Teaching LLM to play Quixo

***S. M. Kunin-Bogoiavlenskij*[1], *A. S. Trushin*[2]**
kuninbogoiavlenskij@gmail.com; tarstars@yandex.ru
[1]Moscow Institute of Physics and Technology; [2]Yandex

This paper presents a study on teaching Large Language Models (LLMs) to play the board game Quixo. We explore various strategies and models to determine the most effective approach for training LLMs to excel in this strategic game. Our work includes implementing MinMax algorithm, leveraging MapReduce for efficient computation, and experimenting with different language models such as YandexGPT7B and torch.nn.Transformer. Additionally, we investigate the impact of multimodal input, including voice commands, and evaluate different notation systems for representing game moves. The findings aim to provide insights into the capabilities of LLMs in strategic gameplay and the effectiveness of various training methods.

Keywords: *Large Language Models; Quixo; MinMax algorithm; MapReduce*

1 Introduction

The field of artificial intelligence has witnessed significant advancements in recent years, particularly in the development of Large Language Models (LLMs). These models have demonstrated remarkable capabilities in understanding and generating human language, making them potential candidates for complex tasks such as playing strategic board games. In this paper, we focus on teaching LLMs to play Quixo, a board game that requires strategic thinking and planning. We aim to understand how clever LLMs are in comprehending the spatial characteristics and strategy involved in the game.

Our study is structured around several key objectives. First, we aim to implement and evaluate the performance of MinMax algorithm, which was proposed in the recent article. To handle the computational complexity, we employ MapReduce to efficiently distribute and process the calculations. Second, we train LLM to play using the acquired full tree of win/loss game states. Then we compare the learning capabilities of different language models, including YandexGPT7B (with 7 billion parameters), torch.nn.Transformer, and convolutional neural networks (CNNs). This comparison will help us understand which models are best suited for learning the intricacies of Quixo.

Additionally, we explore the potential of multimodal input, specifically voice commands, to enhance the interaction and learning process of the LLMs. We also investigate various notation systems for representing game moves, including different visual and verbal methods, to determine the most effective approach for training and communication.

Through this study, we aim to contribute to the understanding of LLMs' capabilities in strategic gameplay and provide insights into effective training methods and input modalities.

References

[1] Satoshi Tanak, Francois Bonnet, Sebastien Tixeuil, Yasumasa Tamura. 2020. Quixo Is Solved. Available at: `https://arxiv.org/pdf/2007.15895`

[2] Anian Ruoss, Gregoire Deletang, Sourabh Medapati, Jordi Grau-Moya, Li Kevin Wenliang, Elliot Catt, John Reid and Tim Genewein. 2024. Gradmaster-Level Chess Without Search. Available at: `https://arxiv.org/pdf/2402.04494v1`

[3] Hongyi Guo, Zhihan Liu, Yufeng Zhang, Zhaoran Wang. 2024. Can Large Language Models Play Games? A Case Study of A Self-Play Approach. Available at: `https://arxiv.org/pdf/2402.04494v1`

[4] Lukas Berglund, Asa Cooper Stickland, Meg Tong, Max Kaufmann, Mikita Balesni, Tomasz Korbak, Owain Evans 2023. The Reverasal Curse: LLMs trained on "A is B" fail to learn "B is A". Available at: https://arxiv.org/pdf/2206.10498

[5] Karthik Valmeekam, Matthew Marquez, Alberto Olmo, Sarath Sreedharan, Subbarao Kambhampati. 2023. PlanBench: An Extensible Benchmark for Evaluating Large Language Models on Planning and Reasoning about Change. Available at: https://arxiv.org/pdf/2206.10498