

# **Teaching LLM to play Quixo**

**Sergey Kunin-Bogoiavlenskij**

Advisor: **Arseniy Trushin**

Phystech School of Applied Mathematics and Computer Science

Moscow Institute of Physics and Technology, Dolgoprudny

May 2025

A thesis submitted in partial fulfillment of the requirements of the degree of

Bachelor of Science

©Sergey Kunin-Bogoiavlenskij, 2025

# Abstract

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 YandexGPT 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.

*Key Words: Large Language Models; Quixo; MinMax; MapReduce.*

# Table of Contents

Abstract . . . . .	i
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Objectives . . . . .	1
1.3 Quixo Rules . . . . .	2
1.3.1 Game Overview . . . . .	2
1.3.2 Gameplay . . . . .	2
1.3.3 Game End . . . . .	3

# Chapter 1

## Introduction

### 1.1 Motivation

Recent advancements in Large Language Models (LLMs) have opened exciting possibilities for applying AI to complex tasks, including strategic game playing. This research investigates the application of LLMs to Quixo, a challenging board game requiring spatial reasoning and strategic planning. While LLMs have demonstrated proficiency in various language-based tasks, their ability to master games with a strong spatial component remains an open question. This study aims to explore the capacity of LLMs to learn and excel at Quixo, providing insights into their strategic learning capabilities and the effectiveness of different training methodologies.

### 1.2 Objectives

Our study is structured around several key objectives. First, we aim to implement and evaluate the performance of the MinMax algorithm, which was proposed in the recent article [4]. To handle the computational complexity, we employ MapReduce to efficiently distribute and process the calculations. Second, we train an LLM to play using the acquired full tree of win/loss game states. Then we compare the learning capabilities of different language models, including YandexGPT,

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.

## 1.3 Quixo Rules

This research focuses on the two-player version of Quixo.

### 1.3.1 Game Overview

Quixo is played on a 5x5 grid of cubes. Initially, all 25 cubes have blank top faces. Players choose to play either crosses (X) or circles (O). The objective is to create a horizontal, vertical, or diagonal line of five cubes bearing the player's symbol.

### 1.3.2 Gameplay

Players take turns choosing and moving a cube according to the following rules:

1. **Cube Selection:** A player selects a blank cube or a cube with their symbol from the periphery of the board. In the first turn, only blank cubes are available. A player cannot select a cube with their opponent's symbol.
2. **Symbol Change:** The selected cube is replaced with a cube bearing the player's symbol on its top face.

3. **Cube Placement:** The player pushes the row or column of cubes from which the selected cube was taken, inserting the new cube at the opposite end. A player cannot place the cube back into the position from which it was just taken.

### **1.3.3 Game End**

The game ends when a player creates and announces a line of five of their symbols. A player who creates a line of five of their opponent's symbols loses.

# Bibliography

- [1] Lukas Berglund, Asa Cooper Stickland, Meg Tong, Max Kaufmann, Mikita Balesni, Tomasz Korbak, and Owain Evans. The reversal curse: LLMs trained on “a is b” fail to learn “b is a”, 2023. Available at: <https://arxiv.org/pdf/2206.10498>.
- [2] Hongyi Guo, Zhihan Liu, Yufeng Zhang, and Zhaoran Wang. Can large language models play games? a case study of a self-play approach, 2024. Available at: <https://arxiv.org/pdf/2402.04494v1>.
- [3] Anian Ruoss, Gregoire Deletang, Sourabh Medapati, Jordi Grau-Moya, Li Kevin Wenliang, Elliot Catt, John Reid, and Tim Genewein. Gradmaster-level chess without search, 2024. Available at: <https://arxiv.org/pdf/2402.04494v1>.
- [4] Satoshi Tanaka, Francois Bonnet, Sebastien Tixeuil, and Yasumasa Tamura. Quixote is solved, 2020. Available at: <https://arxiv.org/pdf/2007.15895>.
- [5] Karthik Valmeekam, Matthew Marquez, Alberto Olmo, Sarath Sreedharan, and Subbarao Kambhampati. PlanBench: An extensible benchmark for evaluating large language models on planning and reasoning about change, 2023. Available at: <https://arxiv.org/pdf/2206.10498>.