

# Proposing and solving olympiad geometry with guided tree search

Chi Zhang<sup>1\*</sup>    Jiajun Song<sup>1</sup>    Siyu Li<sup>1,2</sup>    Yitao Liang<sup>1,2</sup>    Yuxi Ma<sup>1,2</sup>  
Wei Wang<sup>1</sup>    Yixin Zhu<sup>2\*</sup>    Song-Chun Zhu<sup>1,2\*</sup>

<sup>1</sup>Beijing Institute for General Artificial Intelligence (BIGAI), Beijing, 100080, China.

<sup>2</sup>Institute for Artificial Intelligence, Peking University, Beijing, 100080, China.

\*Corresponding author. Email: zhangchi@bigai.ai, yixin.zhu@pku.edu.cn, sczhu@bigai.ai

Mathematics olympiads are prestigious competitions, with problem proposing and solving highly honored. Building artificial intelligence that proposes and solves olympiads presents an unresolved challenge in automated theorem discovery and proving (1), especially in geometry for its combination of numerical and spatial elements. We introduce TongGeometry, a Euclidean geometry system supporting tree-search-based guided problem proposing and solving. The efficient geometry system establishes **the most extensive repository of geometry theorems** to date: within the same computational budget as the existing state-of-the-art (2), TongGeometry discovers **6.7 billion geometry theorems** requiring auxiliary constructions, including **4.1 billion exhibiting geometric symmetry**. Among them, 10 theorems were proposed to regional mathematical olympiads with 3 of TongGeometry’s proposals selected in real competitions, earning spots in a national team qualifying exam or a top civil olympiad in China and the US. Guided by fine-tuned large language models, TongGeometry solved all International Mathematical Olympiad geometry in IMO-AG-30, outperforming gold medalists for the first time. It also surpasses the existing state-of-the-art across a broader spectrum of olympiad-level

**problems. The full capabilities of the system can be utilized on a consumer-grade machine, making the model more accessible and fostering widespread democratization of its use. By analogy, unlike existing systems that merely solve problems like students, TongGeometry acts like a geometry coach, discovering, presenting, and proving theorems.**

**Significance** *This research presents TongGeometry, a system for Euclidean geometry problem proposing and solving, bridging numerical and spatial reasoning. TongGeometry surpasses state-of-the-art methods by autonomously discovering 6.7 billion theorems requiring auxiliary constructions, including 4.1 billion exhibiting geometric symmetry. Its real-world impact is demonstrated by its proposed problems being featured in prestigious math olympiads and outperforming top human competitors in solving International Mathematical Olympiad geometry problems. This achievement represents a step towards both automated theorem discovery and problem-solving, akin to transitioning from a “student” system to a “coach” capable of both generating and solving advanced problems. Importantly, its efficiency and accessibility on consumer-grade machines democratize high-level geometric problem-solving, enabling broader educational and research applications.*

Theorem proving, the journey to navigate from the initial state to the goal by connecting potentially missing intermediate conditions, demonstrates exceptional problem-solving skills. For any individual, winning a prize in the International Mathematical Olympiad (IMO) is a distinct honor and a testament to his / her problem-solving abilities. Since the inception, IMO has become the premier platform for identifying the world’s best mathematical talents. It has also, inadvertently, become the target for Artificial Intelligence (AI) researchers striving to develop theorem-proving models that can match or surpass human competitors in the most prestigious mathematical competition in the world (1, 3–6).

**Proposing olympiad problems** is held in high regard within the mathematics community as well, yet few computational methods have been able to accomplish autonomous problem proposing and rigorous proposal verification. Olympiad problems are often founded on years of intermediate theorems discovered and come with stringent criteria. The most exemplary problems appear deceptively simple and are comprehensible with a basic level of mathematical knowledge but re-

quire exceptional ingenuity to solve completely. Additionally, elegance, such as various forms of symmetry, is typically essential.

Proposing or solving olympiad problems poses considerable challenges for computational methods due to the vast number of branching points and the limited pool of past problems available for heuristics development. Geometry, in particular, stands out among olympiad topics because it involves both numerical and spatial reasoning, unlike other subjects that are text-based and can be more easily adapted with pretrained Large Language Models (LLMs) (7, 8). Besides, contentious axiomatization in Euclidean geometry invalidates either completeness (9) or human-readability of interactive proof assistants applicable in other domains (10, 11). Although specific languages tailored for geometry have been created (2, 12–14), their design limitations leave the space of olympiad problems inadequately explored.

In this work, we introduce TongGeometry<sup>1</sup>, a tree-based system for synthetic Euclidean planar geometry that facilitates human-readable problem proposing via backward tracing and theorem proving through forward chaining. Utilizing this scalable and efficient system and 196 olympiad problems as guiding statistics, we collected 4.1 billion symmetric geometry problems with proofs, among 6.7 billion that all necessitate auxiliary constructions for a complete proof. By analogy, compared to existing “student” systems that only solve problems (2, 13, 14), TongGeometry serves as a “coach” that characterizes the space of geometry using a finite tree and enables guided problem proposing and proving. We also established scoring rubrics for the difficulty and suitability of problems for olympiad competitions and selected a few as proposals. One was selected in 2024 National High School Mathematics League (Beijing), a qualifying competition for Chinese National Team<sup>2</sup>, and two made the shortlist of the 2024 US Ersatz Math Olympiad, a top civil mathematics olympiad in the US<sup>3</sup>.

The generated data contains abundant auxiliary constructions for solving geometry problems. Filling in these auxiliary constructions is crucial for successful geometry theorem proving; these exogenous objects enable a proving system to bridge the gap between the initial state and the goal. We therefore leveraged the data to guide TongGeometry tree search when it is presented a problem to solve. Specifically, we fine-tuned two LLMs (15, 16): one dedicated to suggesting possible search

---

<sup>1</sup>Tong stands for “general” in Chinese.

<sup>2</sup>Team China won the first place in IMO 2019-2023 and the second place in IMO 2024.

<sup>3</sup>The civil competition is organized by Team USA coaches. Team USA won the first place in IMO 2024.

directions and another for estimating the number of steps to go in each direction. When tested on the past 23 years of IMO geometry problems in the IMO-AG-30 benchmark, the **neuro-symbolic system** solves each of them within 38 minutes on a consumer-grade machine with 32 CPU cores and a single GPU card, becoming the first to surpass an average IMO gold medalist (2).

## Assessing TongGeometry

### Evaluating generated problems for olympiads

Using 196 existing olympiad problems as guiding statistics, we performed massive parallel problem search using 10,368 parallel CPU cores. In 30 days of search, TongGeometry traverses 143,379,886 unique paths (170,883,417 in total) in the defined space of geometry, inferring over 1,851,166,755 unique states. On each unique path, TongGeometry finds on average 0.7613 configuration requiring auxiliaries, resulting in 109,157,477 configurations (pairs of context and auxiliaries). Among them, 70,703,508 are unique. After filtering, we ended up with a dataset of 6,688,310,403 problems (triplets of context, goal and auxiliaries), of which 4,096,680,574 are symmetric. See Figure 1(B) for comparison of search efficiency with AlphaGeometry.

With billions of problems available, the challenge of selectively retrieving suitable ones for olympiad proposals remains unresolved due to the absence of an automatic assessment scheme. Inspired by **GeoGen’s methodology** (14), we developed a set of rubrics for problem selection.

Before the problem proposal deadlines for the 2024 National High School Mathematics League (Beijing) and the 2024 US Ersatz Math Olympiad, we enlisted the expertise of a 2023 IMO gold medalist and a student member of Chinese National Team to manually examine a batch of proposals during the initial phase of problem search. 4 proposals were submitted to the 2024 National High School Mathematics League (Beijing), with 1 selected as the only geometry problem in the final competition. Additionally, 6 proposals were submitted to the 2024 US Ersatz Math Olympiad, with 2 making it to the shortlist (Figure 1(E) and Supplementary Text). Of note, AlphaGeometry solved 3 of our 10 proposals.

## Search discovers known lemmas and base configurations

Among the proposed problems, we discover well-known lemmas. Figure 1(D) shows an rotationally symmetric lemma known in the community, where the three green lines are concurrent at the Nine-point center (17, 18). See Supplementary Text for explanation.

To facilitate problem exploration, we implemented a **replay buffer** that allowed **tree search to restart from cached promising intermediate states**. These states include foundational configurations well-known in the community, which serve as scaffolds for constructing more complex problem scenarios. Figure 1(F) illustrates a symmetric example featuring the Mixtilinear Incircle (19) and the Incenter/Excenter Lemma (20). This configuration has been acclaimed as the “richest configuration” by a coach of the USA team (19). See Supplementary Text for explanation.

## Performance on benchmarks

We performed quantitative analysis of TongGeometry on two benchmarks, the IMO-AG-30 dataset that was curated in AlphaGeometry and the newly curated dataset in the development of TongGeometry coined MO-TG-225. The IMO-AG-30 dataset comprises 30 problems, derived by translating 23 years of IMO problems into the domain-specific language of AlphaGeometry. In contrast, the **MO-TG-225 dataset includes 225 mathematical olympiad problems selected from our pool of 196 examples used to calculate search statistics**. Problems in MO-TG-225 have been translated into the domain-specific language of TongGeometry, and none of these problems appear in TongGeometry’s training dataset.

In experiments, we compared TongGeometry with AlphaGeometry, GPT-4 (7) and its o1 variant. Given that TongGeometry and AlphaGeometry employed distinct domain-specific languages, we translated each problem’s original representation into the respective language required by each model. For evaluating GPT-4 and o1, we used the natural language format in which each problem was presented in the competition. All models were assessed with a **standardized 90-minute time limit per problem**.

Table 1(left) shows the performance of different models and human averages on IMO-AG-30. While the latest reasoning-enhanced language model of o1 achieved impressive results on a range of tasks, we observed that large language models still struggle with rigorous mathematical

reasoning in geometry, often generating proofs with erroneous logics and hallucinated intermediate results. AlphaGeometry’s DD+AR approach notably improved upon Wu’s method (10 solves), yet its symbolic reasoning engine remains both redundant and suboptimal. By redesigning and optimizing the deductive database method, TongGeometry’s reasoning backend, using only DD, successfully solved 18 problems — outperforming average IMO contestants. With a learned value heuristics, TongGeometry became the first method to surpass IMO gold medalists, successfully proving all 30 problems in the benchmark. A closer examination revealed that the value model was instrumental in solving the benchmark’s two most challenging problems: IMO 2000 P6 and IMO 2008 P6. Compared to AlphaGeometry, TongGeometry achieved these results on a consumer-grade machine with 32 CPU cores and a single NVIDIA RTX 4090 GPU in a maximum of 38 minutes, whereas AlphaGeometry required 246 CPU cores and 4 NVIDIA V100 GPUs to reduce solve time to under 90 minutes — a resource-intensive setup inaccessible to most users.

Table 1(right) displays the performance of various models on the MO-TG-225 dataset. Unlike the IMO dataset, MO-TG-225 includes problems from diverse sources, making rigorous human evaluation benchmarks unavailable. During the evaluation of GPT-4 and o1, we observed that when prompted to prove a known theorem, such as the Euler line theorem, LLMs frequently assumed the theorem as established and applied it directly without the requested proof. We considered these responses as correct, accounting for a few solves reported in the table. Consistent with results on IMO-AG-30, TongGeometry’s DD backend demonstrated improved problem-solving capability over AlphaGeometry’s DD+AR, reaching a performance level close to AlphaGeometry overall. We noted that AlphaGeometry’s success largely stemmed from its backend engine, with 72.5% of its total solves achieved by DD+AR. In contrast, TongGeometry not only solved a greater proportion of problems (81.3% vs. 45.3%) but also leveraged its neural models to address auxiliary construction effectively, with only 55.2% solved by DD.

Ablation study on the value heuristics suggests that under a resource-constrained environment of 32 cores and 1 graphics card, the value model could squeeze out the last bit of performance, reaching 7.1% (IMO-AG-30) and 3.4% (MO-TG-225) improvement over a policy-only model.

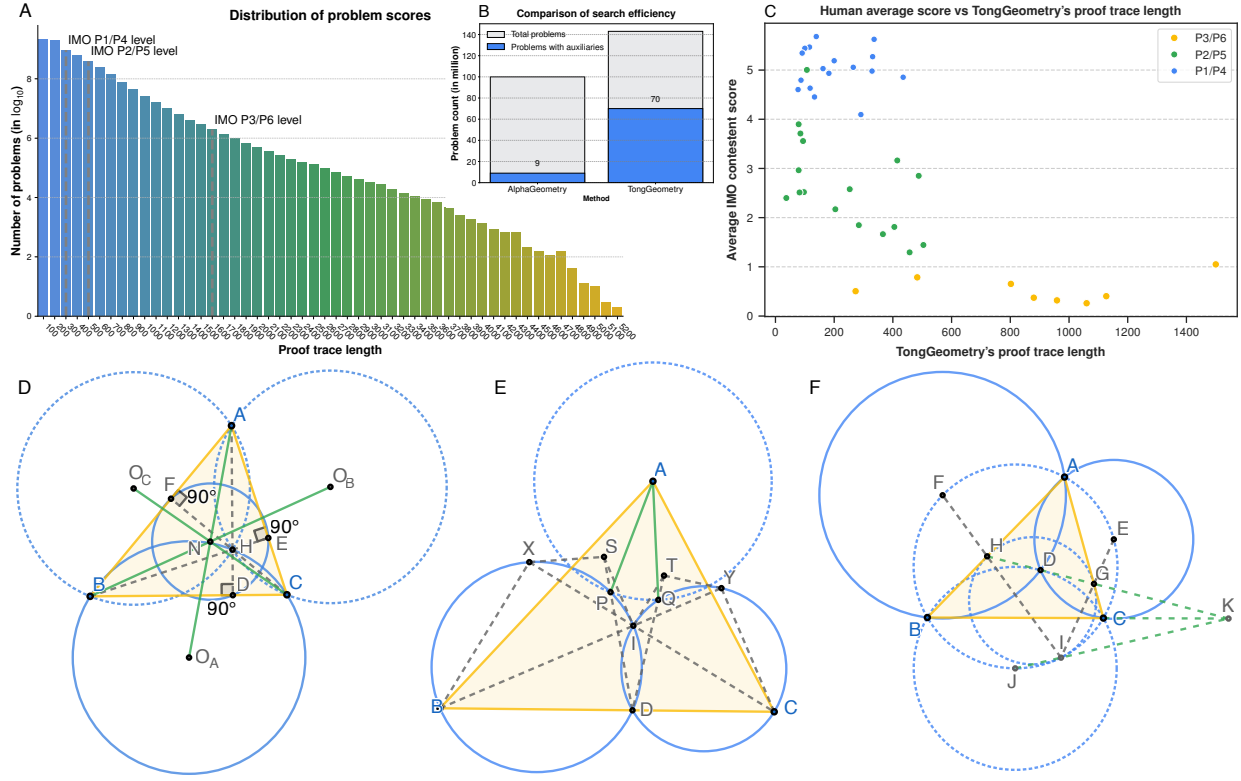
## Expert evaluation on TongGeometry results

Figure 2 shows IMO 2024 P4, the geometry problem in the latest IMO competition, a relatively new problem without many documented solutions at the time of TongGeometry training. TongGeometry generated two equivalent solutions. In retrospect, we observed that the two solutions introduced an auxiliary point identical to what’s presented in the official solution (21). We invited an 2024 IMO gold medalist to evaluate TongGeometry’s solutions, and both were deemed correct.

## Conclusion

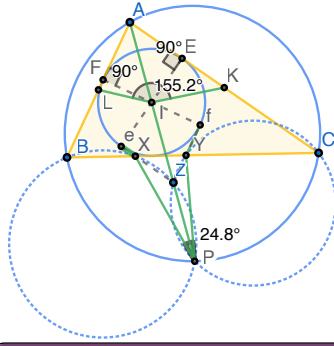
In this work, we introduce TongGeometry, a neuro-symbolic system designed to discover, present, and prove IMO-level geometry problems through guided tree search. We implemented evaluation rubrics for assessing the quality of TongGeometry-proposed problems. Notably, one proposed problem was selected in 2024 National High School Mathematics League (Beijing), a Team China qualifying competition, and two others shortlisted in 2024 US Ersatz Math Olympiad, a top civil math competition in the US. TongGeometry employs actor-critic style inference, where a policy model is learned to complete auxiliaries and a value model to estimate the number of remaining steps until problem solved. This style of problem solving leads to super-gold-medalist performance for the first time on IMO problems, surpasses the previous state-of-the-art, and demonstrates more effective use of neural models for auxiliary construction.

By analogy, TongGeometry functions more like a coach rather than a student merely solving problems — a distinction that opens avenues for future study to further advance the field of geometry and mathematics.



**Figure 1: Generated data, statistics, and their distribution.** (A) The distribution of proof trace length in the generated problems. (B) Comparison of search efficiency of our method and AlphaGeometry. (C) The relationship of proof trace length and problem difficulty. (D) A known lemma discovered by TongGeometry. (E) A symmetric problem proposed by TongGeometry in 2024 USEMO shortlist that asks to prove  $AP$  and  $AQ$  are of the same length. (F) The Mixtilinear Incircle configuration found in TongGeometry's search buffer.





### Problem statement

Let triangle  $ABC$  with incenter  $I$  satisfying  $AB < AC < BC$ . Let  $X$  be a point on line  $BC$ , different from  $C$ , such that the line through  $X$  and parallel to  $AC$  is tangent to the incircle. Similarly, let  $Y$  be a point on line  $BC$ , different from  $B$ , such that the line through  $Y$  and parallel to  $AB$  is tangent to the incircle. Line  $AI$  intersects the circumcircle of triangle  $ABC$  again at  $P$ . Let  $K$  and  $L$  be the midpoints of  $AC$  and  $AB$ , respectively. Prove that  $\angle KIL + \angle YPX = 180^\circ$ .

### Solution 1: $Z = \text{ExtendEqual}(A, I)$

```
[1] contri (AFI, ZfI) because eqangle (AIf, ZIf)[0], ...
[2] contri (AEI, ZeI) because eqangle (AIE, ZIE)[0], ...
...
# Key step YZf collinear, ZXe collinear

[19] eqline (Yf, Zf) because eqangle (If, YfZ)[13], ...
[20] eqline (Xe, Ze) because eqangle (IE, XeZ)[14], ...
...
# Key step CYZP concyclic, BXZP concyclic

[25] eqcircle (None, CPY, None, PYZ) because eqangle (PCY, PZY)[21]
...
[27] eqcircle (None, BPX, None, PXZ) because eqangle (PBX, PZX)[23]
...
# Angle chasing

[33] eqangle (KIL, XPY) because eqangle (AIK, PYX)[31], ...
```

### Solution 2: $Z = \text{IntersectLineLine}(X, e, Y, f)$

```
[1] cong (EF, ef) because contri (EFI, efI)[0]
[2] cong (Pe, Pf) because contri (IPe, IPf)[0]
[3] cong (Ae, Af) because contri (AIE, AIf)[0]
...
# Key step AIZ collinear

[21] eqline (AI, IZ) because eqangle (AIZ, YfZ)[12], ...
[22] eqline (IP, IZ) because eqangle (PIZ, YfZ)[13], ...
...
# Key step I is midpoint of AZ

[26] cong (AI, IZ) because contri (AEI, ZfI)[20]
[27] midp (I, AZ) because cong (AI, IZ)[26], ...
...
# CYZP concyclic, BXZP concyclic, and angle chase

[36] eqcircle (None, BPX, None, PXZ) because eqangle (PBX, PZX)[32]
[37] eqcircle (None, CPY, None, PYZ) because eqangle (PCY, PZY)[33]
...
[46] eqangle (KIL, XPY) because eqangle (AIL, PYX)[44], ...
```

**Figure 2: IMO 2024 P4 proof from TongGeometry.** TongGeometry generated two possible solutions to IMO 2024 P4, a problem that did not have well-documented solutions during TongGeometry training. TongGeometry constructed the same auxiliary point as the official solution. TongGeometry’s proof was deemed correct by a 2024 IMO gold medalist. Since TongGeometry uses full angles, the original statement is equivalent to  $\angle KIL = \angle XPY$ .

**Table 1: Number of problems solved in IMO-AG-30 (left) and MO-AG-225 (right) in ascending order.** TongGeometry w/o value denotes TongGeometry without using the valued model trained with process-based reward learning.

Method	Solves in IMO-AG-30	Solves in MO-TG-225
GPT-4o	0 / 30	2 / 225
o1-preview	0 / 30	6 / 225
DD+AR (AlphaGeometry)	14 / 30	74 / 225
Honorable mentions	14.3 / 30	-
Average IMO contestant	15.2 / 30	-
DD (TongGeometry)	18 / 30	101 / 225
Bronze medalist	19.3 / 30	-
Silver medalist	22.9 / 30	-
AlphaGeometry	25 / 30	102 / 225
Gold medalist	25.9 / 30	-
TongGeometry w/o value	28 / 30	177 / 225
TongGeometry	30 / 30	183 / 225

## References and Notes

1. D. Selsam, *et al.*, IMO Grand Challenge. URL <https://imo-grand-challenge.github.io> (2020).
2. T. H. Trinh, Y. Wu, Q. V. Le, H. He, T. Luong, Solving olympiad geometry without human demonstrations. *Nature* **625** (7995), 476–482 (2024).
3. S. Polu, I. Sutskever, Generative language modeling for automated theorem proving. *arXiv preprint arXiv:2009.03393* (2020).
4. K. Zheng, J. M. Han, S. Polu, Minif2f: a cross-system benchmark for formal olympiad-level mathematics. *arXiv preprint arXiv:2109.00110* (2021).
5. S. Polu, *et al.*, Formal mathematics statement curriculum learning. *arXiv preprint arXiv:2202.01344* (2022).
6. G. Lample, *et al.*, Hypertree proof search for neural theorem proving. *Advances in neural information processing systems* **35**, 26337–26349 (2022).
7. J. Achiam, *et al.*, Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).
8. H. Touvron, *et al.*, Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* (2023).
9. W. W. R. Ball, *A short account of the history of mathematics* (Courier Corporation) (1960).
10. F. Marić, *et al.*, Formalizing IMO problems and solutions in Isabelle/HOL. *arXiv preprint arXiv:2010.16015* (2020).
11. L. d. Moura, S. Ullrich, The Lean 4 theorem prover and programming language, in *Automated Deduction—CADE 28: 28th International Conference on Automated Deduction, Virtual Event, July 12–15, 2021, Proceedings 28* (Springer) (2021), pp. 625–635.
12. S.-C. Chou, X.-S. Gao, J.-Z. Zhang, An introduction to geometry expert, in *Automated Deduction—CADE-13: 13th International Conference on Automated Deduction New Brunswick, NJ, USA, July 30–August 3, 1996 Proceedings 13* (Springer) (1996), pp. 235–239.

13. C. Sangwin, A brief review of GeoGebra: dynamic mathematics. *MSor Connections* **7** (2), 36 (2007).
14. P. Bak, R. S. Krajči, M. M. Rolínek, *Automated generation of planar geometry Olympiad problems*, Ph.D. thesis, Master Thesis (2020).
15. D. Guo, *et al.*, DeepSeek-Coder: When the Large Language Model Meets Programming—The Rise of Code Intelligence. *arXiv preprint arXiv:2401.14196* (2024).
16. J. Hoffmann, *et al.*, Training compute-optimal large language models. *arXiv preprint arXiv:2203.15556* (2022).
17. J. S. Mackay, History of the Nine-point Circle. *Proceedings of the Edinburgh Mathematical Society* **11**, 19–57 (1892).
18. E. Chen, *Euclidean geometry in mathematical olympiads*, vol. 27 (American Mathematical Soc.) (2021).
19. E. Chen, A Guessing Game: Mixtilinear Incircles. URL <https://web.evanchen.cc/handouts/Mixt-GeoGuessr/Mixt-GeoGuessr.pdf> (2015).
20. E. Chen, The Incenter/Excenter Lemma. URL <https://web.evanchen.cc/handouts/Fact5/Fact5.pdf> (2016).
21. T. O. Committee, the Problem Selection Committee of IMO 2024, 65th International Mathematical Olympiad Problems with Solutions. URL <https://www.imo2024.uk/solutions> (2024).
22. W.-t. Wu, On the decision problem and the mechanization of theorem-proving in elementary geometry, in *Selected Works Of Wen-Tsun Wu* (World Scientific), pp. 117–138 (2008).
23. S.-C. Chou, *Mechanical geometry theorem proving* (Springer) (1988).
24. S.-C. Chou, *Proving elementary geometry theorems using Wu’s algorithm*, Ph.D. thesis, University of Texas at Austin (1984).
25. S.-C. Chou, W. F. Schelter, Proving geometry theorems with rewrite rules. *Journal of Automated Reasoning* **2**, 253–273 (1986).

26. S.-C. Chou, X.-S. Gao, J.-Z. Zhang, A deductive database approach to automated geometry theorem proving and discovering. *Journal of Automated Reasoning* **25** (3), 219–246 (2000).
27. S.-C. Chou, X. Gao, J.-Z. Zhang, *Machine proofs in geometry: Automated production of readable proofs for geometry theorems*, vol. 6 (World Scientific) (1994).
28. S.-C. Chou, X.-S. Gao, J.-Z. Zhang, Automated generation of readable proofs with geometric invariants: II. Theorem proving with full-angles. *Journal of Automated Reasoning* **17** (3), 349–370 (1996).
29. D. Hilbert, *The foundations of geometry* (Open court publishing Company) (1902).
30. A. Tarski, A decision method for elementary algebra and geometry, in *Quantifier elimination and cylindrical algebraic decomposition* (Springer), pp. 24–84 (1998).
31. N. Matsuda, K. Vanlehn, Gramy: A geometry theorem prover capable of construction. *Journal of Automated Reasoning* **32**, 3–33 (2004).
32. K. Wang, Z. Su, Automated geometry theorem proving for human-readable proofs, in *Twenty-Fourth International Joint Conference on Artificial Intelligence* (2015).
33. S. Schulz, E—a brainiac theorem prover. *Ai Communications* **15** (2-3), 111–126 (2002).
34. C. Weidenbach, *et al.*, SPASS Version 3.5, in *Automated Deduction—CADE-22: 22nd International Conference on Automated Deduction, Montreal, Canada, August 2-7, 2009. Proceedings 22* (Springer) (2009), pp. 140–145.
35. L. De Moura, N. Bjørner, Z3: An efficient SMT solver, in *International conference on Tools and Algorithms for the Construction and Analysis of Systems* (Springer) (2008), pp. 337–340.
36. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. *nature* **521** (7553), 436–444 (2015).
37. S. Loos, G. Irving, C. Szegedy, C. Kaliszyk, Deep network guided proof search. *arXiv preprint arXiv:1701.06972* (2017).
38. C. Wei, M. Sun, W. Wang, Proving Olympiad Algebraic Inequalities without Human Demonstrations. *arXiv preprint arXiv:2406.14219* (2024).

39. D. Silver, *et al.*, Mastering the game of Go with deep neural networks and tree search. *nature* **529** (7587), 484–489 (2016).
40. D. Silver, *et al.*, Mastering the game of go without human knowledge. *nature* **550** (7676), 354–359 (2017).
41. D. Silver, *et al.*, A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* **362** (6419), 1140–1144 (2018).
42. N. Megill, D. A. Wheeler, *Metamath: a computer language for mathematical proofs* (Lulu.com) (2019).
43. E. Aygün, *et al.*, Proving theorems using incremental learning and hindsight experience replay, in *International Conference on Machine Learning* (PMLR) (2022), pp. 1198–1210.
44. V. Firoiu, *et al.*, Training a first-order theorem prover from synthetic data. *arXiv preprint arXiv:2103.03798* (2021).
45. B. Liu, *et al.*, Tinygsm: achieving 80% on gsm8k with small language models. *arXiv preprint arXiv:2312.09241* (2023).
46. Y. Wu, *et al.*, Autoformalization with large language models. *Advances in Neural Information Processing Systems* **35**, 32353–32368 (2022).
47. J. Uesato, *et al.*, Solving math word problems with process-and outcome-based feedback. *arXiv preprint arXiv:2211.14275* (2022).
48. H. Lightman, *et al.*, Let’s verify step by step. *arXiv preprint arXiv:2305.20050* (2023).

## Acknowledgments

This project is a two-year long collaboration between Beijing Institute for General Artificial Intelligence and Peking University. Along the journey, we would like to thank Jingyu Xie (Winter Camp) and Xingjian Liang (2023 IMO Gold Medalist) for assistance in problem examination, Patrik Bak (GeoGen author) for very thoughtful and detailed discussion for designing automatic rubrics and

generating good proposals, Evan Chen (USA IMO coach) for assistance in problem evaluation in USEMO, Chen Zhen for crafting beautiful figures, Qing Li and Junqi Wang for early discussion, Liangru Xiang and Aoyang Qin for consistent assistance, and Xianbang Wang (2024 IMO Gold Medalist) for generated proof assessment.

**Funding:** C. Z., J. S., S. L., Y. L., Y. M., W. W., and S.-C. Z. were supported in part by Beijing Institute for General Artificial Intelligence (BIGAI) (No. KY20220060). All authors were supported in part by the National Science and Technology Major Project (No. 2022ZD0114900). Y. M. and Y. Z. were supported in part by the National Natural Science Foundation of China (No. 62376031).

**Author contributions:** Conceptualization: C. Z., Y. L.

Data curation: C. Z.

Formal analysis: C. Z., J. S., S. L.

Funding acquisition: Y. Z., S.-C. Z.

Investigation: C. Z.

Methodology: C. Z., Y. L.

Project administration: C. Z., W. W., S.-C. Z.

Software: C. Z., J. S.

Resources: Y. M., Y. Z., S.-C. Z.

Supervision: W. W., S.-C. Z.

Validation: C. Z.

Visualization: C. Z.

Writing - original draft: C. Z.

Writing - review & editing: C. Z., Y. L., Y. Z.

**Competing interests:** There are no competing interests to declare.

**Supplementary Materials for**  
**Proposing and solving olympiad geometry with guided tree**  
**search**

Chi Zhang\*, Jiajun Song, Siyu Li, Yitao Liang, Yuxi Ma

Wei Wang, Yixin Zhu\*, Song-Chun Zhu\*

\*Corresponding author. Email: zhangchi@bigai.ai, yixin.zhu@pku.edu.cn, sczhu@bigai.ai



## Supplementary Text

### Automated geometry theorem proposing and proving

GeoGen (14) is currently the only known program capable of automatically proposing olympiad geometry theorems. Like our approach, GeoGen begins with a blank canvas and randomly draws geometric elements. However, GeoGen is more accurately described as a theorem guesser rather than a theorem prover. While it can verify whether a fact holds numerically after constructing a diagram, it lacks the ability to construct auxiliary elements necessary for rigorous proof. As a result, the proving process relies heavily on human involvement. In contrast, TongGeometry not only proposes problems but also autonomously proves theorems.

Automatic geometry theorem proving methods can be broadly categorized into two approaches: algebraic methods and search-based technique. Wu's method (22, 23) is the most prominent in the former line of research, in which a geometry theorem is formulated using a system of algebraic equations. The conclusion is considered proven if the system, after reduction (24, 25), has a zero remainder. While Wu's method can theoretically check the validity of any equality fact for geometry problems, there are two major concerns that hinder the wider adoption of Wu's method in geometry theorem proving. For one thing, Wu's method is purely algebraic and of limited interest to human readers. The proof process is mechanical, with reduction and equation solving carrying little semantic meaning corresponding to the diagrams, much as how contemporary mathematicians view analytical proving as less elegant. For another, Wu's method is inefficient in terms of space and time complexity, limiting its practical use.

With the development of deductive database (26), area method (27), and full angles (28), the research community is more invested in the second, more human-like approach for geometry theorem proving. The Lean interactive theorem prover (11) has also incorporated geometry libraries; yet these libraries based on either completed Euclid's axioms, Hilbert's axioms (29), or Tarski's axioms (30) are very limited in their ability to prove olympiad-level problems. Besides, proof using axiomatic geometry is less intuitive and not as easily understandable as plain synthetic geometry. In addition, none of the methods properly address how the auxiliaries should be constructed to advance reasoning, other than using manually designed heuristics (31, 32). AlphaGeometry (2) sidesteps this challenge using a data-driven method similar to ours. However, their data engine is

inadequately designed to quantify the space of olympiad geometry, leaving it underexplored and resulting in performance that falls short of top-tier competition standards.

## **Learning-based methods for theorem proving**

The pursuit of automated theorem proving has its roots in the early development of logical systems. Building on this foundation, a range of powerful logic solvers, such as E (33), SPASS (34), and Z3 (35), have been developed. Recent advancements in interactive proof assistants, like Lean (11), have enabled the resolution of more complex theorems, facilitated by human guidance. Simultaneously, the AI community has been working towards minimizing human intervention in the theorem proving process by generating programmatic input, thus pushing the boundaries of automation. Following the breakthroughs in deep learning (36), learning-based approaches have become the predominant focus of research in this area.

The deep learning era has ushered in two distinct streams of research in learning-based theorem proving: one focusing on learning-guided search and the other on improved data curation.

In learning-guided search, deep learning models serve as hint providers for the next steps taken by proof assistants (37, 38). Due to the absence of a direct mechanism to evaluate the utility of these actions, the log probability of the action is often employed as a heuristics (3). Utilizing best-first search or beam search, these systems have successfully solved a few simple olympiad problems of algebra and number theory (4) and certain geometry problems (2). The success of learning a value function and its integration with a policy function in Go (39–41) have inspired the community to explore learning a utility function for theorem proving actions. For instance, HyperTree (6) employs AlphaZero-style Monte Carlo Tree Search to learn both a policy model and a critics model to guide the proof search.

High-quality data is considered a crucial element for the success of learning-based methods. However, in contrast to other domains, mathematical data is substantially more scarce. The most widely used datasets, such as Metamath (42) and Mathlib (11), contain fewer than 60,000 human-written theorems combined. The miniF2F benchmark (4, 5) includes only a few hundred validation and test problems. Quality mathematical data is difficult to find naturally; for instance, the IMO produces just six problems per year. This scarcity has led researchers to explore synthetic data as an alternative. Models trained on data generated from specially designed symbolic systems

have demonstrated impressive performance (2, 43, 44). Additionally, advancements in LLMs have prompted investigations into their use as data generators (45). Informal proofs have also been used to augment data via autoformalization (46). An emerging technique in data curation is process-based supervision (47, 48), where each step of the proof is supervised. However, this approach requires more extensive annotation efforts compared to outcome-based supervision.

### Explanation on the Nine-point Center

In Figure 1(D), in  $\triangle ABC$ ,  $D$ ,  $E$ , and  $F$  are the feet corresponding to  $A$ ,  $B$ , and  $C$ , respectively. Let  $H$  be the orthocenter of  $\triangle ABC$  and  $O_A$  be the circumcenter of  $\triangle BCH$ . TongGeometry finds that the circumcenter  $N$  of  $\triangle DEF$  happens to be the midpoint of  $AO_A$ . The fact builds on several important theorems in geometry literature: the circumcircle of  $\triangle DEF$  is the famous Nine-point circle (17) and the circumcenter of  $\triangle ABC$  and the nine-point circle center of  $\triangle ABC$  are homothetic with respect to its orthocenter in a ratio of 1:2 (18). Via angle chasing, one can prove  $O_A$  and the circumcenter of  $\triangle ABC$  are symmetric with respect to  $BC$ , hence the result. Note that the original problem is mirror-symmetric in that  $B - C$  and  $E - F$  are symmetric pairs. In addition, the problem can be made rotationally symmetric, by constructing circumcircles of  $\triangle ACH$  and  $\triangle ABH$  and claiming that  $N$  is the midpoint of  $AO_A$ ,  $BO_B$  and  $CO_C$ .

### Explanation on the Mixtilinear Incircle

In Figure 1(F), the base configuration is described as follows: in  $\triangle ABC$ , let  $D$  be its incenter, and  $E$  and  $F$  the circumcenters of  $\triangle ACD$  and  $\triangle ABD$ , respectively. Points  $H$  and  $G$  lie on  $AB$  and  $AC$ , respectively, such that  $H$ ,  $D$ , and  $G$  are collinear and perpendicular to  $AD$ . This foundational setup reveals multiple intrinsic geometric properties. Notably,  $ABCEF$  are concyclic, with  $E$  and  $F$  marking the arc midpoints. The intersection of  $EG$  and  $FH$ , labeled  $I$ , forms a circle with  $H$  and  $G$ , which is tangent to  $AB$ ,  $AC$ , and  $\odot(ABC)$  at  $H$ ,  $G$ , and  $I$ , respectively. Denoting the circumcenter of  $\triangle BCD$  as  $J$ , the lines  $HG$ ,  $BC$ , and  $IJ$  are concurrent at a point  $K$ .

## Limitation

Detailed analysis suggests that current models still fall short in sequentially constructing a series of interrelated actions, failing on problems requiring multiple new points to be constructed. Such a problem could be rooted in the data-driven nature of neural models: the longer the number of steps to take, the (exponentially) fewer the data points. Additionally, a disconnect exists between the natural language representation of geometry problems and the domain-specific languages used in computational settings. While the computer programs use direct sequential construction, how to correctly convert natural language representation to computer construction remains an under-explored issue. Human-proposed problems frequently obscure straightforward constructions with constraints, introducing intentional complexities that contribute to the problem's difficulty.

## IMO 2024 P4 Full Solutions

Z = ExtendEqual(A, I)

- [1] contri (AFI, ZfI) because eqangle (AIF, ZIf)[0], cong (FI, If)[0], cong (AI, IZ)[0]
- [2] contri (AEI, ZeI) because eqangle (AIE, ZIe)[0], cong (EI, Ie)[0], cong (AI, IZ)[0]
- [3] simtri (EFf, ZfI) because eqangle (EeF, ZIf)[0], eqratio (Ee, IZ, Fe, If)[0]
- [4] simtri (EFf, ZeI) because eqangle (EeF, eIZ)[0], eqratio (Ee, IZ, Fe, Ie)[0]
- [5] simtri (ACZ, AKI) because eqangle (CAZ, KAI)[0], eqratio (AC, AK, AZ, AI)[0]
- [6] simtri (ABZ, ALI) because eqangle (BAZ, LAI)[0], eqratio (AB, AL, AZ, AI)[0]
- [7] eqangle (AFI, ZfI) because contri (AFI, ZfI)[1]
- [8] eqangle (AEI, ZeI) because contri (AEI, ZeI)[2]
- [9] eqangle (FEf, fZI) because simtri (EFf, ZfI)[3]
- [10] eqangle (FEf, IZe) because simtri (EFf, ZeI)[4]
- [11] eqangle (AIK, AZC) because simtri (ACZ, AKI)[5]
- [12] eqangle (AIL, AZB) because simtri (ABZ, ALI)[6]
- [13] eqangle (IFf, YfZ) because eqangle (AFI, ZfI)[7], eqangle (AFf, YfI)[0]
- [14] eqangle (IEe, XeZ) because eqangle (AEI, ZeI)[8], eqangle (AEe, XeI)[0]
- [15] eqangle (AZf, PCY) because eqangle (FEf, YCP)[0], eqangle (FEf, fZI)[9], eqangle (AZf, IZf)[0]
- [16] eqangle (AZe, PBX) because eqangle (AZe, IZe)[0], eqangle (FEf, IZe)[10], eqangle (FEf, PBX)[0]
- [17] eqangle (AIK, PZC) because eqangle (AIK, AZC)[11], eqangle (AZC, PZC)[0]
- [18] eqangle (AIL, PZB) because eqangle (AIL, AZB)[12], eqangle (AZB, PZB)[0]
- [19] eqline (Yf, Zf) because eqangle (IFf, YfZ)[13], eqline (FI, Ff)[0]
- [20] eqline (Xe, Ze) because eqangle (IEe, XeZ)[14], eqline (EI, Ee)[0]
- [21] eqangle (PCY, PZY) because eqangle (AZY, PZY)[0], eqangle (AZY, AZf)[19], eqangle (AZf, PCY)[15]
- [22] eqangle (CYf, XYZ) because eqangle (AYZ, AYf)[19], eqangle (AYC, AYZ)[0]
- [23] eqangle (PBX, PZX) because eqangle (AZX, PZX)[0], eqangle (AZX, AZe)[20], eqangle (AZe, PBX)[16]
- [24] eqangle (BXe, YXZ) because eqangle (AXZ, AXe)[20], eqangle (AXB, AXY)[0]
- [25] eqcircle (None, CPY, None, PYZ) because eqangle (PCY, PZY)[21]
- [26] eqangle (CYP, XYP) because eqangle (PYZ, PYf)[19], eqangle (CYf, XYZ)[22]

[27] eqcircle (None, BPX, None, PXZ) because eqangle (PBX, PZX)[23]  
 [28] eqangle (BXP, YXP) because eqangle (PXZ, PXe)[20], eqangle (BXe, YXZ)[24]  
 [29] eqangle (CYP, CZP) because eqcircle (None, CPY, None, PYZ)[25]  
 [30] eqangle (BXP, BZP) because eqcircle (None, BPX, None, PXZ)[27]  
 [31] eqangle (AIK, PYX) because eqangle (AIK, PZC)[17], eqangle (CYP, CZP)[29], eqangle (CYP, XYP)[26]  
 [32] eqangle (AIL, PXY) because eqangle (AIL, PZB)[18], eqangle (BXP, BZP)[30], eqangle (BXP, YXP)[28]  
 [33] eqangle (KIL, XPY) because eqangle (AIK, PYX)[31], eqangle (AIL, PXY)[32]

Z = IntersectLineLine(X, e, Y, f)

[1] cong (EF, ef) because contri (EFI, efi)[0]  
 [2] cong (Pe, Pf) because contri (IPe, IPf)[0]  
 [3] cong (Ae, Af) because contri (AIE, AIf)[0]  
 [4] simtri (AEF, Zfe) because eqangle (AFE, feZ)[0], eqangle (AEF, efZ)[0]  
 [5] eqcircle (None, IZe, None, IZf) because eqangle (IeZ, IfZ)[0]  
 [6] eqratio (AE, Zf, EF, ef) because simtri (AEF, Zfe)[4]  
 [7] eqangle (ZIE, Zfe) because eqcircle (None, IZe, None, IZf)[5]  
 [8] eqangle (IZf, Ief) because eqcircle (None, IZe, None, IZf)[5]  
 [9] contri (PZe, PZf) because cong (Ze, Zf)[0], cong (Pe, Pf)[2]  
 [10] contri (AZe, AZf) because cong (Ze, Zf)[0], cong (Ae, Af)[3]  
 [11] cong (AE, Zf) because eqratio (AE, Zf, EF, ef)[6], cong (EF, ef)[1]  
 [12] eqangle (AIZ, YfZ) because eqangle (ZIE, Zfe)[7], eqangle (AIE, Yfe)[0]  
 [13] eqangle (PIZ, YfZ) because eqangle (ZIE, Zfe)[7], eqangle (PIe, Yfe)[0]  
 [14] eqangle (IAL, IZY) because eqangle (IAL, Ief)[0], eqangle (IZf, Ief)[8], eqangle (IZY, IZf)[0]  
 [15] eqangle (IZY, XBP) because eqangle (Ief, XBP)[0], eqangle (IZf, Ief)[8], eqangle (IZY, IZf)[0]  
 [16] eqangle (IAK, YZI) because eqangle (IAK, feI)[0], eqangle (IZf, Ief)[8], eqangle (IZY, IZf)[0]  
 [17] eqangle (IZY, PCY) because eqangle (Ief, PCY)[0], eqangle (IZf, Ief)[8], eqangle (IZY, IZf)[0]  
 [18] eqangle (PZe, fZP) because contri (PZe, PZf)[9]  
 [19] eqangle (AZe, fZA) because contri (AZe, AZf)[10]  
 [20] contri (AEI, ZfI) because cong (AE, Zf)[11], perp (AEI)[0], perp (IfZ)[0], cong (EI, If)[0]  
 [21] eqline (AI, IZ) because eqangle (AIZ, YfZ)[12], eqline (Yf, Zf)[0]  
 [22] eqline (IP, IZ) because eqangle (PIZ, YfZ)[13], eqline (Yf, Zf)[0]  
 [23] eqangle (PZX, YZP) because eqangle (PZX, PZe)[0], eqangle (PZe, fZP)[18], eqangle (PZY, PZf)[0]  
 [24] eqangle (AZP, PZA) because eqangle (PZe, fZP)[18], eqangle (AZe, fZA)[19]  
 [25] eqangle (AZY, CAZ) because eqangle (AZe, ZAC)[0], eqangle (AZe, fZA)[19], eqangle (AZY, AZf)[0]  
 [26] cong (AI, IZ) because contri (AEI, ZfI)[20]  
 [27] midp (I, AZ) because cong (AI, IZ)[26], eqline (AI, IZ)[21]  
 [28] eqratio (AB, AL, AZ, AI) because midp (I, AZ)[27], midp (L, AB)[0]  
 [29] eqratio (AC, AK, AZ, AI) because midp (I, AZ)[27], midp (K, AC)[0]  
 [30] eqangle (BAZ, LAI) because eqangle (AZY, ZAB)[0], eqangle (AZY, IZY)[21], eqangle (IAL, IZY)[14]  
 [31] eqangle (CAZ, KAI) because eqangle (AZY, CAZ)[25], eqangle (AZY, IZY)[21], eqangle (IAK, YZI)[16]  
 [32] eqangle (PBX, PZX) because eqangle (IZY, XBP)[15], eqangle (IZY, PZY)[22], eqangle (PZX, YZP)[23]  
 [33] eqangle (PCY, PZY) because eqangle (IZY, PCY)[17], eqangle (IZY, PZY)[22]  
 [34] simtri (ABZ, ALI) because eqratio (AB, AL, AZ, AI)[28], eqangle (BAZ, LAI)[30]  
 [35] simtri (ACZ, AKI) because eqratio (AC, AK, AZ, AI)[29], eqangle (CAZ, KAI)[31]  
 [36] eqcircle (None, BPX, None, PXZ) because eqangle (PBX, PZX)[32]  
 [37] eqcircle (None, CPY, None, PYZ) because eqangle (PCY, PZY)[33]

[38] eqangle (AIL, AZB) because simtri (ABZ, ALI)[34]  
 [39] eqangle (AIK, AZC) because simtri (ACZ, AKI)[35]  
 [40] eqangle (BXP, BZP) because eqcircle (None, BPX, None, PXZ)[36]  
 [41] eqangle (CYP, CZP) because eqcircle (None, CPY, None, PYZ)[37]  
 [42] eqangle (AZB, PXY) because eqangle (AZB, PZB)[24], eqangle (BXP, BZP)[40], eqangle (BXP, YXP)[0]  
 [43] eqangle (AZC, PYX) because eqangle (AZC, PZC)[24], eqangle (CYP, CZP)[41], eqangle (CYP, XYP)[0]  
 [44] eqangle (AIL, PXY) because eqangle (AIL, AZB)[38], eqangle (AZB, PXY)[42]  
 [45] eqangle (AIK, PYX) because eqangle (AIK, AZC)[39], eqangle (AZC, PYX)[43]  
 [46] eqangle (KIL, XPY) because eqangle (AIL, PXY)[44], eqangle (AIK, PYX)[45]

## TongGeometry Proposals to 2024 USEMO

1. Figure S2: Let  $\triangle ABC$  be an acute-angled triangle,  $D$  be its incenter,  $E$  be its  $B$ -excenter and  $F$  its  $A$ -excenter.  $G$  is the center of  $\odot(CBE)$  and  $H$  is the center of  $\odot(ACF)$ . A circumcircle centered at  $I$  passes through  $\triangle DEF$ .  $J$  is the foot of  $I$  to  $CD$ . Prove  $GHIJ$  are concyclic.
2. Figure S3: Let  $\triangle ABC$  be an acute-angled triangle and  $D$  be its incenter. Let  $E$  be the circumcenter of  $\odot(ACD)$  and  $F$  be that of  $\odot(BCD)$ . Project  $D$  to  $CE$  to get  $G$  and project  $D$  to  $CF$  to get  $H$ . Denote the other intersection of  $EC$  with  $\odot(BCD)$  as  $I$  and that of  $FC$  with  $\odot(ACD)$  as  $J$ .  $ID$  and  $JD$  intersect with  $HG$  at  $K$  and  $L$ , respectively. Prove  $I$  and  $J$  and the other intersection of  $LC$  with  $\odot(CDB)$  and that of  $KC$  with  $\odot(DCA)$  concyclic.
3. Figure S4: Let  $\triangle ABC$  be an acute-angled triangle and  $D$  be its incenter.  $E$  is on  $CD$  such that  $\angle EAC$  is a right angle.  $F$  is on  $CD$  such that  $\angle FBC$  is a right angle. The circumcircle of  $\triangle AED$  meets  $BC$  again at  $I$  and the circumcircle of  $\triangle BFD$  meets  $BC$  again at  $J$ . Let  $K$  be the midpoint of  $IJ$ . Denote the second intersection of  $IF$  with  $\odot(BFD)$  and that of  $EJ$  with  $\odot(AED)$  as  $L$  and  $M$ , respectively. Prove  $KL = KM$ .
4. Figure S5 (**shortlisted**): Let  $\triangle ABC$  be an acute-angled triangle and  $D$  be its incenter. The projection of  $D$  on  $AB$ , denoted as  $E$ , forms two circumcircles with  $AD$  and  $BD$ , centered at  $F$  and  $G$  respectively.  $AD$  intersects  $\odot(BDE)$  again at  $I$  and  $BD$  intersects  $\odot(ADE)$  again at  $H$ . Denote the antipode of  $G$  in  $\odot(GID)$  as  $M$  and that of  $F$  as  $L$ . Prove  $ME$  and  $LE$  intersect  $\odot(BDE)$  and  $\odot(ADE)$  at the same distance from  $C$ .
5. Figure S6 (**shortlisted**): Let  $\triangle ABC$  be a scalene triangle,  $ABCD$  and  $ACBE$  are two parallelograms. Circle  $E$  and  $D$  pass  $A$ . An arbitrary line through  $A$  meets the two circles again at

$Q$  and  $P$ . Let  $Q'$  and  $P'$  denote the reflections of  $Q$  against  $B$  and  $P$  against  $C$ , respectively.  $F$  being the intersection of  $DC$  and  $EB$ , what conditions should  $PQ$  and  $AF$  satisfy to make  $DEQ'P'$  concyclic?

6. Figure S7: Let  $\triangle ABC$  be an acute-angled triangle and  $D$  be  $AB$ 's midpoint.  $E$  is on line  $AB$  and  $CD \perp CE$ . The line linking the midpoint  $H$  of minor arc  $AE$  in  $\odot(ACE)$  and the midpoint  $I$  of minor arc  $BE$  in  $\odot(BCE)$  intersects  $\odot(ACE)$  again at  $J$  and  $\odot(BCE)$  again at  $K$ .  $JE$  meets  $\odot(BCE)$  again at  $L$  and  $KE$  meets  $\odot(ACE)$  at  $M$ . Prove the circumcenter of  $\odot(IJL)$ , the circumcenter of  $\odot(HKM)$  and  $C$  and  $D$  are concyclic.

### TongGeometry Proposals to 2024 NHSML (Beijing)

- Figure S8: Let  $\triangle ABC$  be an acute-angled triangle,  $D, E, F$  be feet corresponding to  $A, B, C$ , respectively.  $G$  is  $\triangle ABC$ 's orthocenter.  $H$  is the circumcenter of  $\odot(DEF)$ .  $I$  is the circumcenter of  $\odot(BCG)$ . Prove  $H$  is the mid point of  $AI$ .
- Figure S9 (**selected**): Let  $\triangle ABC$  be an acute-angled triangle,  $D, E, F$  be feet corresponding to  $A, B, C$ , respectively.  $G$  is  $\triangle ABC$ 's orthocenter.  $H$  is a point on  $AB$  such that  $DH$  is parallel to  $AC$ .  $I$  is the circumcenter of  $\odot(ADH)$ .  $J$  is another intersection of  $AC$  and  $\odot(I)$ .  $K$  is the intersection of  $CF$  and a line through  $J$  parallel to  $EF$ . Prove  $IF = IK$ .
- Figure S10: Let  $\triangle ABC$  be an acute-angled triangle,  $D$  be  $B$ 's foot.  $E$  is on  $AC$  such that  $\angle EBC$  is a right angle.  $F$  is the reflection of  $E$  over  $B$ .  $G$  is the reflection of  $F$  over  $C$ . Denote the intersection of  $BG$  and  $AC$  as  $I$  and the mid point of  $BI$  as  $J$ .  $H$  is the circumcenter of  $\odot(BDF)$ . Prove  $HIEJ$  is a parallelogram.
- Figure S11: Let  $\triangle ABC$  be an acute-angled triangle and  $D$  the intersection of  $AB$  and  $\angle C$ 's internal angle bisector.  $E$  is the projection of  $A$  on  $BC$ .  $F$  is the reflection of  $A$  over  $E$ .  $G$  the reflection of  $F$  over  $C$ .  $H$  is the circumcenter of  $\odot(ACG)$ .  $I$  is another intersection of  $CD$  with  $\odot(H)$ .  $J$  is the projection of  $I$  on  $BC$ .  $K$  is the reflection of  $A$  over  $J$ .  $L$  the reflection of  $C$  over  $J$ .  $KC$  intersects with  $\odot(H)$  at  $M$ .  $O$  is the mid point of  $CM$ .  $P$  is the reflection of  $E$  over  $O$ . Prove  $HL$  is parallel to  $JP$ .

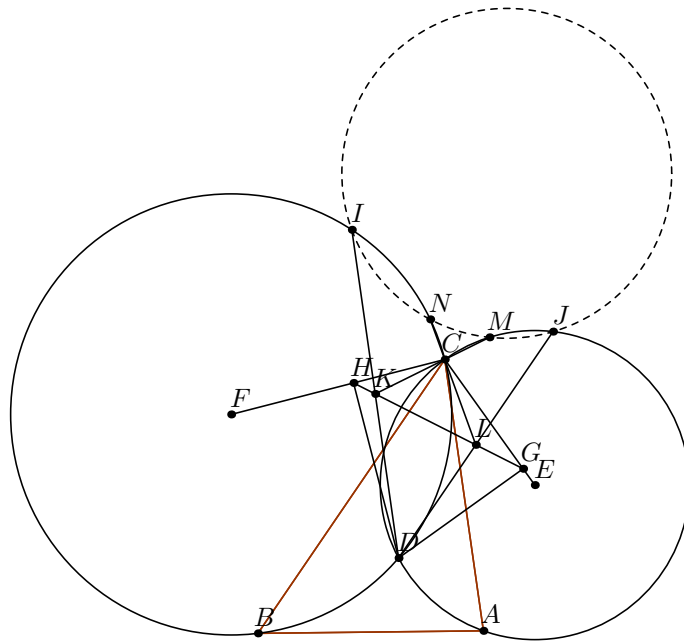
### **Proof of Proposal Acceptance**

Figure S12 is the actual test in 2024 NHSML (Beijing). The second problem is our proposal.

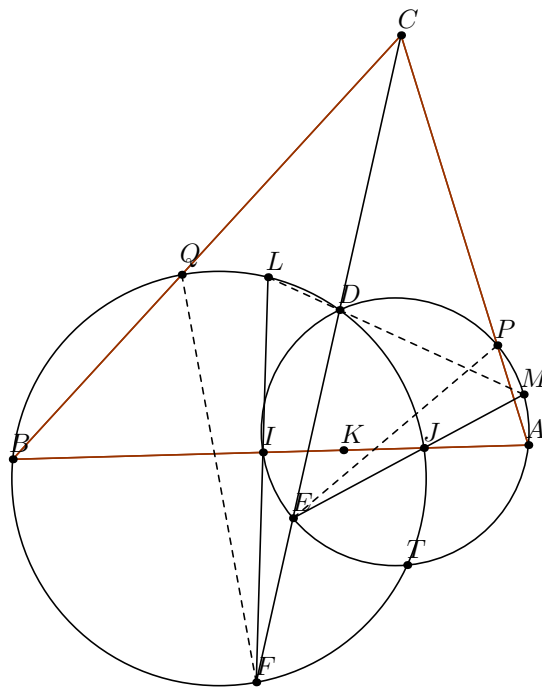
Figure S13 is the receipt from the organizer of 2024 USEMO confirming shortlisting our proposals.



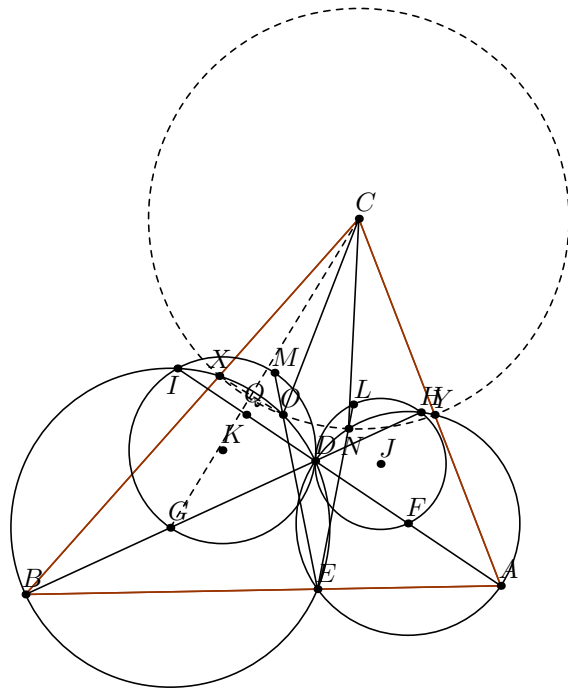




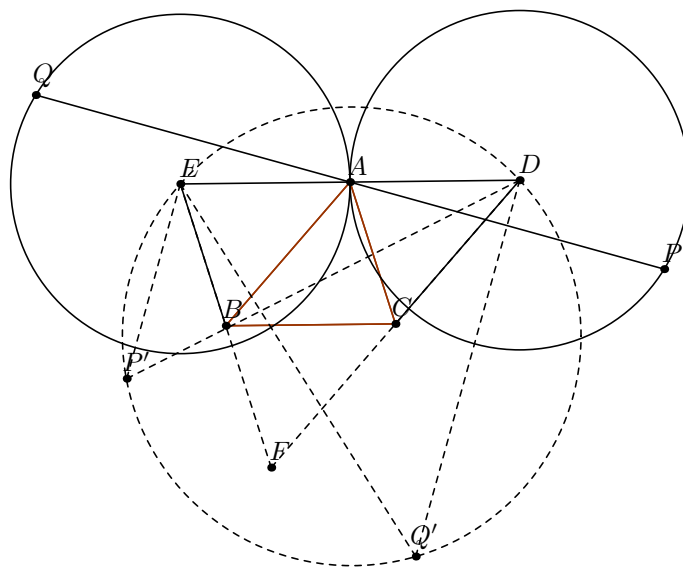
**Figure S3: Problem 2 of TongGeometry’s proposal to 2024 USEMO.**



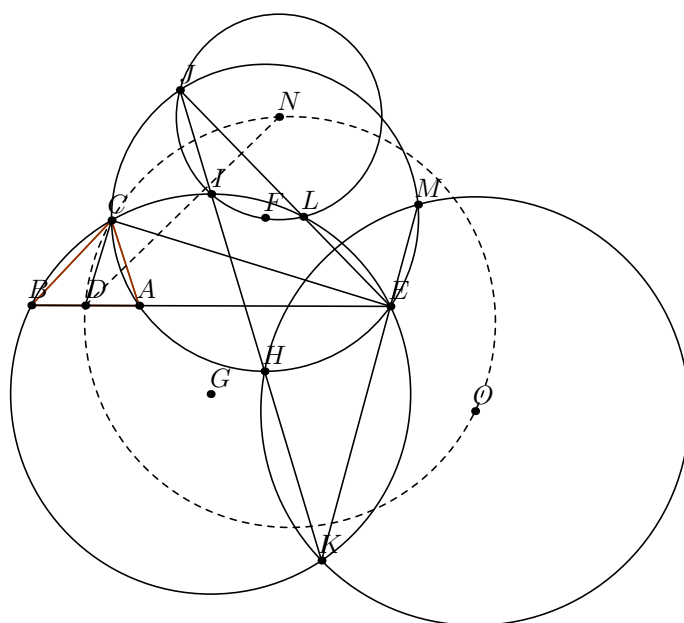
**Figure S4: Problem 3 of TongGeometry’s proposal to 2024 USEMO.**



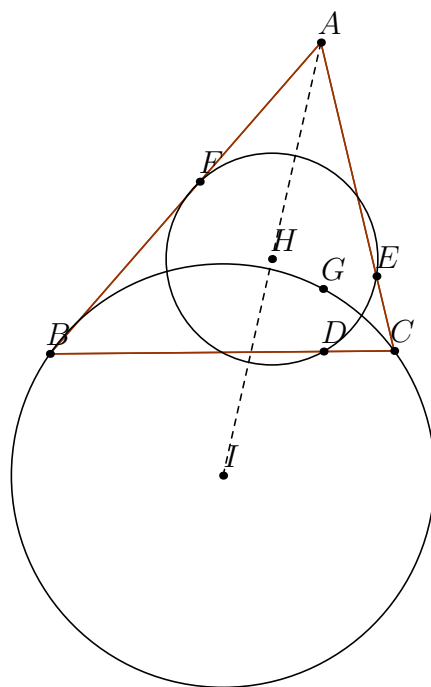
**Figure S5: Problem 4 of TongGeometry's proposal to 2024 USEMO.**



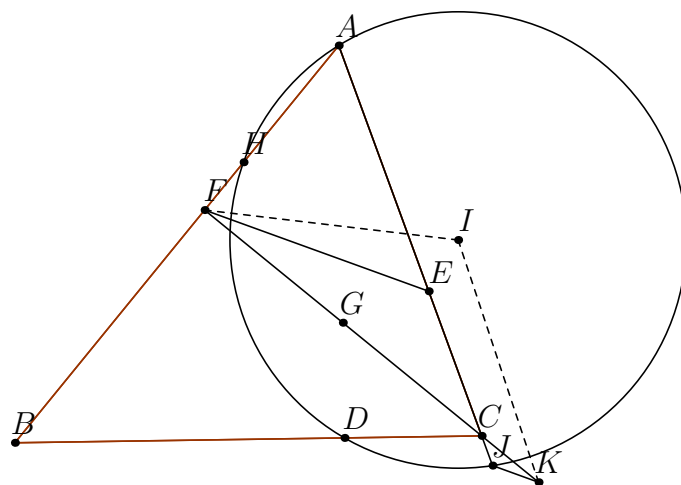
**Figure S6: Problem 5 of TongGeometry's proposal to 2024 USEMO.**



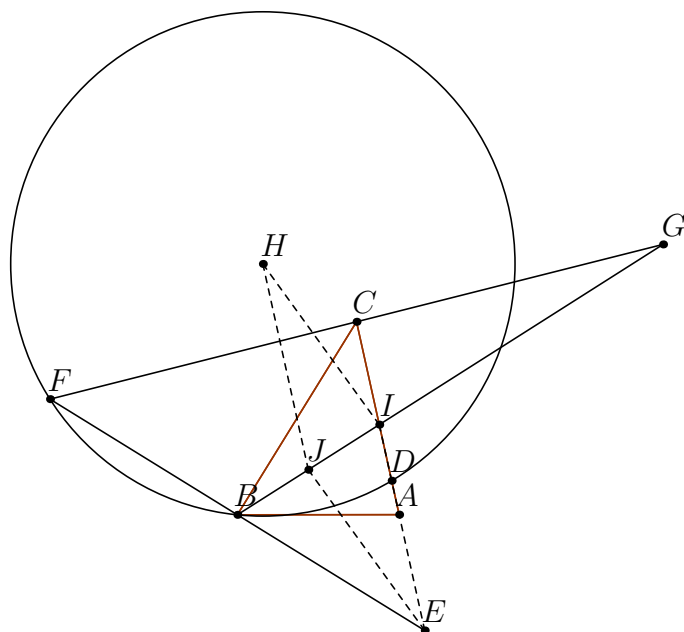
**Figure S7: Problem 6 of TongGeometry's proposal to 2024 USEMO.**



**Figure S8: Problem 1 of TongGeometry's proposal to 2024 NBSML (Beijing).**



**Figure S9: Problem 2 of TongGeometry's proposal to 2024 NHSM (Beijing).**



**Figure S10: Problem 3 of TongGeometry's proposal to 2024 NHSM (Beijing).**



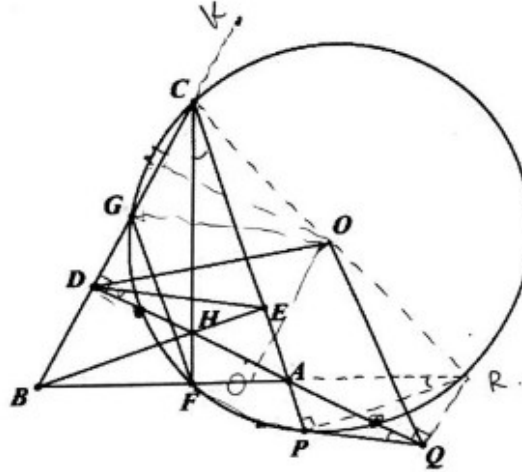
2024 年全国高中联赛北京赛区预赛二试试题

考试时间: 9:40-12:30

一 (40 分) 设  $a, b, c$  是三个正数, 求证:

$$\frac{2a}{\sqrt{2a^2+b^2+c^2}} + \frac{2b}{\sqrt{a^2+2b^2+c^2}} + \frac{2c}{\sqrt{a^2+b^2+2c^2}} \leq \frac{3\sqrt{2}(a+b+c)}{\sqrt{5a^2+5b^2+5c^2+ab+bc+ca}}.$$

二 (40 分) 如图所示, 锐角  $\triangle ABC$  的三条高线  $AD, BE, CF$  交于点  $H$ , 过点  $F$  作  $FG \parallel AC$  交直线  $BC$  于点  $G$ , 设  $\triangle CFG$  的外接圆为  $\odot O$ ,  $\odot O$  与直线  $AC$  的另一个交点为  $P$ , 过  $P$  作  $PQ \parallel DE$  交直线  $AD$  于点  $Q$ , 连接  $OD, OQ$ . 求证:  $OD = OQ$ .



三 (50 分) 有  $n$  个球队参加比赛, 球队之间的比赛计划已经安排好了. 但是每场比赛的主场客场还没有分配好. 这时每个球队都上报了自己能够接受的客场比赛的最大次数. 最终组委会发现这些次数加在一起恰好是比赛的总场次, 并且组委会还发现任意挑出若干支球队, 他们能够接受的客场次数之和都要大于等于他们之间的比赛总场次. 请问组委会能否安排好主客场使得每支球队都满意, 请证明你的结论.

四 (50 分) 设  $a_1, a_2, \dots, a_n$  为  $n$  个两两不同的正整数且  $a_1 a_2 \cdots a_n$  恰有 4048 个质因数.

如果  $a_1, a_2, \dots, a_n$  中任意多个数相乘均不是一个整数的 4049 次方, 求  $n$  的最大值.

**Figure S12: Our proposal in 2024 NHSML (Beijing).** The second problem is our proposal, the only geometry in the competition.



**Evan Chen** evan@evanchen.cc ...



Sat, May 11, 11:07 AM



to me ▼

Hi Chi,

I've workshopped some of the proposals and returned some comments in the file here. Note I edited statements for p4-p6 a lot. I hope those edits can be useful to you.

Probably for USEMO shortlist, I will try p4 + p5 and see how it goes. The USEMO shortlist this year is really geo-heavy (not even including the problems you sent, it was already like half geometry, and mostly hard geometry), so there will be some competition, but you might get lucky. I didn't end up shortlisting p6 since there were a lot of geo in that difficulty tier, and I don't want to overwhelm my small review panel.

You'll hear back about p4 + p5 in about two months time or so. You're free to use remaining four problems for anything, those are returned to you now.

**Figure S13: Receipt of our proposals to 2024 USEMO.** The receipt confirms shortlisting two of our proposals.