AlphaGeometry: Solving olympiad geometry without human demonstrations

Amit Kumar

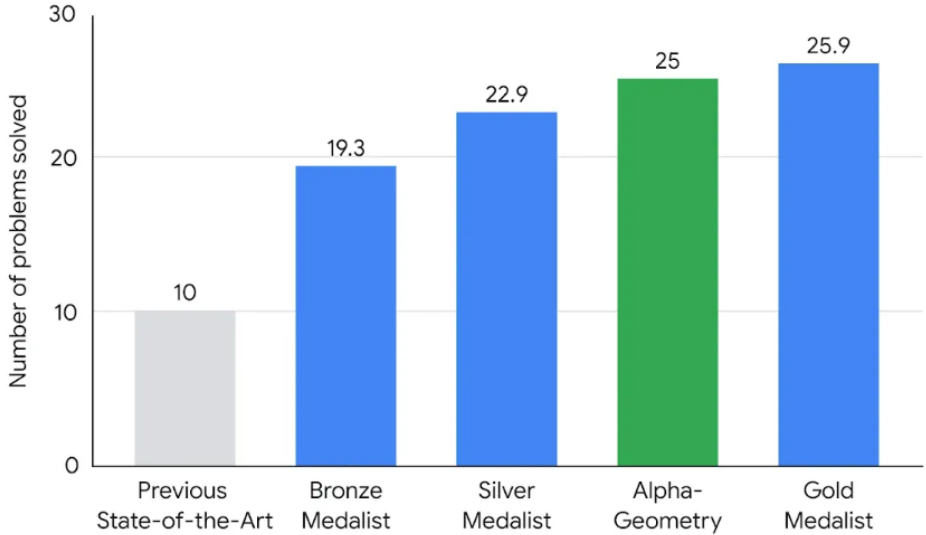
AI/NLP Engineer at E42.ai

Table of contents

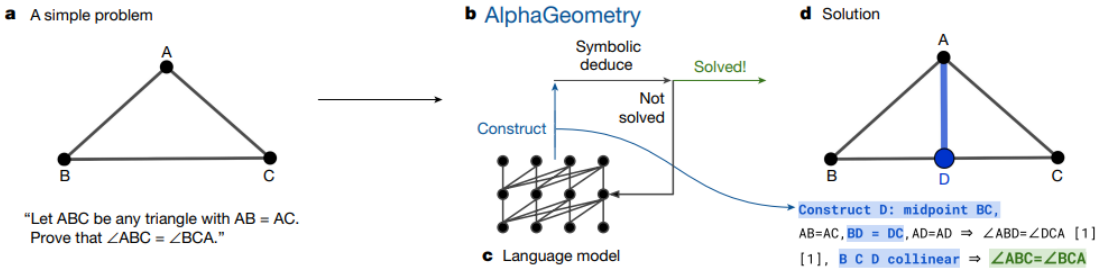
Introduction

- Proving olympiad-level math theorems is hard, even for top young mathematicians. Most AI can't do it because translating human proofs to machine format is costly and especially difficult in geometry.
- AlphaGeometry is a new AI system that doesn't need human-written proofs. It creates its own huge dataset of geometry theorems and proofs, then uses a special language model to guide a logic engine through solving new problems.
- AlphaGeometry is good! It solves 25 out of 30 recent olympiad problems, better than any other AI and close to an average IMO gold medalist.
- AlphaGeometry can explain its proofs in a way humans can understand and even discovered a new theorem!

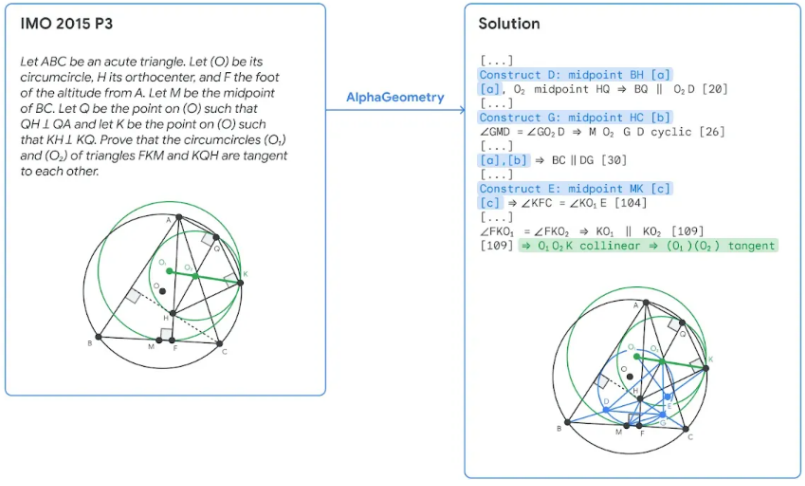
Approaching the Olympiad gold-medalist standard



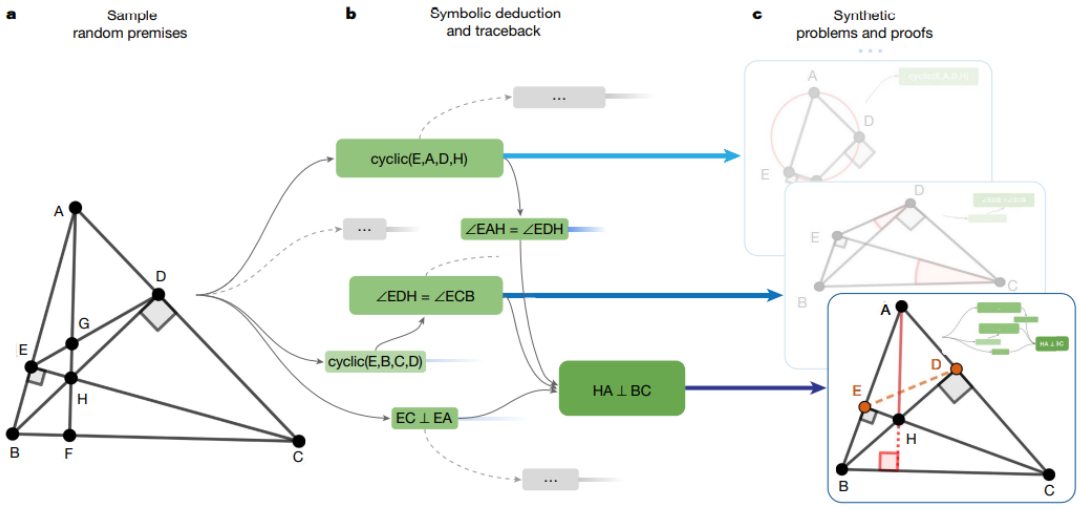
AlphaGeometry



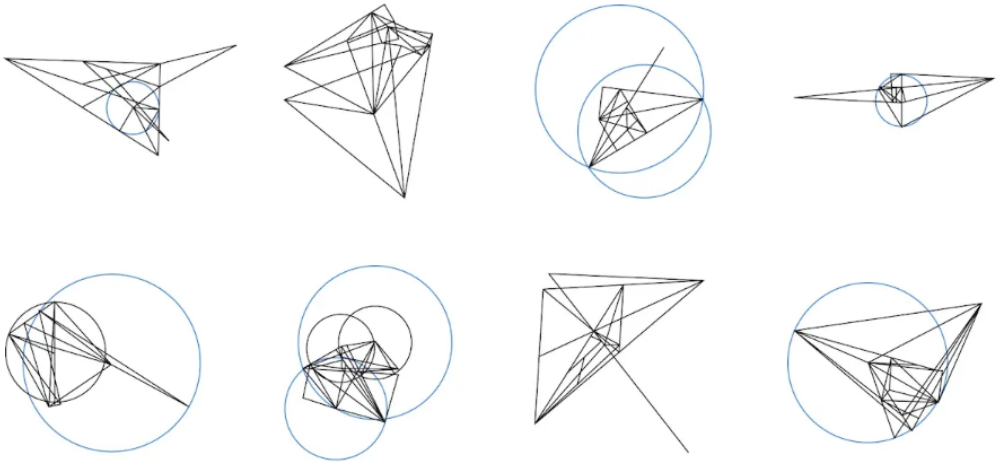
- AlphaGeometry solving an Olympiad problem: Problem 3 of the 2015 International Mathematics Olympiad (left) and a condensed version of AlphaGeometry's solution (right). The blue elements are added constructs. AlphaGeometry's solution has 109 logical steps.



Synthetic data generation process

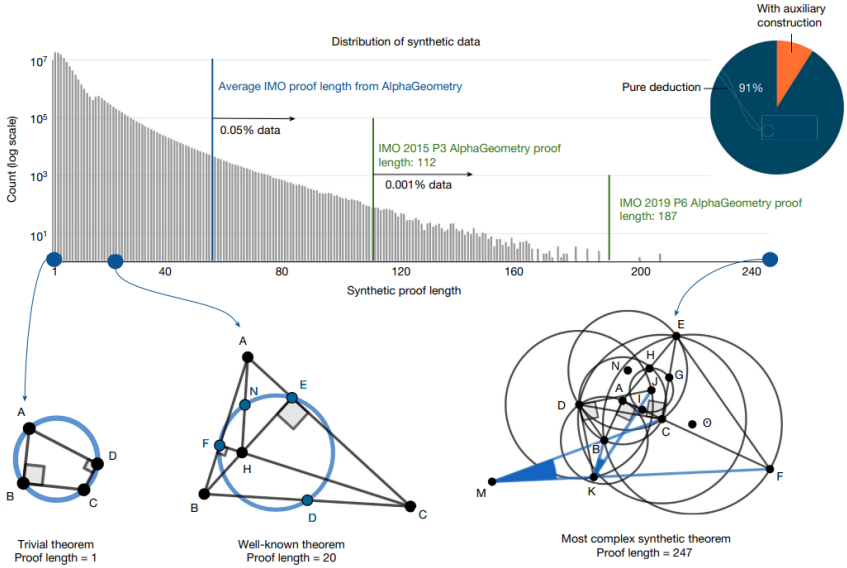


Synthetic Data



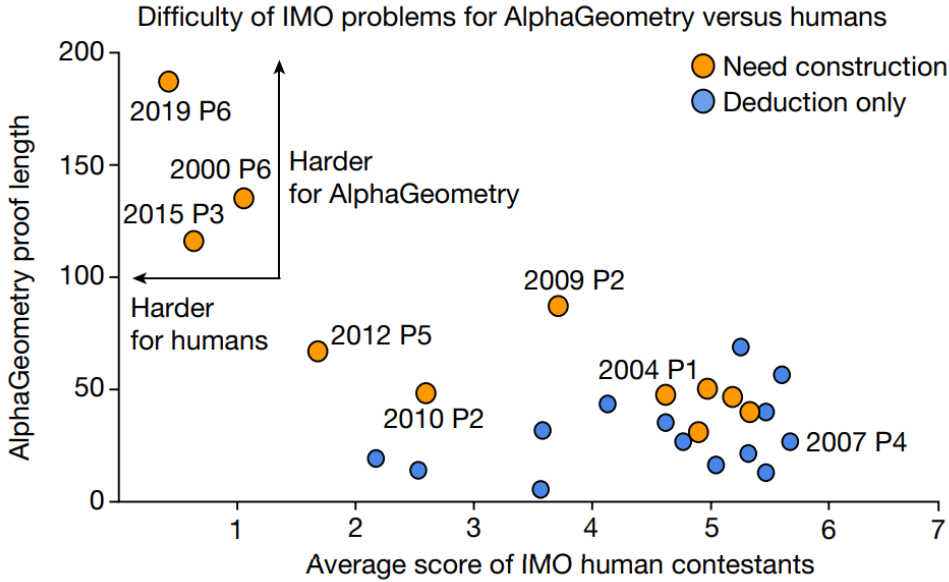
Visual representations of the synthetic data generated by AlphaGeometry

Experimental Results



Method		Problems solved (out of 30)
Computer algebra	Wu's method[21] (previous state of the art)	10
	Gröbner basis[20]	4
Search (human-like)	GPT-4 (ref. 25)	0
	Full-angle method[30]	2
	Deductive database (DD)[10]	7
	DD + human-designed heuristics[17]	9
	DD + AR (ours)	14
	DD + AR + GPT-4 auxiliary constructions	15
	DD + AR + human-designed heuristics	18
	AlphaGeometry	25

Main results on IMO-AG-30 test benchmark


Difficulty of IMO problems for AlphaGeometry versus humans
Need construction
Deduction only
2019 P6
2000 P6
2015 P3
Harder
for AlphaGeometry
Harder
for humans
2009 P2
2012 P5
2004 P1
2010 P2
2007 P4
AlphaGeometry proof length
Average score of IMO human contestants
200
150
100
50
0
1
2
3
4
5
6
7

Conclusion

1. **Modular Deductive Reasoning**: AlphaGeometry demonstrates the effectiveness of decomposing complex mathematical challenges into manageable subproblems. It then utilizes a combination of deduction and construction techniques to achieve solutions, highlighting the importance of structured, step-by-step reasoning in tackling intricate tasks.
2. **Strategic Synthetic Data Generation**: The success of AlphaGeometry underscores the potential of synthetic data generation, particularly when designed with worst-case scenarios in mind. By pre-training on a diverse set of synthetically generated theorems and proofs, the system equips itself to handle a wider range of real-world problems not yet encountered. (ICR)
3. **LLM Guidance**: AlphaGeometry showcases the value of employing large language models (LLMs) not simply as answer generators, but as strategic guides for symbolic deduction engines. By leveraging the LLM's knowledge of geometric relationships and problem-solving patterns, the system navigates the vast solution space more efficiently and effectively. (Agents)

References

- Solving olympiad geometry without human demonstrations.
- AlphaGeometry: An Olympiad-level AI system for geometry
- AlphaGeometry - Github

Thank You