# Towards automatic detection of geometric difficulty of geometric problems

Zoltán Kovács<sup>1</sup> András Kerekes<sup>2</sup>

<sup>1</sup>PHDL, Linz, Austria + RISC, Hagenberg, Austria

<sup>2</sup>Radnóti Miklós High School, Szeged, Hungary





ADG Conference, 1 August 2025 CADE Stuttgart, Germany



#### **Abstract**

We compare three different options to define geometric difficulty  $C_g$  of geometry problems.

By computing correlation between the achieved values and the algebraic complexity  $C_a$ , and, alternatively, ad-hoc human intuition  $C_h$ , we confirm that our former definition seems acceptable.

#### Why is this interesting?

- ► Ranking geometry theorems → Wos' Research Problem 31 (1988): What properties can be identified to permit an automated reasoning program to find new and interesting theorems, as opposed to proving conjectured theorems? Also studied e.g. by Colton, Bundy and Walsh; Puzis, Gao and Sutcliffe; Gao, Goto and Cheng; Gao, Li and Cheng; and Quaresma, Graziani and Nicoletti.
- Abar and de Almeida: research among university students to learn their personal opinions on geometric difficulty, classifying complexity before and after attempting to actually prove the statements.
- ► Algebraic complexity may be different from human perception and geometric complexities → a good measure that works in general? Having such a measure would help to classify problem setting in classroom situations, exams, or mathematics contests.
- ▶ Generating new geometry problems → automated estimation of complexity.



# Possible definitions of complexity

Algebraic, Z.K., T. Recio and M. P. Vélez, 2023

Let C be a geometric construction with hypotheses  $H_1, H_2, \ldots, H_r$  and thesis T. Prove the statement

$$H_1 \wedge H_2 \wedge \ldots \wedge H_r \Rightarrow T$$

by considering an algebraic proof

$$1 = f_1 h_1 + f_2 h_2 + \ldots + f_r h_r + f_{r+1} \cdot (zt - 1)$$

of the statement, where  $h_1, h_2, \ldots, h_r$  are algebraic translations of hypotheses  $H_1, H_2, \ldots, H_r$ , and t is an algebraic translation of thesis T. Then, the *algebraic complexity*  $C_a$  of the statement is the maximal degree of  $f_r$ , that is

$$C_a := \max_{i=1}^{r+1} \deg f_i.$$

# Possible definitions of complexity

Geometric, S. Bartha, A.K., Z.K. and T. Recio, 2024

Let C be a geometric construction with hypotheses  $H_1, H_2, \ldots, H_r$  and thesis T. Prove the statement

$$H_1 \wedge H_2 \wedge \ldots \wedge H_r \Rightarrow T$$

by using the Geometric Deductive Database method (GDD) in Java Geometry Expert and consider the number of steps of the proofs, exported as a directed acyclic graph in GraphViz format. Then, the *geometric complexity*  $C_g$  of the statement is the number of steps (technically, the *number of lines*) in the output.

# Possible definitions of complexity

Geometric (3 variations), A.K., Z.K., present contribution

Let C be a geometric construction with hypotheses  $H_1, H_2, \ldots, H_r$  and thesis T. Prove the statement

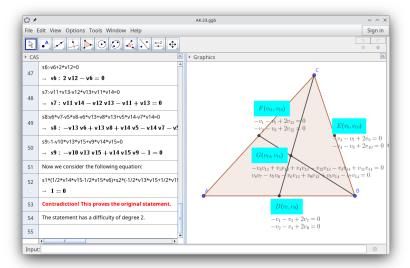
$$H_1 \wedge H_2 \wedge \ldots \wedge H_r \Rightarrow T$$

by using the Geometric Deductive Database method (GDD) in Java Geometry Expert and consider the following:

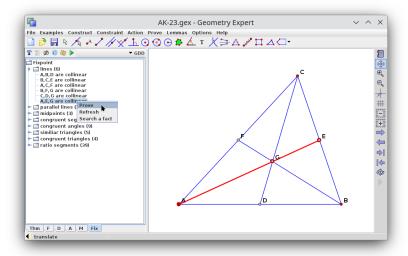
- longest directed path ("Ldp"),
- number of vertices ("#v"),
- number of edges ("#e")

of the obtained directed acyclic graph.

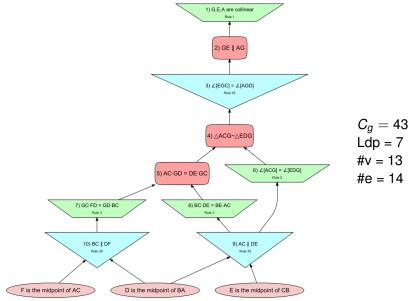
Medians of a triangle intersect each other in one point, algebraic complexity (GD),  $C_a = 2$ 



Medians of a triangle intersect each other in one point (JGEX)



Medians of a triangle intersect each other in one point, geometric complexities



Computing  $C_g$  by counting the lines of the GraphViz output for the GDD proof

```
AK-23.av
node [shape = box, color = black, style = filled];
edge [dir = back, tooltip = " "]:
 [ label = <1) G.F.A are collinear<BR/>
 FONT POINT-SIZE="10">Rule 1</FONT>>, tooltip = "If AB | BC, then A, B and C are collinear,", style = "filled", shape = in
 [ label = <2) GE | AG>, tooltip =
                                      ". style = "rounded, filled". shape = box, fillcolor = "#FFA0A0", URL="https://githu
FONT POINT-SIZE="10">Rule 33</FONT>>. tooltip = "SSS (side-side-side) similarity.\nIf two triangles have all three pairs
 [ label = <4) AACG~AEDG>, tooltip = " ", style = "rounded, filled", shape = box, fillcolor = "#FFA0A0", URL="https://gi
 [ label = <5) AC-GD = DE-GC>, tooltip = " ", style = "rounded, filled", shape = box, fillcolor = "#FFA9A0", URL="https:
 FONT POINT-SIZE="10">Rule 3</FONT>>. tooltip = "If AB | CD and E is the intersection of AC and BD, then EA / EC = EB / EC
.0 [ label = <10) BC || DF<BR/>
FONT POINT-SIZE="10">Rule 35</FONT>>. tooltip = "Middle Connection Theorem for Triangles.\nThe line connects the midpoint
                          [ fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                           fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                           fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                          fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                           fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                          fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
```

Computing  $C_g$  by counting the lines of the GraphViz output for the GDD proof

```
[----] 1 L:[ 1+42 43/ 44] *(4312/4313b) 0010 0x00A
node [shape = box, color = black, style = filled];
edge [dir = back, tooltip = " "]:
  -> 10:
 0 -> "D is the midpoint of BA":
5 -> 8:
 -> 6:
 [ label = <1) G.E.A are collinear<BR/>
      POINT-SIZE="10">Rule 1</FONT>>, tooltip = "If AB | BC, then A, B and C are collinear,", style = "filled", shape = in
                                          ', style = "rounded, filled", shape = box, fillcolor = "#FFA0A0", URL="https://githu
 FONT POINT-SIZE="10">Rule 33</FONT>>. tooltip = "SSS (side-side-side) similarity.\nIf two triangles have all three pairs
  [ label = <4) AACG~AEDG>, tooltip = " ", style = "rounded, filled", shape = box, fillcolor = "#FFA0A0", URL="https://git
   label = <5) AC-GD = DE-GC>, tooltip = " ", style = "rounded, filled", shape = box, fillcolor = "#FFA0A0". URL="https://
 FONT POINT-SIZE="10">Rule 2</FONT>>, tooltip = "For two angles \angle \lceil \lceil 1, \rceil \rceil, \angle \lceil \lceil 3, \rceil \rceil. If \lceil 1 \rceil = \lceil 3 \rceil and \lceil 2 \rceil \rceil \lceil 4 \rceil, then \angle \lceil \lceil 4, \rceil \rceil.
 FONT POINT-SIZE="10">Rule 3</FONT>>. tooltip = "If AB | CD and E is the intersection of AC and BD, then EA / EC = EB / ED
 FONT POINT-SIZE="10">Rule 3</FONT>>, tooltip = "If AB | CD and E is the intersection of AC and BD, then EA / EC = EB / ED
 FONT POINT-SIZE="10">Rule 35</FONT>>, tooltip = "Middle Connection Theorem for Triangles.\nThe line connects the midpoint
 FONT POINT-SIZE="10">Rule 35</FONT>>. tooltip = "Middle Connection Theorem for Triangles.\nThe line connects the midpoint
                              fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                              fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                              fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                              fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                              fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                             fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
```

Computing  $C_g$  by counting the lines of the GraphViz output for the GDD proof

```
node [shape = box, color = black, style = filled];
edge [dir = back, tooltip = " "]:
1 -> 2:
 -> 10:
10 -> "F is the midpoint of AC":
10 -> "D is the midpoint of BA";
5 -> 8:
8 -> 9:
9 -> "D is the midpoint of BA":
4 -> 6:
1 [ label = <1) G.E.A are collinear<BR/>
<FONT POINT-SIZE="10">Rule 1</FONT>>, tooltip = "If AB | BC, then A, B and C are collinear,", style = "filled", shape = in
2 [ label = <2) GE | AG>, tooltip = " ", style = "rounded, filled", shape = box, fillcolor = "#FFA0A0", URL="https://githul
3 [ label = <3) \( \( \text{EGC} \) = \( \text{AGD} \) < BR/>
<FONT POINT-SIZE="10">Rule 33
FONT>>. tooltip = "SSS (side-side-side) similarity.\nIf two triangles have all three pairs
4 [ label = <4) AAGG~AEDG>, tooltip = " ", style = "rounded, filled", shape = box, fillcolor = "#FFA0A0", URL="https://git
   label = <5) AC-GD = DE-GC>, tooltip = " ", style = "rounded, filled", shape = box, fillcolor = "#FFA0A0". URL="https:/
5 [ label = <6) \( \text{[ACG]} = \( \text{[EDG]} \) < BR/>
<FONT POINT-SIZE="10">Rule 2</FONT>>, tooltip = "For two angles ∠[l1,l2], ∠[l3,l4]. If l1 = l3 and l2 | l4, then ∠[l1,l2]
<FONT POINT-SIZE="10">Rule 3</FONT>>. tooltip = "If AB | CD and E is the intersection of AC and BD, then EA / EC = EB / ED
8 [ label = <8) BC·DE = BE·AC<BR/>
<FONT POINT-SIZE="10">Rule 3</FONT>>, tooltip = "If AB | CD and E is the intersection of AC and BD, then EA / EC = EB / ED
<FONT POINT-SIZE="10">Rule 35</FONT>>, tooltip = "Middle Connection Theorem for Triangles.\nThe line connects the midpoint
10 [ label = <10) BC | DF<BR/>
<PONT POINT-SIZE="10">Rule 35</FONT>>, tooltip = "Middle Connection Theorem for Triangles.\nThe line connects the midpoint:
                           fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                            fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                            fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                            fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip =
                            fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip = " " ];
"D is the midpoint of BA" [
                            fillcolor = "#F7CAC9", shape = oval, style = "filled", tooltip = " "
```

Database of GeoGebra files, joint work with Soma Bartha, June 2024, Linz, Austria



Discussion, joint work with Tomás Recio, February 2025, Madrid, Spain



Joint work with S. Bartha and T. Recio, published in May 2025



Joint work with S. Bartha and T. Recio, published in May 2025

#### **Abstract**

We elaborate, as a benchmark, a list of geometry theorems that are regularly used in Hungarian high schools, in specialized tracks for mathematics. Then, these statements are implemented as GeoGebra files, and algebraically proven (with certification) in GeoGebra Discovery. Next, these files are imported in Java Geometry Expert (JGEX), and geometrically proven with the associated Geometry Deductive Database method. In both programs, through quite different approaches in each case, we define and implement an algorithmic measure of the difficulty of a geometric theorem. Then, we compute the corresponding measures on the selected list of statements, and we analyze the obtained collection of pairs of difficulty grades, finding a moderate positive correlation (0.51) between GeoGebra Discovery's algebraic, and JGEX's deductive difficulty formulations. A correlation that seems to support the adequacy of the chosen approximation in our on-going work to algorithmically establish some *ad-hoc* measure of the difficulty of geometric statements that could be reasonably coincidental with the appreciation of human expert users (e.g. teachers).

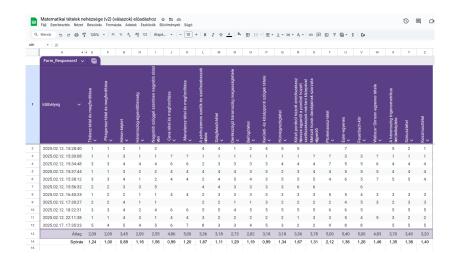
**Keywords** Automated theorem prover  $\cdot$  Algebraic methods  $\cdot$  Deductive Database  $\cdot$  GeoGebra Discovery  $\cdot$  Java Geometry Expert  $\cdot$  Euclidean Geometry

# Comparison of difficulty of 14 classroom problems

Theorem	$C_a$	$C_g$	$C_h$	Ldp	#v	#e
Thales' ○ thm	4	13	2.1	2	4	3
Converse of Thales' ○ thm	4	13	2.1	1	3	2
Midline thm	2	10		1	3	2
Intercept thm	7	8	3.4	1	2	1
Theorem of inscr. $\angle$ in a $\bigcirc$	7		3.2	1	2	1
∠ of cyclic quadrilaterals			3.2	1	2	1
Rel. betw. inscr. $\angle$ and $\angle$ at the $\odot$			3.2	2	5	4
Tangent-secant thm	6	16	3.8	3	5	4
Wallace-Simson thm	16	53	4.8	8	20	19
Euler's line	7	120	5.4	11	34	46
9-point ○	18	94	5	5	23	32
Circumcenter of a $\triangle$	2	25		3	9	8
Barycenter of a △	5	43		5	9	10
Apollonius' thm	7		3.2	5	12	11

#### Results of the questionnaire

 $C_h$  is taken as the average of the answers (line 13), stddev is shown in line 14



#### A.K.'s classmates

Some of them participated in filling in a questionnaire



#### Correlation computations

	$C_g$	$C_h$	Ldp	#v	#e
Ca	0.58	0.71	0.47	0.57	0.54
$C_g$ $C_h$		0.84	0.88	0.98	0.99
$C_h$			0.81	0.86	0.85
Ldp				0.94	0.90
#v					0.99

The correlation computation was performed via Pearson's correlation formula for samples,

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}}$$

where n is the sample size, and  $x_i$  and  $y_i$  are the individual sample points indexed with i,  $1 \le i \le n$ , moreover,  $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$  is the mean for sample points  $x_1, \ldots, x_n$ , the same for  $\bar{y}$ , respectively.

#### Conclusion and future work

- ▶ C<sub>g</sub> ~ C<sub>a</sub> is the closest geometric complexity.
- #v and #e are very close to C<sub>g</sub>.
- $ightharpoonup C_h$  is, surprisingly, between  $C_a$  and  $C_g$ .
- Both the amount of theorems and the number of participants in the questionnaire are quite low. The research should be therefore repeated
  - in the same school, in different classes of young learners,
  - in other Hungarian schools, in special math program classes,
  - in other countries,

to work on a larger dataset.

 $ightharpoonup C_a$  could be defined differently, e.g. by taking the maximal number of terms in the  $f_i$  in the formula

$$1 = f_1h_1 + f_2h_2 + \ldots + f_rh_r + f_{r+1} \cdot (zt - 1).$$



# THANK YOU!

#### References I



C. Abar and M. V. de Almeida.

Levels of estimation of the complexity of a geometric statement, Feb. 2025.

Presentation at I Workshop IAxEM, 20 February 2025, Universidad Nebrija, Madrid.



S. Bartha, A. Kerekes, Z. Kovács, and T. Recio.

Automated analysis of the difficulty of secondary school geometry theorems. *Annals of Mathematics and Artificial Intelligence*, 2025.



S. Colton, A. Bundy, and T. Walsh.

On the notion of interestingness in automated mathematical discovery. *International Journal of Human-Computer Studies*, 53(3):351–375, 2000.



E. R. Gansner and S. C. North.

An open graph visualization system and its applications to software engineering. Software—Practice and Experience, 30(11):1203–1233, 2000.



H. Gao, Y. Goto, and J. Cheng.

Explicitly epistemic contraction by predicate abstraction in automated theorem finding: A case study in nbg set theory.

In N. T. Nguyen, editor, Intelligent Information and Database Systems, 7th Asian Conference, ACIIDS 2015, Bali, Indonesia, March 23-25, 2015, Proceedings, Part I, Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science), Vol. 9011, pages 593–602, Cham, 2015. Springer.

#### References II



H. Gao, J. Li, and J. Cheng.

Measuring interestingness of theorems in automated theorem finding by forward reasoning: A case study in Tarski's geometry.

In 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), pages 168–173, 2018.



A. Kerekes and Z. Kovács.

Java Geometry Expert: Visualizing proofs as a directed acyclic graph, Feb. 2025.

Presentation at I Workshop IAxEM, 20 February 2025, Universidad Nebrija, Madrid.



Z. Kovács, T. Recio and M. P. Vélez.

Showing Proofs, Assessing Difficulty with GeoGebra Discovery.

Electronic Proceedings in Theoretical Computer Science, 398:43–52, 2024.



Z. Kovács.

Automated detection of interesting properties in regular polygons.

Mathematics in Computer Science, 14:727–755, 2020.



Y. Puzis, Y. Gao, and G. Sutcliffe.

Automated generation of interesting theorems.

In G. C. J. Sutcliffe, editor, *Proceedings of the Nineteenth International Florida Artificial Intelligence Research Society Conference (FLAIRS 2006*), pages 49–54, 2006.

#### References III



P. Quaresma, P. Graziani, and S. M. Nicoletti.

Considerations on approaches and metrics in automated theorem generation/finding in geometry.

In Proceedings ADG 2023, 2023.



P. Todd.

A method for the automated discovery of angle theorems.





L. Wos.

The problem of automated theorem finding.

Journal of Automated Reasoning, 10:137–138, 1993.



Z. Ye, S. C. Chou, and X. S. Gao.

An Introduction to Java Geometry Expert.

In Automated Deduction in Geometry, pages 189–195. Springer Science + Business Media, 2011.