Applications of Large Language Models in Special Sectors

Supervisor: Prof.CHEN Kani

Students: HAN, Liuruo

#### Abstract

This project focuses on developing a specialized AI model using multi-agents structure, to enhance mathematical logic and reasoning capabilities of basic large language models. Through the integration of specialized agents and multi-agent systems, the AI assistant, can provide accurate solutions, detailed explanations, and deeper insights into mathematical concepts for relatively simple math problems.

#### Introduction

The ability to comprehend and solve mathematical problems is a fundamental skill for students across various educational levels. While existing large-scale language models have demonstrated exceptional language understanding capabilities, their proficiency in mathematical and logical reasoning remains limited. This project aims to address this gap by developing a powerful AI model specialized in mathematical logic and reasoning to serve as an AI assistant. The objective is to generate detailed articles that solve mathematical problems, aiding students in better understanding mathematical concepts and facilitating problem-solving.

In recent years, significant advancements have been made in the field of artificial intelligence, particularly with large-scale language models like ChatGPT. These models have exhibited impressive language generation capabilities and have been widely employed in various applications, including natural language understanding, conversation generation, and text completion tasks. However, their effectiveness in mathematical problem solving, which requires rigorous logic and reasoning, has been relatively limited.

To overcome this limitation, it is essential to explore and develop AI models that excel in mathematical logic and reasoning. By augmenting the capabilities of ChatGPT with specialized agents, we aim to enhance its ability to solve mathematical problems and assist students in their learning process.

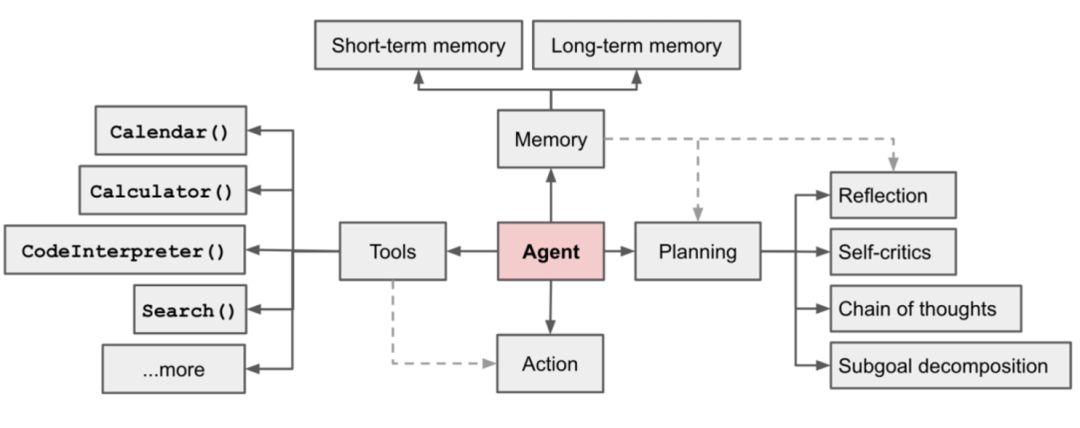


Figure1. Working principle of Agents.

To guide our project, we conducted a comprehensive competitive analysis to gain insights into existing products with similar functionalities. Through this analysis, we studied the features and evaluated the effectiveness of these products in the context of mathematical problem solving. The findings from this study served as a foundation for determining the desired functionalities for our AI assistant.

Furthermore, we delved into relevant research papers to expand our knowledge and gain insights into state-of-the-art techniques in AI-assisted mathematical problem solving. One notable paper, "Solving Olympiad Geometry without Human Demonstrations" provided valuable insights into leveraging AI to solve mathematical problems. The concepts from this paper were shared and discussed with our research team, aiming to inspire new ideas and adopt effective strategies for our project.

With the primary objective of developing an AI assistant that can generate comprehensive articles to solve mathematical problems, we aim to incorporate advanced mathematical logic and reasoning capabilities. This AI assistant will provide detailed step-by-step solutions and explanations, leveraging the power of ChatGPT and specialized agents. Our goal is to create an AI assistant that not only generates accurate solutions but also enhances students' understanding of the underlying mathematical concepts.

In summary, this research report introduces a project focused on building an AI assistant with strong mathematical logic and reasoning abilities. By leveraging the existing language understanding capabilities of ChatGPT and incorporating specialized agents, our aim is to facilitate mathematical problem solving for students. The competitive analysis, along with the insights gained from relevant research papers, has guided our project's direction and inspired innovative approaches. Through this project, we hope to contribute to the advancement of AI in the field of mathematical education, ultimately helping students better comprehend and solve mathematical problems.

#### Methodology

In the context of our project, we have developed a multi-agent system consisting of various agents that collectively work towards achieving our objectives--solve math problems. These agents play distinct roles and collaborate to enhance the capabilities of ChatGPT, enabling it to fulfill specific tasks with greater effectiveness and relevance.

The following table outlines the different roles and corresponding jobs of the agents:

|  |  |  |
| --- | --- | --- |
| Agents | Role | Job |
| analyzer | Math expert | Analyze the problem step by step |
| computation | Python expert | Use to execute python code to solve math problems |
| derivation | Math expert | Do mathematical derivations based on the given thought using rigor mathematical expressions |
| designer | N/A | Get the format of the article |
| editor | N/A | Get the html layout of the article |
| explainer | Math teacher expert in explaining in simple yet accurate terms | Explain how to solve a math problem |
| input\_checker | System/user | Check whether the input is a valid math problem |
| judger | System/user | Judge whether a problem needs rigor mathematical transformation |
| pulisher | N/A | Save newspaper html |
| questioner | Math expert | Question the thoughts and analysis process on a certain problem |
| reflection | Math expert | Criticize the usefulness of thought |
| thinker | Math expert | Think about what to do next for solving the question you are given based on current progress and feedback. |
| writer | Excellent explanatory blog writer | Write a well-written and engaging article about how to solve a math problem. |

Through the coordinated efforts of these agents, ChatGPT can assume different roles and adapt its responses accordingly. This multi-agent approach enhances the AI assistant's ability to solve mathematical problems comprehensively and offer valuable insights to students. By leveraging the unique capabilities of each agent, we aim to provide students with accurate solutions, detailed explanations, and a deeper understanding of mathematical concepts.

#### Implementation

In this section, we will briefly showcase our work so far, as well as the difficulties and challenges we have encountered. Additional sample programs are demonstrated in the Appendix section.

1. ***Competitor analysis***

During my competitive research, I explored a variety of math learning products including Khan Academy, Khanmigo, Eduade.AI, Synthesis Tutor, MathSolver, Brainiac Buddy, and mathGPT. I analyzed their features, user interface, target audience, and limitations to gain a better understanding of the current market and identify areas where my product can stand out.

To begin my research, I first identified the key features that are essential for a successful math learning product. I then explored each of the products in detail, noting their strengths and weaknesses. I paid particular attention to their user interface, as I believe that a well-designed interface is crucial to engaging users and promoting learning.

Through my research, I identified a need for a math learning product with high accuracy and a focus on interactive games and emotional motivation. I realized that many of the existing products on the market were lacking in these areas and that there was an opportunity for my product to stand out by addressing these needs.

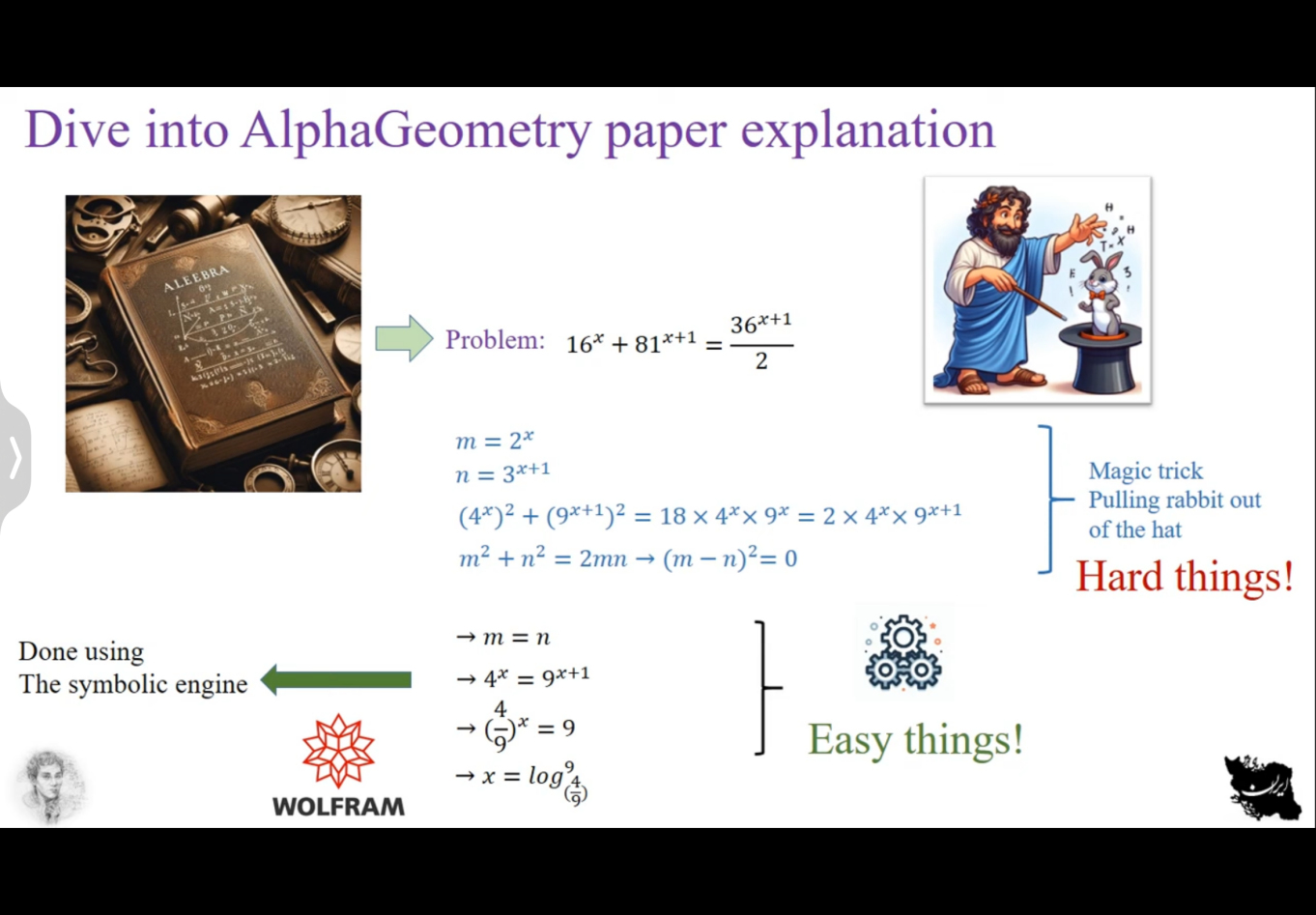
Overall, my competitive research has helped me establish a clear direction for the development of my product and increased my confidence in its potential success. I now have a better understanding of the market and the needs of my target audience, which will allow me to create a product that truly meets their needs and exceeds their expectations.

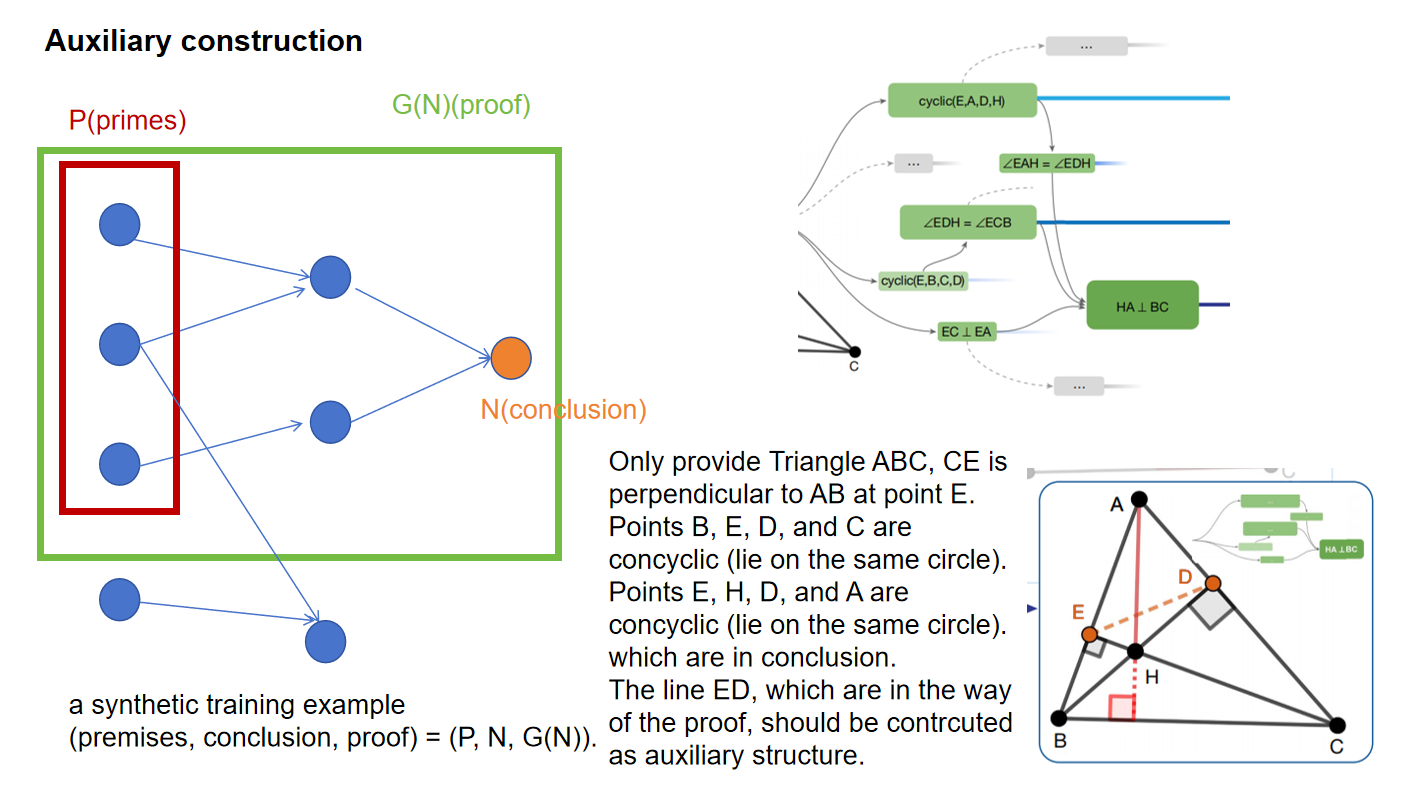
1. ***Establishment of agents***

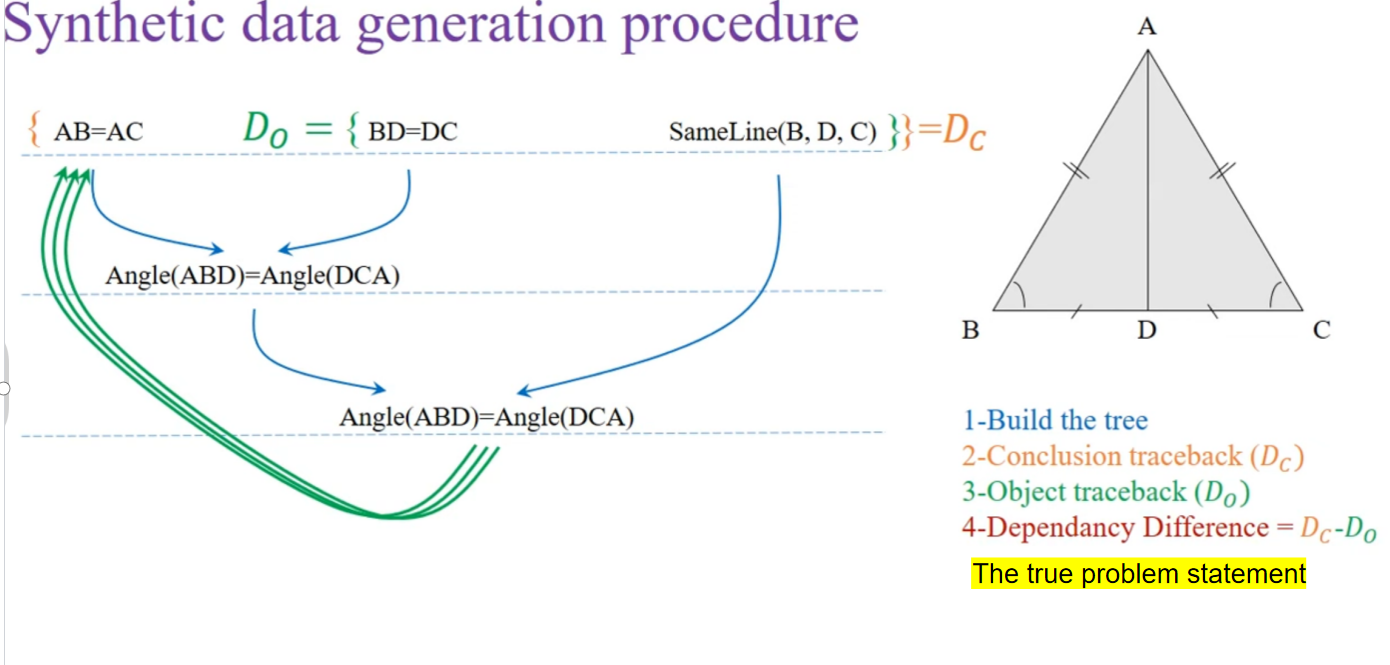
******

1. ***Further improvement by intensive reading of the paper named “Solving Olympiad Geometry without Human Demonstrations”***

The research paper introduces AlphaGeometry, a neuro-symbolic system designed to solve olympiad-level geometry problems without the need for human demonstrations. This innovative approach combines neural language modeling with symbolic deduction engines to navigate complex geometry theorems and proofs. By synthesizing millions of theorems and proofs across varying levels of complexity, AlphaGeometry achieves remarkable success in solving challenging geometry problems.







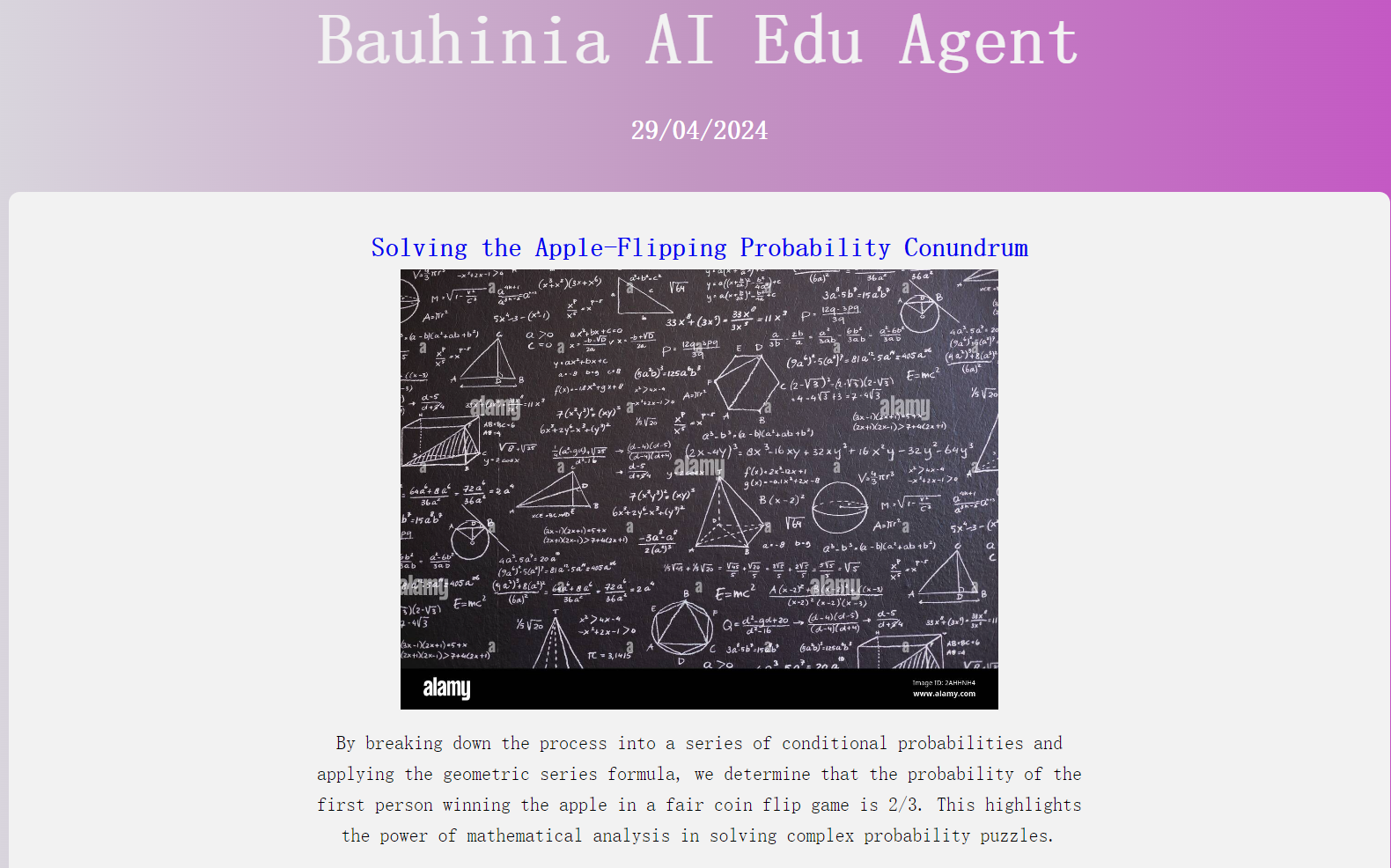
We try to use this logic and are still finding a way out.

#### Results

I ask one math question: Two people are eating apples, and they decide by flipping a coin. The person who gets heads first gets to eat. What is the probability that the person who flips first will eat the apple?



After about one minutes, the answer comes out as an article. The answer is correct.



#### Conclusion and future development

#### The utilization of a multi-agent framework has proven to be valuable in harnessing the capabilities of large-scale models through the formulation and refinement of prompts. Moving forward, there are plans to further explore the applicability of the "Solving Olympiad Geometry without Human Demonstrations" paper and how its principles can be incorporated into our own work. This includes investigating methods to generate mathematical problem data and corresponding answers autonomously, as well as developing techniques to efficiently convert human language descriptions of problems into symbolic expressions that can be understood by the symbolic engine, thereby leveraging the powerful auxiliary line generation capabilities and proof abilities of AlphaGeometry for plane geometry.

#### By delving deeper into these areas, we aim to enhance the functionality and performance of our system, ultimately providing advanced problem-solving assistance and a broader range of mathematical insights to users.

#### Reference

[1] Trinh, T.H., Wu, Y., Le, Q.V. et al. Solving olympiad geometry without human demonstrations. Nature 625, 476–482 (2024). https://doi.org/10.1038/s41586-023-06747-5