

Title: Gold-medalist Performance in Solving Olympiad Geometry with AlphaGeometry2

Speakers: Yuri Chervonyi

Collection/Series: Theory + AI Symposium

Date: April 08, 2025 - 4:00 PM

URL: <https://pirsa.org/25040075>

Abstract:

We present AlphaGeometry2, a significantly improved version of AlphaGeometry introduced in Trinh et al. (2024), which has now surpassed an average gold medalist in solving Olympiad geometry problems. To achieve this, we first extend the original AlphaGeometry language to tackle harder problems involving movements of objects, and problems containing linear equations of angles, ratios, and distances. This, together with support for non-constructive problems, has markedly improved the coverage rate of the AlphaGeometry language on International Math Olympiads (IMO) 2000-2024 geometry problems from 66% to 88%. The search process of AlphaGeometry2 has also been greatly improved through the use of Gemini architecture for better language modeling, and a novel knowledge-sharing mechanism that enables effective communication between search trees. Together with further enhancements to the symbolic engine and synthetic data generation, we have significantly boosted the overall solving rate of AlphaGeometry2 to 84% for all geometry problems over the last 25 years, compared to 54% previously. AlphaGeometry2 was also part of the system that achieved silver-medal standard at IMO 2024 this [https URL](https://pirsa.org/25040075). Last but not least, we report progress towards using AlphaGeometry2 as a part of a fully automated system that reliably solves geometry problems directly from natural language input.

Esc



AI achieves silver-medal standard solving International Mathematical Olympiad problems

25 JULY 2024

AlphaProof and AlphaGeometry teams

[Share](#)

Google Deepmind at IMO 2024

- International Mathematical Olympiad (IMO) - math competition for high school students
 - 6 problems from 4 categories: algebra, geometry, number theory, combinatorics
- AlphaProof:
 - Reinforcement-learning based system for formal math reasoning (Lean)
 - Couples a language model with the AlphaZero RL algorithm (superhuman performance in Go)
 - Formal - no hallucinations
 - Solved 3 (algebra, number theory)
- AlphaGeometry2:
 - Neuro-symbolic system for geometry (language model + symbolic engine)
 - Solved geometry problem in 30 sec
- Total solved 4 out of 6 - 28 points out of 42 (1 point away from gold)
- Graded by:
 - Prof Sir Timothy Gowers, an IMO gold medalist and Fields Medal winner
 - Dr Joseph Myers, a two-time IMO gold medalist and Chair of the IMO 2024 Problem Selection Committee

Score on IMO 2024 problems



Graph showing performance of our AI system relative to human competitors at IMO 2024. We earned 28 out of 42 total points, achieving the same level as a silver medalist in the competition.

AlphaGeometry2 [arXiv:2502.03544]

Google DeepMind

2025-3-4

Gold-medalist Performance in Solving Olympiad Geometry with AlphaGeometry2

Yuri Chervonyi^{*,1,◊}, Trieu H. Trinh^{*,1,◊}, Miroslav Olšák^{†,1,2}, Xiaomeng Yang^{‡,1}, Hoang Nguyen^{1,3}, Marcelo Menegali¹, Junehyuk Jung^{1,4}, Vikas Verma¹, Quoc V. Le¹ and Thang Luong^{1,◊}

¹Google DeepMind, ²University of Cambridge, ³Georgia Institute of Technology, ⁴Brown University
This work was conducted entirely at Google DeepMind by all authors.

We present AlphaGeometry2, a significantly improved version of AlphaGeometry introduced in (Trinh et al., 2024), which has now surpassed an average gold medalist in solving Olympiad geometry problems. To achieve this, we first extend the original AlphaGeometry language to tackle harder problems involving movements of objects, and problems containing linear equations of angles, ratios, and distances. This, together with support for non-constructive problems, has markedly improved the coverage rate of the AlphaGeometry language on International Math Olympiads (IMO) 2000-2024 geometry problems from 66% to 88%. The search process of AlphaGeometry2 has also been greatly improved through the use of Gemini architecture for better language modeling, and a novel knowledge-sharing mechanism that enables effective communication between search trees. Together with further enhancements to the symbolic engine and synthetic data generation, we have significantly boosted the overall solving rate of AlphaGeometry2 to 84% for *all* geometry problems over the last 25 years, compared to 54% previously. AlphaGeometry2 was also part of the system that achieved silver-medal standard at IMO 2024 <https://dpm.d.ai/imo-silver>. Last but not least, we report progress towards using AlphaGeometry2 as a part of a fully automated system that reliably solves geometry problems directly from natural language input.

AlphaGeometry2: Results

System description	IMO-AG-50 solved	IMO-AG-30 solved
OpenAI o1	0	0
Gemini thinking	0	0
AG1 DDAR (Trinh et al., 2024)	14	14
AG2 DDAR	16	15
TongGeometry DD (Zhang et al., 2024)	-	18
Average bronze medalist	27.1	19.3
Wu with AG1 DDAR (Sinha et al., 2024)	-	21
Average silver medalist	33.9	22.9
AG1 (Trinh et al., 2024)	27	25
Average gold medalist	40.9	25.9
Wu + AG1 (Sinha et al., 2024)	-	27
TongGeometry w/o value (Zhang et al., 2024)	-	28
AG2 with AG1 setup	38	28
TongGeometry full setting (Zhang et al., 2024)	-	30
AG2 full setting	42	30

Table 4 | Evaluation on IMO-AG-50 benchmark. IMO-AG-50 contains *all* IMO 2000-2024 geometry problems, while IMO-AG-30 introduced in ([Trinh et al., 2024](#)) contains only a subset formalizable in terms of the AG1 language.

MathArena: Evaluating LLMs on Uncontaminated Math Competitions



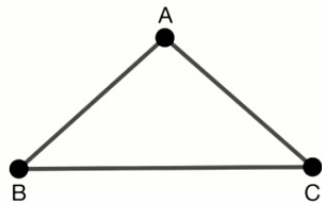
[Click on a cell to see the raw model output.](#)

Overall AIME 2025 I AIME 2025 II HMMT February 2025 **USAMO 2025**

Model	Acc	Cost	1	2	3	4	5	6
gemini-2.5-pro	24.40%	N/A	93%	0%	4%	50%	0%	0%
DeepSeek-R1	4.76%	\$2.03	7%	0%	0%	21%	0%	0%
gemini-2.0-flash-thinking	4.17%	N/A	21%	0%	0%	0%	4%	0%
Claude-3.7-Sonnet (Think)	3.65%	\$9.03	7%	7%	0%	0%	0%	8%
QwQ-32B	2.98%	\$0.42	18%	0%	0%	0%	0%	0%
o1-pro (high)	2.83%	\$203.44	7%	0%	0%	0%	4%	6%
o3-mini (high)	2.08%	\$1.11	7%	2%	0%	0%	0%	4%

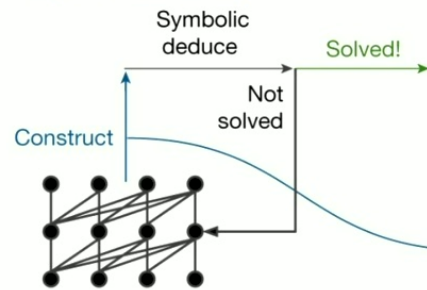
AlphaGeometry: Overview

a A simple problem



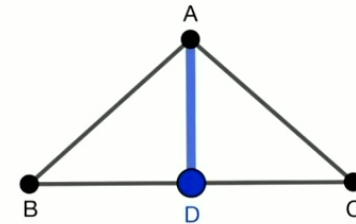
"Let ABC be any triangle with $AB = AC$.
Prove that $\angle ABC = \angle BCA$."

b AlphaGeometry



c Language model

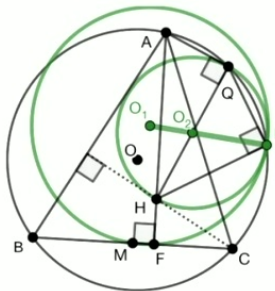
d Solution



Construct D: midpoint BC,
 $AB=AC, BD = DC, AD=AD \Rightarrow \angle ABD=\angle DCA$ [1]
 [1], **B C D collinear** $\Rightarrow \angle ABC=\angle BCA$

e IMO 2015 P3

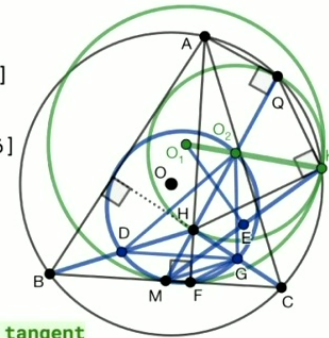
"Let ABC be an acute triangle. Let (O) be its circumcircle, H its orthocenter, and F the foot of the altitude from A. Let M be the midpoint of BC. Let Q be the point on (O) such that $QH \perp QA$ and let K be the point on (O) such that $KH \perp KQ$. Prove that the circumcircles (O_1) and (O_2) of triangles FKM and KQH are tangent to each other."



Alpha-
Geometry

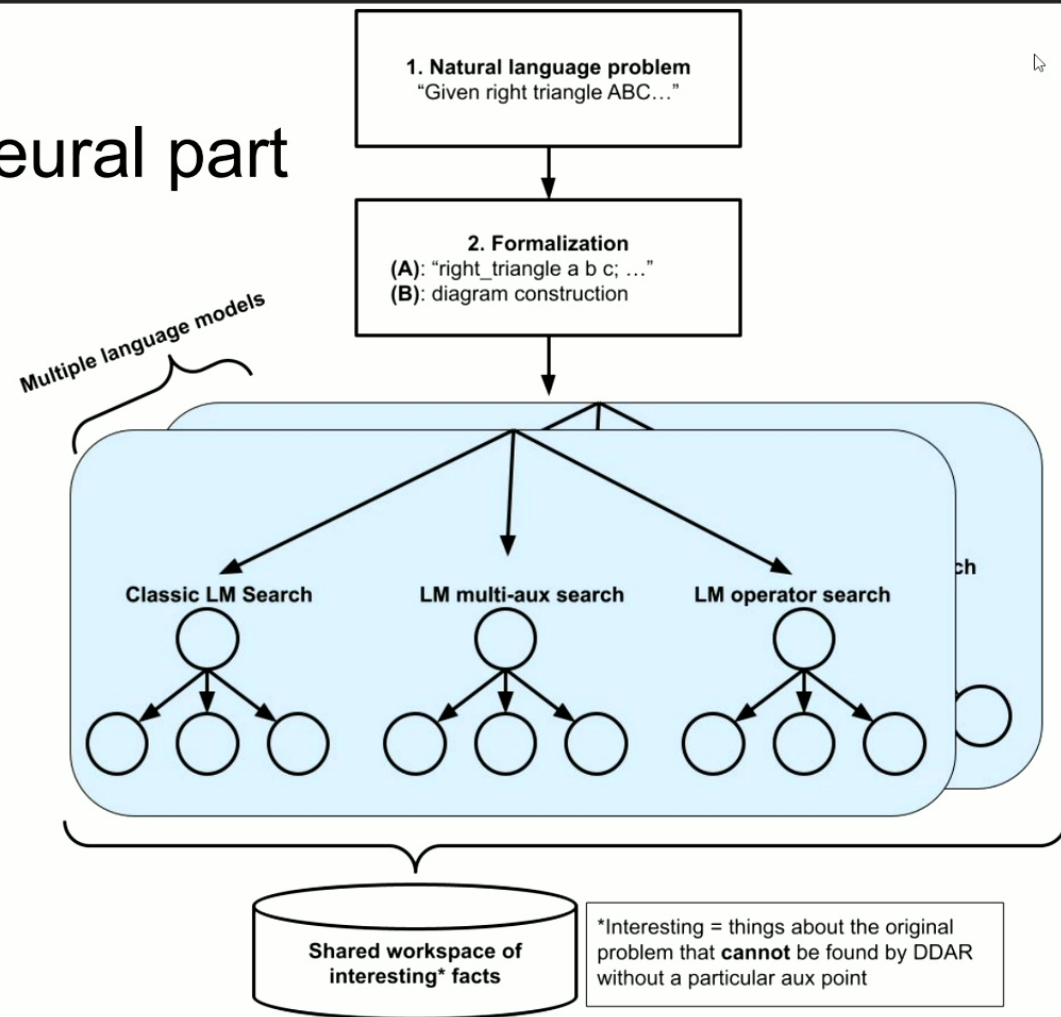
f Solution

...
Construct D: midpoint BH [a]
 [a], O_2 midpoint HQ $\Rightarrow BQ \parallel O_2D$ [20]
 ...
Construct G: midpoint HC [b] ...
 $\angle GMD = \angle GO_2D \Rightarrow M O_2 G D$ cyclic [26]
 ...
 [a], [b] $\Rightarrow BC \parallel DG$ [30]
 ...
Construct E: midpoint MK [c]
 ..., [c] $\Rightarrow \angle KFC = \angle KO_1E$ [104]
 ...
 $\angle FKO_1 = \angle FKO_2 \Rightarrow KO_1 \parallel KO_2$ [109]
 [109] $\Rightarrow O_1 O_2 K$ collinear $\Rightarrow (O_1)(O_2)$ tangent



AlphaGeometry2: Neural part

- Language/foundation models are trained on synthetic data - text (+ image)
- Data is generated by the symbolic engine



AlphaGeometry2: Symbolic part

- Algorithm:
 - Set up all constraints from the problem statement `ddar.force_pred(pred)`, for all `pred` in predicates
 - Compute all facts from the problem statement obtained at step 1) `ddar.deduction_closure()`
 - Search for the problem goal among all the facts obtained at step 2) `ddar.check_pred(goal)`
- Used to:
 - Complete the proof given LM output
 - Generate synthetic data to train LM
- Implementation:
 - C++
 - Very fast ~1 sec per problem

AlphaGeometry2: Example

