A New Set of Problems for the AlphaGeometry Engine

Vladmir Sicca[1], Rasul Tutunov[2], Shangling Jui[1], Jun Wang[2,3], and Haitham Bou Ammar[2,3]

[1]Huawei Lagrange Center for Mathematics and Computing
[2]Huawei London Research Center
[3]University College London

May 13, 2024

Abstract

In this paper we report our improvement to AlphaGeometry's language to allow the insertion of constant angles and ratios, as well as of angle and ratio equivalences, as hypothesis and proof goals of problems to an extent not covered in the original work. By doing that, we extend the capabilities of the software to solve new problems in school level, prove classic theorems, and even solve some mathematical olympiad problems from national competitions.

In January of 2024 the paper [1] brought to light AlphaGeometry (AG), an efficient reasoning engine capable of executing a wide range of logical deductions through a set of axioms and theorems from Euclidean plane geometry. The software receives a set of hypothesis given in an appropriate language, aiming at an also formally-given goal. Such deductive system is composed by what are called in the original paper a deductive derivation and an algebraic reasoning modules (DDAR). The complete system is also enhanced by a language model trained to offer new constructions within the well-established language in order to advance the derivation engine in case it gets stuck. The original engine achieved good results on a selected set of problems from the International Mathematics Olympiad (IMO), a known competition of high-school level mathematics which problems are oriented towards finding proofs, not executing calculations. The engine created is specialized in those kinds of problems.

As it would be expected, such a system has limitations, which can be represented by problems it is unable to solve. A limitation that is harder to detect is that of problems that can be fed to the engine, but it simply happens to not be able to find a solution within given time/search constraints. But there is a class of problems that the original AlphaGeometry language is not equipped to even represent, specifically some hypothesis and goals cannot be written within the original language. In our work we extend the set of definitions and of predicates to allow the engine to admit problems that it already can process, but for which the language was lacking before.

Specifically, the original code could not represent hypothesis of the following form:

- "Given points A, B, C, D, E, F, and G, let H be a point such that $\frac{AB}{CD} = \frac{EF}{GH}$."
- "Assume segments AB and CD are in a proportion of $3 : 2$ with each other."

It also had limitations on representing the hypothesis that two angles are the same in some cases. We added new constructions that allow for the hypothesis $\angle ABC = \angle CDE$ to be implemented regardless of the situation of previously defined points.

It also could not solve problems that asked to prove that an angle or a ratio had a given value, as in

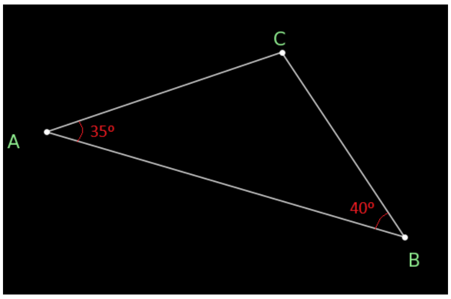

Figure 1: Red notation added to AG-generated picture.

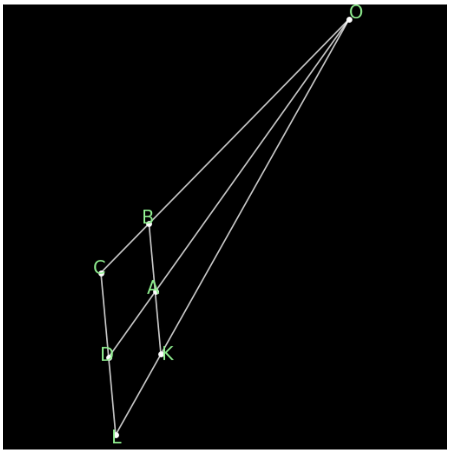

Figure 2: AG-generated picture.

- "Prove that $\angle ABC = 30°$."
- "Prove that $\frac{AB}{CD} = \frac{3}{2}$."

Interestingly enough, most of the necessary wiring was already present within the code, so by adding new constructions, without changing the reasoning of the DDAR deduction engine, a new collection of problems can be solved. Such problems are indeed rare at IMO exams, but they are easily recognizable by geometry students from a very early level.

A very standard example is that of finding the third angle of a triangle given the other two. Consider the following problem:

"In triangle ABC, we have angle $A = 35°$ and angle $B = 40°$. Prove that angle $C = 105°$." (See Figure 1).

As simple as it is, and even it being a paradigmatic problem for AG's angle chasing capabilities, this simple problem could not be previously fed into the original software, but our improvements allow it now.

Our additions also allow for prescribing an equality of ratios. For example, we can now prove the following version of Thales theorem:

"If AB and CD are parallel segments, K is a point in the line AB, and L is a point in the line CD such that $\frac{DL}{CL} = \frac{AK}{BK}$, prove that lines AD, BC, and KL are concurrent lines." (See Figure 2).

That problem also could not be proposed before, but is solved with our improvements.

Finally, although such problems are not usually in the style of IMO exams, our changes do allow the

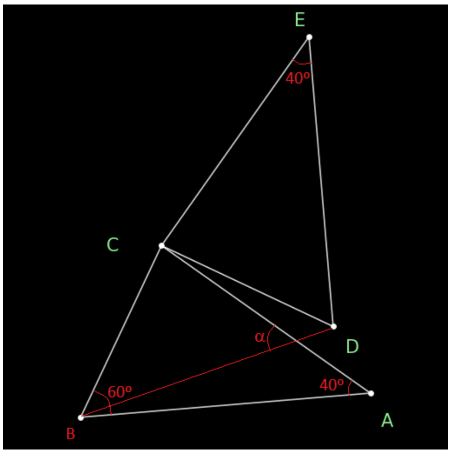

Figure 3: Red notation added to AG-generated picture.

treatment of problems in other olympiads that are beyond reach of the original AG. A good example are problems of the 1st phase of the Brazilian Mathematical Olympiad (OBM). For example, we can now solve Problem 2 of the 2001 exam. Of course, some reformulation of the question is needed to fit the existing definitions, and the original question of "calculate angle α" has to be turned into "prove $\alpha = 55°$". But we do not consider they change the nature of the problem, specially because the engine has to specifically calculate the angle in order to prove the claim, and could detect the mistake in case we tried to prove α was equal to any other value. Our formulation of the problem is the following:

"Consider the segment AB. Let C be such that angle $\angle BAC = 40°$ and angle $\angle ABC = 60°$. Let E be a point such that $\angle CED = 40°$ and $DC \perp CB$. Prove that the smallest angle between lines BD and AC is 55°." (See Figure 3).

Finally, the new functionalities allow us also to have a better picture of the reasoning limitations of the engine when it comes to processing angles and ratios, in particular of the algebraic reasoning module. For example, the reasoning engine can now face the question of proving that the internal angle of a regular hexagon is 120°, but fails to find a solution, despite it being approachable through a linear system of equations on the angles. The formulation of the question we used is equivalent to the following:

"In the circle of center O going through point A, consider the points B, C, D, E, and F such that $OA = AB = BC = CD = DE = EF$. Prove that the smallest angle between the lines AF and AB measures 120°."

The possibility of addressing this kind of questions, unavailable in the original version of AlphaGeometry, reveals limitations in its engine, which can inform future ameliorations to turn the machine into a broader, more powerful, geometric solver.

References

[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