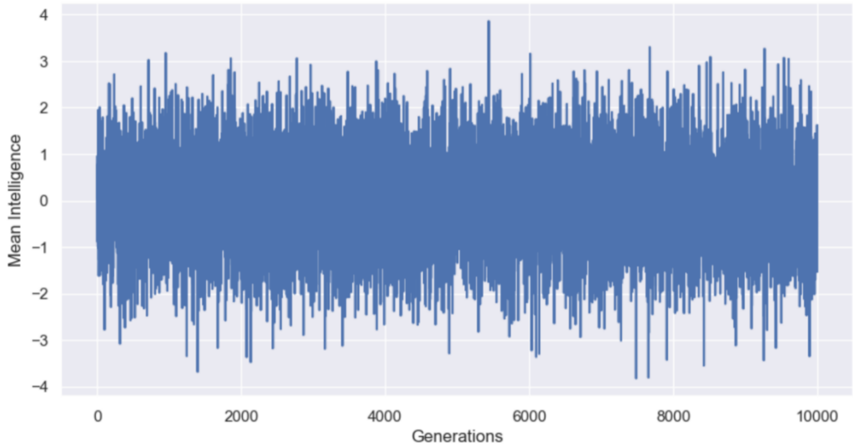

Figure 3. Change of mean intelligence over generations (Figure Credits: Original).

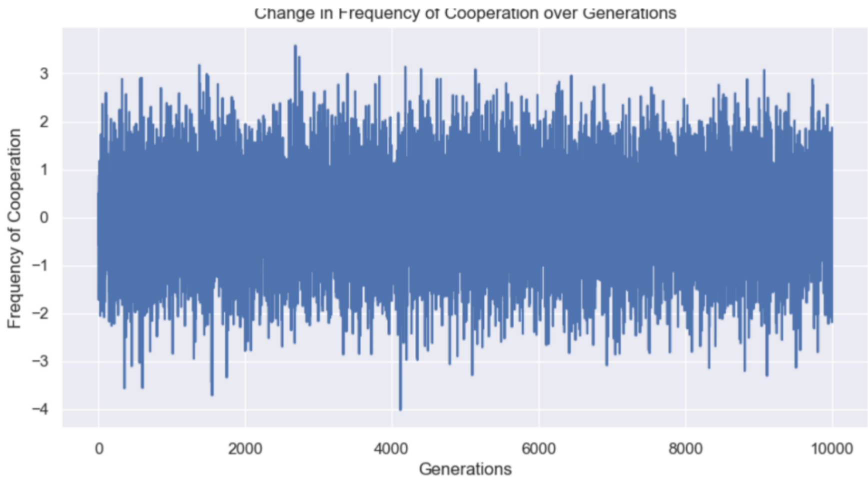

Figure 4. Change of frequency over generations (Figure Credits: Original).

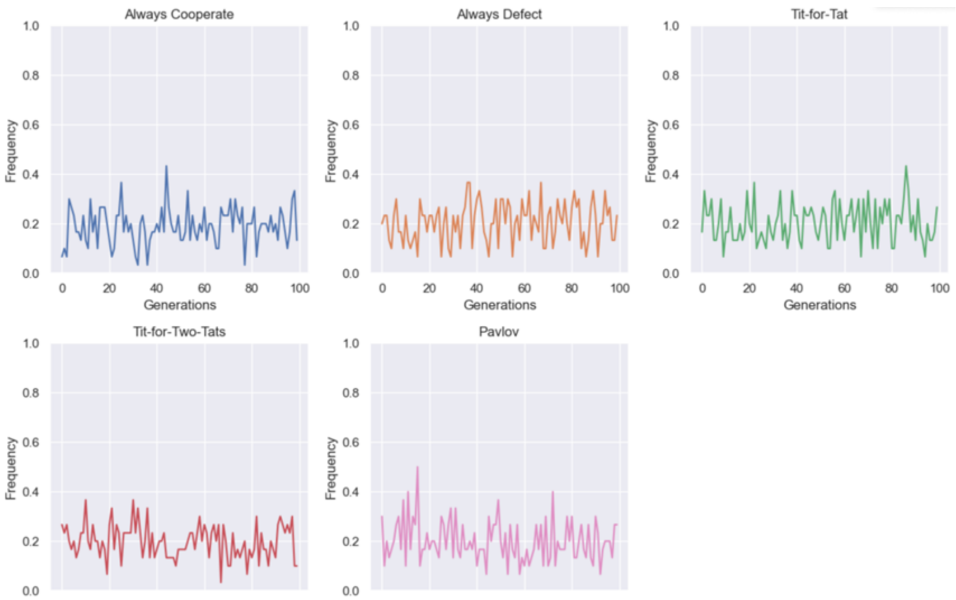

Figure 5. Change of other indexes over generations (Figure Credits: Original).

4. Discussion

Before delving into the application of NEAT algorithm and LSTM model in multi-agent game environment, it is necessary to understand the existing research results and insights provided in this field. Historically, the study of MAS has focused on how to optimize the decision-making and behavior patterns of agents in complex and dynamically changing environments.

Applications of evolutionary algorithms: The NEAT algorithm proposed by Kenneth O. Stanley and Risto Miikkulainen in 2002 is a significant improvement over traditional evolutionary algorithm [3]. The NEAT algorithm not only adjusts the weights in the neural network, but also optimizes the network structure itself to adapt to the changing environment and task requirements. This innovation provides a more flexible and scalable solution for multi-agent games, capable of naturally handling scenarios of increasing complexity. In addition, the NEAT algorithm effectively maintains genetic diversity through its unique species division and gene coding technology, thus preventing the problem of premature elimination of excellent innovations.

Impact of Long Short-Term Memory Networks (LSTMs): In 1997, the LSTM model proposed by Hochreiter and Schmidhuber greatly advanced the study of sequence data processing [4]. Through its internal gating mechanism, LSTM effectively solves the gradient vanishing or explosion problem encountered by traditional recurrent neural networks when processing long sequences. In multi-agent systems, these properties of LSTMs make them an ideal tool for predicting agent behavior, especially in game theory environments where historical behavior needs to be analyzed to predict future actions.

Game Theory and Agent Behavior: Game theory frameworks, such as the IPD and ISD, have been the main methods for studying cooperative and competitive behavior in the study of multi-agent systems. Through these frameworks, researchers are able to explore how agents adjust their strategies in repeated interactions, and how these strategies affect group dynamics and evolutionary processes.

Integration perspective: In recent years, research has shown that a single model often struggles to adequately capture all the dynamics in a complex environment. Therefore, the research on multi-model ensemble methods is gradually emerging. By integrating different types of models such as NEAT and LSTM, the behavior of agents in multi-agent systems can be simulated and predicted more comprehensively, thereby improving the accuracy of prediction and the stability of the system.

5. Conclusion

This work applies LSTM to tackle NEAT problems. The results show that the ability of the NEAT algorithm to optimize and adjust the network structure through natural selection and genetic algorithm principles, as well as the excellent ability of LSTM to deal with long-term dependency problems in complex sequences, enable the developed new game agent to not only learn the opponent's strategy, but also predict and adapt to the behavior changes of the opponent in the iterative game. Experimental results show that the proposed method performs well in different game environments, and has obvious advantages and practicability compared with existing methods.

References

[1] Dorri, A., Kanhere, S. S., & Jurdak, R. (2018). Multi-agent systems: A survey. Ieee Access, 6, 28573-28593.

[2] Balaji, P. G., & Srinivasan, D. (2010). An introduction to multi-agent systems. Innovations in multi-agent systems and applications-1, 1-27.

[3] Stanley, K. O., & Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. Evolutionary computation, 10(2), 99-127.

[4] Graves, A., & Graves, A. (2012). Long short-term memory. Supervised sequence labelling with recurrent neural networks, 37-45.

[5] Jiang, Z., Balu, A., Hegde, C., & Sarkar, S. (2017). Collaborative deep learning in fixed topology networks. Advances in Neural Information Processing Systems, 30, 1-11.

DOI: 10.54254/2755-2721/83/2024GLG0061

[6] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: A search space odyssey. IEEE transactions on neural networks and learning systems, 28(10), 2222-2232.
[7] Raihani, N. J., & Bshary, R. (2011). Resolving the iterated prisoner's dilemma: theory and reality. Journal of Evolutionary Biology, 24(8), 1628-1639.
[8] Greenwood, G. W., & Chopra, S. (2011). A numerical analysis of the evolutionary iterated snowdrift game. IEEE Congress of Evolutionary Computation, 2010-2016).
[9] Detailed Explanation and Practice of the NEAT (Neuro Evolution of Augmenting Topologies) Algorithm (Based on NEAT-Python) URL: https://blog.csdn.net/LOVEmy134611/article/details/115624709. Last Accessed: 2024/08/26
[10] Graves, A., Mohamed, A. R., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. IEEE international conference on acoustics, speech and signal processing, 6645-6649.