Causal-R: A Causal-Reasoning Geometry Problem Solver for Optimized Solution Exploration

Anonymous Author(s)
Affiliation
Address
`email`

A Related works

Automated mathematical problem solving has always been a hot topic in the community 8; 4; 19, where geometry problem solving obtains increasing attention these years due to its complexity, requiring multiple crucial capabilities of machine. Earlier methods 6; 7; 23 such as Wu's Method 21 built fundamental approaches for understanding and proving geometry problems. Recent advancements in neural networks and symbolic systems have spurred increasing researches for breakthroughs in problem solving performances of automated geometry problem solvers, which can be categorized into two mainlines.

Neural-based methods predict executable or human-readable sequences based on the extracted multi-modal features of raw problem. For example, Chen et al. (2021) made the first attempt and established a baseline to uniformly process textual and visual data by encoder-decoder framework and generate program sequences, which is then executed by external symbolic engine. Since then, continuous improvements have been made on the problem types 2, information extraction 25; 12, and cross-modal information alignment 14; 12. With the exploitation of large language models (LLMs), many studies have evaluated and developed the performances of LLMs on mathematic tasks 11; 18, including geometry problem solving 24. Some works focus on improving the multi-modal problem understanding capabilities of multi-modal LLMs 15, especially on geometric diagrams 3; 27; 26. Other works try to enhance the problem reasoning and solving skills of MLLMs by integration of external symbolic engines 16; 5, such as ToRA 9 and MathCoder 20.
However, these methods fall short of guaranteeing the accuracy and interpretability of generated solution paths due to the inherent shortage of autoregressive models, likely introducing redundant and wrong steps.

Symbolic-based methods continuously apply predefined theorem rules on parsed geometric conditions until the problem goal is solved. Inter-GPS 13 proposes to use a pretrained neural model to predict a sequence of potential needed theorems, improving the efficiency based on brute-search strategy. Subsequently, several works aim at enhancing the capabilities of neural models to analyze the pattern and make improved predictions. For example, Peng et al. (2023) and Zou et al. (2024) utilize reinforcement learning to enable the model with step-wise prediction power before each step of theorem application. Following them, Huang et al. (2024) builds global hologram of geometric conditions and pattern holograms of theorems and uses neural models to predict appropriate theorem based on pattern matching at each step. Recent work Pi-GPS 28 mainly focuses on resolving the textual ambiguities with diagrammatic information, significantly improving the applicability of symbolic-based methods. As an outstanding work, AlphaGeometry 19 highlights the potentials of symbolic models on mathematical proof problems even at Olympic-level, encouraging researchers to keep exploring. E-GPS 22 is the first work that emphasizes the importance of explainability of solutions, advancing the practical application of geometry problem solvers.
However, most of these works mainly focus on solving the geometry problems instead of reasoning for optimized solutions, ignoring the importance of reasonable solutions in real application scenarios. Besides E-GPS, they are unable to ensure the interpretability and accuracy of solutions (*e.g.*, without

introducing redundant steps), leaving the exploration of shorter and multiple solutions an ongoing research problem.

B Overall framework

We provide an overall framework of our Causal-R in fig. 1.

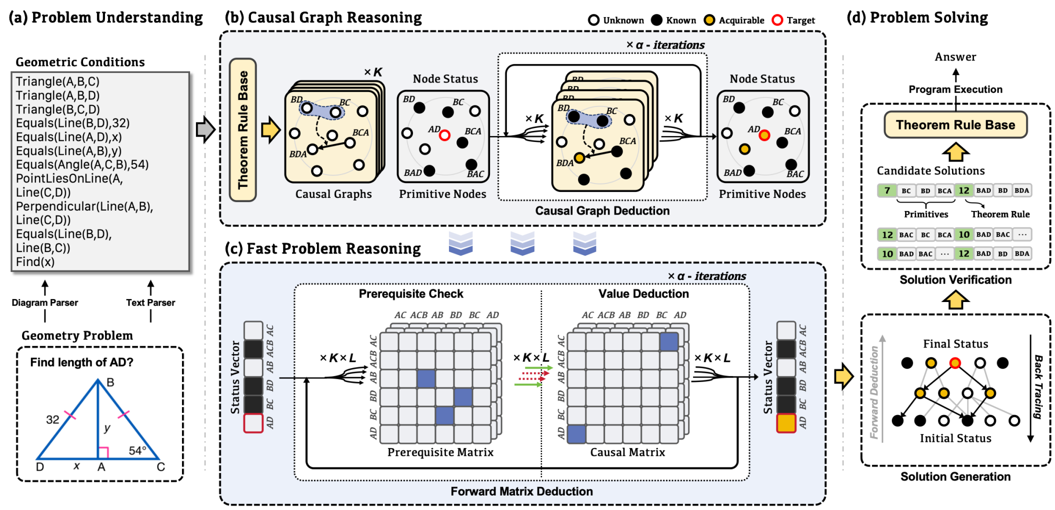

Figure 1: The overall framework of our Causal-R. It mainly contains four parts: (a) in problem understanding stage, the raw geometry problem is parsed into formalized geometric conditions; (b) causal graph reasoning provides the fundamental theory for causal deduction based on causal graphs; (c) in fast problem reasoning stage, the causal graph deduction is implemented by forward matrix deduction to reason for solutions to final problem goal in iterations; (d) the solutions are generated by tracing back from the target node, and are then verified to obtain the final answer.

C Theorem rule simplification

We provide a detailed illustration of how we simplify the theorem rule into unified causal deduction representation forms in table 1 and table 2. Note that we only list partial transformations w.r.t. each theorem rule for clarity and brevity. For example, in *Angle Sum of Triangle* in table 1, we only list one deduction path, while the deduction path remains valid after the alteration of n_a, n_b and n_c. Similar situations exist in *Law of Sines* in table 2, too. These deduction paths are extracted from the equations between these values of nodes, where any one of them can be obtained if the other values are known. More detailed implementation of causal graph deduction can be found in codes.

Table 1: The simplified causal deduction paths of original theorem rule base. For clarity and brevity, only partial deduction paths w.r.t. each rule are presented. Δ is used as placeholder to represent the same content within each rule.

Geometry Theorem/Definition	Simplified Causal Deduction
Circle Definition	For radius OA (n_a), OB (n_b) of circle O: $\tilde{n_a} \leftarrow e_i^c(\tilde{n_b})$
Thales Theorem	Point A, B, C on circle O, for $\angle$ABC (n_b), $\angle$AOC (n_o): $\tilde{n_b} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\cdot)$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\tilde{n_o})$
Inscribed Angle Theorem	Point A, B, C on circle O, for $\angle$ABC (n_b), $\angle$AOC (n_o): $\tilde{n_b} \leftarrow e_i^c(\tilde{n_o})$
Parallel Lines Theorem	AB // CD, P on AB, Q on CD, for $\angle$APQ (n_p) and $\angle$DQP (n_q): $\tilde{n_p} \leftarrow e_i^c(\tilde{n_q})$
Angle Sum of Triangle	For $\angle$ABC (n_b), $\angle$ACB (n_c), $\angle$BAC (n_a) in $\triangle$ABC: $\tilde{n_a} \leftarrow e_i^c(\tilde{n_b}, \tilde{n_c})$
Isosceles Triangle Theorem (Side)	For AB (n_{b1}), AC (n_{c1}), $\angle$ABC (n_{b2}), $\angle$ACB (n_{c2}) in $\triangle$ABC: $\tilde{n_{b1}} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n_{c1}})$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\mathbf{rel}(n_{b2}, n_{c2}))$
Isosceles Triangle Theorem (Angle)	For AB (n_{b1}), AC (n_{c1}), $\angle$ABC (n_{b2}), $\angle$ACB (n_{c2}) in $\triangle$ABC: $\tilde{n_{b2}} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n_{c2}})$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\mathbf{rel}(n_{b1}, n_{c1}))$
Equilateral Triangle Theorem	For AB (n_{c1}), AC (n_{b1}), BC (n_{a1}) $\angle$ABC (n_{b2}), $\angle$ACB (n_{c2}), $\angle$BAC (n_{a2}) in $\triangle$ABC: $\tilde{n_{a2}} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\cdot), \tilde{n_{b2}} \leftarrow \mathbf{g}(\Delta) \cdot e_{i+1}^c(\cdot), \tilde{n_{c2}} \leftarrow \mathbf{g}(\Delta) \cdot e_{i+2}^c(\cdot)$ $\mathbf{g}(\Delta) \leftarrow e_j^p(\mathbf{rel}(n_{a1}, n_{b1}), \mathbf{rel}(n_{a1}, n_{c1}))$
Triangle's Center of Gravity	If Q on AB, M on BC, N on AC, CQ,AM,BN intersects at P, for AQ (n_{c1}), BQ (n_{c2}), BM (n_{a1}), CM (n_{a2}), AN (n_{b1}), CN (n_{b2}), CP (n_{q1}), PQ (n_{q2}), AP (n_{m1}), PM (n_{m2}), BP (n_{n1}), PN (n_{n2}): $\tilde{n_{q2}} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n_{q1}}), \tilde{n_{m2}} \leftarrow \mathbf{g}(\Delta) \cdot e_{i+1}^c(\tilde{n_{m1}})$, $\tilde{n_{n2}} \leftarrow \mathbf{g}(\Delta) \cdot e_{i+2}^c(\tilde{n_{n1}})$ $\mathbf{g}(\Delta) \leftarrow e_j^p(\mathbf{rel}(n_{a1}, n_{a2}), \mathbf{rel}(n_{b1}, n_{b2}), \mathbf{rel}(n_{c1}, n_{c2}))$
Congruent Triangle Theorem (Proving)	For AB (n_c), AC (n_b), $\angle$BAC (n_{a1}), BC (n_{a2}) in $\triangle$ABC and DE (n_f), DF (n_e), $\angle$EDF (n_{d1}), EF (n_{d2}) in $\triangle$DEF (*SAS*): $\tilde{n_{d2}} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n_{a2}})$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\mathbf{rel}(n_c, n_f), \mathbf{rel}(n_b, n_e), \mathbf{rel}(n_{a1}, n_{d1}))$
Congruent Triangle Theorem	For AB (n_{c1}), AC (n_{b1}), BC (n_{a1}), $\angle$ACB (n_{c2}), $\angle$ABC (n_{b2}), $\angle$BAC (n_{a2}) in $\triangle$ABC and DE (n_{f1}), DF (n_{e1}), EF (n_{d1}), $\angle$DFE (n_{f2}), $\angle$DEF (n_{e2}), $\angle$EDF (n_{d2}) in $\triangle$DEF: $\tilde{n_{f2}} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n_{c2}}), \tilde{n_{e2}} \leftarrow \mathbf{g}(\Delta) \cdot e_{i+1}^c(\tilde{n_{b2}})$, $\tilde{n_{d2}} \leftarrow \mathbf{g}(\Delta) \cdot e_{i+2}^c(\tilde{n_{a2}})$ $\mathbf{g}(\Delta) \leftarrow e_j^p(\mathbf{rel}(n_{c1}, n_{f1}), \mathbf{rel}(n_{b1}, n_{e1}), \mathbf{rel}(n_{a1}, n_{d1}))$
Tangent Secant Theorem	AB and circle O tangent at point B, AM intersects circle O at N, for AB (n_b), AN (n_n) and AM (n_m): $\tilde{n_b} \leftarrow e_i^c(\tilde{n_m}, \tilde{n_n})$
Chord Theorem	A,B,C,D on circle O, AB and CD intersects at point M, for AM (n_a), BM (n_b), CM (n_c), DM (n_d): $\tilde{n_a} \leftarrow e_i^c(\tilde{n_b}, \tilde{n_c}, \tilde{n_d})$
Angle Bisector Theorem	In $\triangle$ABC, M on BC, for AB (n_{b1}), BM (n_{b2}), AC (n_{c1}), CM (n_{c2}), $\angle$BAM (n_{a1}), $\angle$CAM (n_{a2}): $\tilde{n_{b1}} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n_{b2}}, \tilde{n_{c1}}, \tilde{n_{c2}})$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\mathbf{rel}(n_{a1}, n_{a2}))$

Table 2: The rest simplified causal deduction paths of original theorem rule base, following table 1. For clarity and brevity, only partial deduction paths w.r.t. each rule are presented. Δ is used as placeholder to represent the same content within each rule.

Geometry Theorem/Definition	Simplified Causal Deduction
Pythagoras Theorem	For AB (n_c), AC (n_b), BC (n_a) in $\triangle$ABC, $\angle$ACB=90°: $\tilde{n}_c \leftarrow e_i^c(\tilde{n}_a, \tilde{n}_b)$
Law of Sines	For AB (n_{c1}), $\angle$ACB (n_{c2}), AC (n_{b1}), $\angle$ABC (n_{b2}) in $\triangle$ABC: $\tilde{n}_{c1} \leftarrow e_i^c(\tilde{n}_{c2}, \tilde{n}_{b1}, \tilde{n}_{b2})$
Law of Cosines	For AB (n_{c1}), AC (n_b), BC (n_a), $\angle$ACB (n_{c2}) in $\triangle$ABC: $\tilde{n}_{c1} \leftarrow e_i^c(\tilde{n}_a, \tilde{n}_b, \tilde{n}_{c2})$
Similar Triangle Theorem (Proving)	For $\angle$ABC (n_b), $\angle$ACB (n_c), $\angle$BAC (n_a) in $\triangle$ABC and $\angle$DEF (n_e), $\angle$DFE (n_f), $\angle$EDF (n_d) in $\triangle$DEF: $\tilde{n}_a \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n}_d)$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\mathbf{rel}(n_b, n_e), \mathbf{rel}(n_c, n_f))$
Similar Triangle Theorem	For $\angle$ABC (n_{b1}), $\angle$ACB (n_{c1}), $\angle$BAC (n_a), AB (n_{c2}), AC (n_{b2}) in $\triangle$ABC and $\angle$DEF (n_{e1}), $\angle$DFE (n_{f1}), $\angle$EDF (n_d), DE (n_{f2}), DF (n_{e2}) in $\triangle$DEF: $\tilde{n}_{c2} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n}_{b2})$ $\mathbf{g}(\Delta) \leftarrow$ $e_i^p(\mathbf{rel}(n_{b1}, n_{e1}), \mathbf{rel}(n_{c1}, n_{f1}), \mathbf{rel}(n_a, n_d), \mathbf{rel}(n_{f2}, n_{e2}))$
Similar Polygon Theorem	For AB (n_a), BC (n_b), EF (n_e), FG (n_f) in similar polygons ABCD and EFGH: $\tilde{n}_a \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n}_e)$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\mathbf{rel}(n_b, n_f))$
Median Line Theorem	Point M on AB and point N on AC, for AM (n_{m1}), BM (n_{m2}), AN (n_{n1}), CN (n_{n2}), MN (n_{a1}), BC (n_{a2}) in $\triangle$ABC: $\tilde{n}_{a1} \leftarrow \mathbf{g}(\Delta) \cdot e_i^c(\tilde{n}_{a2})$ $\mathbf{g}(\Delta) \leftarrow e_i^p(\mathbf{rel}(n_{m1}, n_{m2}), \mathbf{rel}(n_{n1}, n_{n2}))$
Area Equation Theorem	In $\triangle$ABC, point M on BC, AM$\perp$BC, for BC (n_a), AM (n_m), Area_$\triangle$ABC (n_s): $\tilde{n}_s \leftarrow e_i^c(\tilde{n}_a, \tilde{n}_m)$
Angle Sum of Polygon	For $\angle$ABC (n_b), $\angle$BCD (n_c), $\angle$CDA (n_d), $\angle$DAB (n_a) in Quadrilateral ABCD: $\tilde{n}_a \leftarrow e_i^c(\tilde{n}_b, \tilde{n}_c, \tilde{n}_d)$

D Case analysis

In order to provide more detailed qualitative analysis, we conduct some cases from Causal-R and present them in fig. 2. In case (a), the model is able to generate two solutions, where the order of theorem application is consistent but the detailed involved primitives are different. Both solutions are reasonable and feasible and lead to correct answer. In case (b), two different solutions with inverted order of theorem applications are obtained. The solutions are simple even if the geometric conditions of the problem seems complex. On one hand, it shows the superiority of Causal-R to explore shorter solutions. On the other hand, it also benefits from the design of symbolic-based strategy which can precisely match the needed conditions within a large set of complicated and redundant geometric conditions. Case (c) is a successful case that demonstrates the applicability of our method in situations that involve proving theorems. However, as we analyze this geometry problem, we find that it can be solved simply by using *Geometric Mean Theorem*, which is not yet contained in the existing theorem rule base. Case (d) is a failure case that our model fails to obtain the solutions and answer, which mainly attributes to the incorrect parsing of the geometric problem and incomplete design of symbolic system. It remains a difficult task to parse the content with irregular representations such as shaded area. These two cases both indicate the necessity of further developing the symbolic system for adapting to more question types and more theorems.

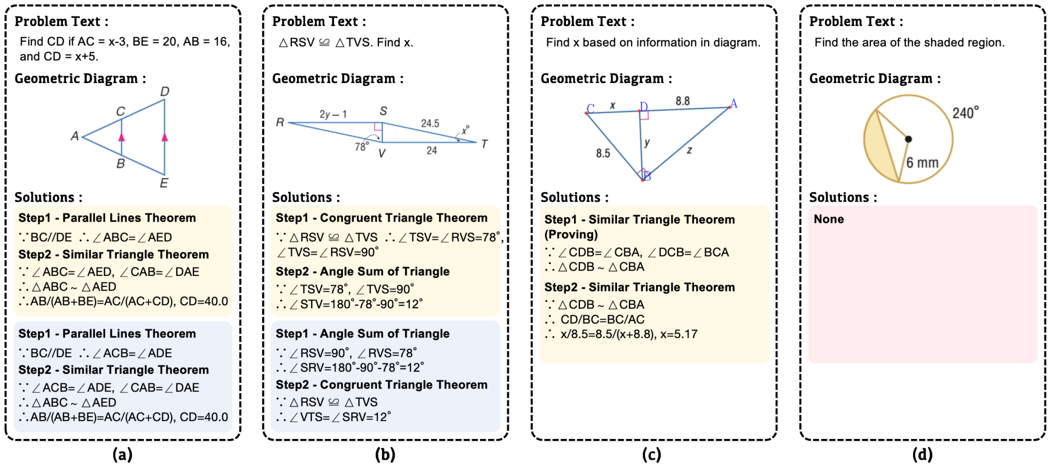

Figure 2: Four typical cases from Causal-R. (a), (b) and (c) are successful cases and (d) is a failure case.

E Limitations

While certain advantages and performance improvements have been obtained by our Causal-R, we also identify some limitations, along with potential solutions, from the following aspects: (1) The effectiveness of causal graph reasoning is influenced by the design of theorem rule base $\mathcal{KB}$. It not only limits the upper bound of problem solving performances (*i.e.*, the geometry problem can not be solved if the needed theorem rule is not contained in $\mathcal{KB}$), but also affects the comprehensive outcome of transformation from theorem rules to causal graphs. This can be resolved by future development and refinement of $\mathcal{KB}$, providing a more standard and larger theorem rule base. Note that once the $\mathcal{KB}$ is determined, the deduction paths in causal graph construction are also determined (such as presented in table 1 and table 2) and do not change for different geometry problems. (2) For more intricate geometry problems, when the combinations of geometric primitives and theorem rules become extreme large, the temporary storage requirement for matrix-based deduction also increases. It can be optimized through both refinement of $\mathcal{KB}$ and application of more efficient matrix operation methodologies. (3) Currently, the method does not support theorem rules that involve constructing geometric primitives (*e.g.*, connect point `A` and point `B` as a new line primitive `AB`). One possible solution is to apply such rules before the causal graph construction to augment the base geometric primitive nodes for a wider solution space that includes the action of constructing new primitives.

References

Jiaqi Chen, Jianheng Tang, Jinghui Qin, Xiaodan Liang, Lingbo Liu, Eric Xing, and Liang Lin. Geoqa: A geometric question answering benchmark towards multimodal numerical reasoning. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 513–523, 2021.

Jiaqi Chen, Tong Li, Jinghui Qin, Pan Lu, Liang Lin, Chongyu Chen, and Xiaodan Liang. Unigeo: Unifying geometry logical reasoning via reformulating mathematical expression. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 3313–3323, 2022.

Seunghyuk Cho, Zhenyue Qin, Yang Liu, Youngbin Choi, Seungbeom Lee, and Dongwoo Kim. Geodano: Geometric vlm with domain agnostic vision encoder. *arXiv preprint arXiv:2502.11360*, 2025.

Shang-Ching Chou, Xiao-Shan Gao, and Jing-Zhong Zhang. Automated generation of readable proofs with geometric invariants: I. multiple and shortest proof generation. *Journal of Automated Reasoning*, 17(3): 325–347, 1996.

Daocheng Fu, Zijun Chen, Renqiu Xia, Qi Liu, Yuan Feng, Hongbin Zhou, Renrui Zhang, Shiyang Feng, Peng Gao, Junchi Yan, et al. Trustgeogen: Scalable and formal-verified data engine for trustworthy multi-modal geometric problem solving. *arXiv preprint arXiv:2504.15780*, 2025.

Wenbin Gan and Xinguo Yu. Automatic understanding and formalization of natural language geometry problems using syntax-semantics models. *International Journal of Innovative Computing, Information and Control*, 14 (1):83–98, 2018.

Wenbin Gan, Xinguo Yu, Ting Zhang, and Mingshu Wang. Automatically proving plane geometry theorems stated by text and diagram. *International Journal of Pattern Recognition and Artificial Intelligence*, 33(07): 1940003, 2019.

Herbert Gelernter, James R Hansen, and Donald W Loveland. Empirical explorations of the geometry theorem machine. In *Papers presented at the May 3-5, 1960, western joint IRE-AIEE-ACM computer conference*, pages 143–149, 1960.

Zhibin Gou, Zhihong Shao, Yeyun Gong, Yujiu Yang, Minlie Huang, Nan Duan, Weizhu Chen, et al. Tora: A tool-integrated reasoning agent for mathematical problem solving. In *The Twelfth International Conference on Learning Representations*.

Litian Huang, Xinguo Yu, Feng Xiong, Bin He, Shengbing Tang, and Jiawen Fu. Hologram reasoning for solving algebra problems with geometry diagrams. *arXiv preprint arXiv:2408.10592*, 2024.

Xiaoyuan Li, Wenjie Wang, Moxin Li, Junrong Guo, Yang Zhang, and Fuli Feng. Evaluating mathematical reasoning of large language models: A focus on error identification and correction. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 11316–11360, 2024a.

Zhong-Zhi Li, Ming-Liang Zhang, Fei Yin, and Cheng-Lin Liu. Lans: A layout-aware neural solver for plane geometry problem. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 2596–2608, 2024b.

Pan Lu, Ran Gong, Shibiao Jiang, Liang Qiu, Siyuan Huang, Xiaodan Liang, and Song-chun Zhu. Inter-gps: Interpretable geometry problem solving with formal language and symbolic reasoning. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 6774–6786, 2021.

Maizhen Ning, Qiu-Feng Wang, Kaizhu Huang, and Xiaowei Huang. A symbolic characters aware model for solving geometry problems. In *Proceedings of the 31st ACM International Conference on Multimedia*, pages 7767–7775, 2023.

Maizhen Ning, Zihao Zhou, Qiufeng Wang, Xiaowei Huang, and Kaizhu Huang. Gns: Solving plane geometry problems by neural-symbolic reasoning with multi-modal llms. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 24957–24965, 2025.

Yicheng Pan, Zhenrong Zhang, Pengfei Hu, Jiefeng Ma, Jun Du, Jianshu Zhang, Quan Liu, Jianqing Gao, and Feng Ma. Enhancing the geometric problem-solving ability of multimodal llms via symbolic-neural integration. *arXiv preprint arXiv:2504.12773*, 2025.

Shuai Peng, Di Fu, Yijun Liang, Liangcai Gao, and Zhi Tang. Geodrl: A self-learning framework for geometry problem solving using reinforcement learning in deductive reasoning. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 13468–13480, 2023.

Ankit Satpute, Noah Gießing, André Greiner-Petter, Moritz Schubotz, Olaf Teschke, Akiko Aizawa, and Bela Gipp. Can llms master math? investigating large language models on math stack exchange. In *Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval*, pages 2316–2320, 2024.

Trieu H Trinh, Yuhuai Wu, Quoc V Le, He He, and Thang Luong. Solving olympiad geometry without human demonstrations. *Nature*, 625(7995):476–482, 2024.

Ke Wang, Houxing Ren, Aojun Zhou, Zimu Lu, Sichun Luo, Weikang Shi, Renrui Zhang, Linqi Song, Mingjie Zhan, and Hongsheng Li. Mathcoder: Seamless code integration in llms for enhanced mathematical reasoning. In *ICLR*, 2024.

Wu Wen-Tsun. Basic principles of mechanical theorem proving in elementary geometries. *Journal of automated Reasoning*, 2:221–252, 1986.

Wenjun Wu, Lingling Zhang, Jun Liu, Xi Tang, Yaxian Wang, Shaowei Wang, and Qianying Wang. E-gps: Explainable geometry problem solving via top-down solver and bottom-up generator. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13828–13837, 2024.

Xinguo Yu, Mingshu Wang, Wenbin Gan, Bin He, and Nan Ye. A framework for solving explicit arithmetic word problems and proving plane geometry theorems. *International Journal of Pattern Recognition and Artificial Intelligence*, 33(07):1940005, 2019.

Jiaxin Zhang, Zhong-Zhi Li, Ming-Liang Zhang, Fei Yin, Cheng-Lin Liu, and Yashar Moshfeghi. Geoeval: Benchmark for evaluating llms and multi-modal models on geometry problem-solving. In *Findings of the Association for Computational Linguistics ACL 2024*, pages 1258–1276, 2024a.

Ming-Liang Zhang, Fei Yin, and Cheng-Lin Liu. A multi-modal neural geometric solver with textual clauses parsed from diagram. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 3374–3382, 2023.

Renrui Zhang, Dongzhi Jiang, Yichi Zhang, Haokun Lin, Ziyu Guo, Pengshuo Qiu, Aojun Zhou, Pan Lu, Kai-Wei Chang, Yu Qiao, et al. Mathverse: Does your multi-modal llm truly see the diagrams in visual math problems? In *European Conference on Computer Vision*, pages 169–186. Springer, 2024b.

Renrui Zhang, Xinyu Wei, Dongzhi Jiang, Ziyu Guo, Shicheng Li, Yichi Zhang, Chengzhuo Tong, Jiaming Liu, Aojun Zhou, Bin Wei, et al. Mavis: Mathematical visual instruction tuning with an automatic data engine. *arXiv preprint arXiv:2407.08739*, 2024c.

Junbo Zhao, Ting Zhang, Jiayu Sun, Mi Tian, and Hua Huang. Pi-gps: Enhancing geometry problem solving by unleashing the power of diagrammatic information. *arXiv preprint arXiv:2503.05543*, 2025.

Jia Zou, Xiaokai Zhang, Yiming He, Na Zhu, and Tuo Leng. Fgeo-drl: Deductive reasoning for geometric problems through deep reinforcement learning. *Symmetry*, 16(4):437, 2024.