

Geometry Automated Theorem Provers Systems Competition

0.2 Report

Pedro Quaresma
CISUC, Department of Mathematics
University of Coimbra, Portugal
&
Nuno Baeta
CISUC,
University of Coimbra, Portugal

Geometry Automated Theorem Provers Systems Competition

0.2 Report

Pedro Quaresma
CISUC/Department of Mathematics
University of Coimbra
E-mail: pedro@mat.uc.pt

Nuno Baeta
CISUC,
University of Coimbra, Portugal
e-mail: nmsbaeta@gmail.com

November, 2019

Abstract

The implementation of a competition between Geometry Automated Theorem Provers (GATP) would allow to create a test bench for GATP developers to improve the existing ones and to propose new ones. It would also allow to establish a ranking for GATP that could be used by “clients” (e.g. developers of educational e-learning systems) to choose the best implementation for a given intended use.

With the implementation of Geometry Automated theorem provers Systems Competition 0.2 (GASC 0.2), a sort of test-run, we intend to start the route towards a geometry automated theorem provers systems competition.

Chapter 1

Introduction

To be able to compare the different methods and implementations, a competition will have the virtue of pushing towards the standardisation of the input language, the standardisation of test sets, the direct comparability and the easier exchange of ideas and algorithmic techniques. The results of such a competition will also constitute a showcase, where potential users will look for the best GATP for their goals.

The first step toward a **G**eometry **A**utomated theorem provers **S**ystems **C**ompetition (GASC) was given at ThEdu'19, the 8th International Workshop on Theorem proving components for Educational software, August 2019, Natal, Brazil. At ThEdu'19 a presentation was made and a first trial, GASC 0.1, was conducted in a local computer (the first author laptop) using two scripts: one to launch the competition and follow it and another script to see the results in a never ending loop.

Chapter 2

GASC 0.2

In the preparation of ThEdu’19 post-proceedings at EPTCS, incorporating all the comments received during the presentation, e.g. the Toolympics reference [3] the GASC 0.2 was implemented. The major difference between GASC 0.1 and GASC 0.2 is in the use of an Web server to support the competition [1].

The server that supported GASC 0.2 was a Linux system, `Linux 4.9.0-2-amd64 #1 SMP Debian 4.9.18-1 (2017-03-30) x86_64 GNU/Linux`. The desktop computer motherboard is a Intel(R) Core(TM) i7-4770 CPU @ 3.40GHz with 16GiB of RAM.

The GATPs that entered GASC 0.2 were:

GCLC *GCLC* is a tool for visualising objects and notions of geometry and other fields of mathematics, by generating figures and animations in the *gc* language, it has a built-in geometry theorem prover that can automatically prove a range of complex problems. The GATP module implements the *Area Method*, the *Wu’s method* and the *Gröbner Basis Method* [7, 8, 9].

GCLC AM — Implementation of the area method [9];

GCLC WM — Implementation of the Wu’s method [6];

GCLC GBM — Implementation of the Gröbner bases method [6].

CoqAM The formalisation of the area method using the proof assistant *Coq* was done by implementing the decision procedure as a *Coq* tactic and formalising all theorems needed by the method. The implementation guarantee the soundness of the method implementation, i.e., the proofs generated by the tactic are always correct [2, 9, 12].

Coq AM — Implementation of the Area Method in *Coq* [9, 12].

GeoGebra’s portfolio prover an embedded prover system in GeoGebra that is capable of using multiple internal back ends for proving theorems [10].

GeoGebra RM — Implementation of the Recio’s exact check method [11];

GeoGebra BM — Implementation of the Gröbner Basis method [4, 5].

The database TGTP, Thousand of Geometric problems for geometric Theorem Provers [13] was used as a source of geometric conjectures (problems *GEONNNN*).

2.1 Results by TGTP Problem

GASC 0.2 - 2019-11-14T15:49:05+01:00

	GCLC			Coq	GeoGebra	
Problem	AM	WM	GBM	AM	RM	BM
GEO0001	0.05	0.00	0.00	0.86	—	—
GEO0002	0.03	0.77	60.00	—	—	—
GEO0003	0.00	0.00	59.99	—	—	—
GEO0004	0.00	0.00	0.00	—	—	—
GEO0005	0.01	0.00	0.47	—	—	—
GEO0006	0.00	0.00	0.01	—	—	—
GEO0007	0.00	0.00	0.00	—	—	—
GEO0008	0.00	0.00	0.00	1.97	—	—
GEO0009	0.00	0.00	0.00	—	—	—
GEO0010	0.08	0.00	0.00	—	—	—
GEO0011	0.08	0.00	0.00	—	—	—
GEO0012	10.25	18.39	59.99	—	—	—
GEO0013	0.00	0.00	0.04	0.62	—	—
GEO0014	0.00	0.00	0.00	—	—	—
GEO0015	0.08	0.05	0.18	0.76	—	—
GEO0020	2.64	60.00	60.06	—	—	—
GEO0021	0.01	0.00	0.05	—	—	—
GEO0022	2.28	0.10	2.55	—	—	—
GEO0023	0.00	0.12	59.99	—	—	—
GEO0024	0.09	0.00	0.01	—	—	—
GEO0025	0.00	0.00	0.00	—	—	—
GEO0027	6.85	0.48	60.00	—	—	—
GEO0028	10.03	60.01	60.00	—	—	—
GEO0080	0.00	0.10	60.00	2.47	0.09	0.297
GEO0170	—	—	—	1.00	—	—
GEO0171	—	—	—	0.81	—	—
GEO0172	—	—	—	9.15	—	—
GEO0173	—	—	—	0.78	—	—
GEO0174	—	—	—	1.82	—	—
GEO0175	—	—	—	0.56	—	—
GEO0176	—	—	—	0.52	—	—
GEO0177	—	—	—	2.06	—	—
GEO0178	—	—	—	8.08	—	—
GEO0179	—	—	—	13.22	—	—
GEO0180	—	—	—	8.67	—	—
GEO0181	—	—	—	1.04	—	—
GEO0182	—	—	—	1.39	—	—
GEO0183	—	—	—	1.53	—	—
GEO0184	0.00	0.00	0.09	1.74	—	—
GEO0185	—	—	—	0.74	—	—

	GCLC			Coq	GeoGebra	
Problem	AM	WM	GBM	AM	RM	BM
GEO0186	—	—	—	1.91	—	—
GEO0187	—	—	—	2.62	—	—
GEO0188	—	—	—	1.28	—	—
GEO0189	—	—	—	0.59	—	—
GEO0190	0.03	0.14	30.88	6.51	—	—
GEO0191	0.16	0.04	0.09	12.68	—	—
GEO0192	0.00	0.04	0.03	0.89	—	—
GEO0193	0.00	0.01	0.09	1.23	—	—
GEO0194	0.00	0.01	0.09	1.85	—	—
GEO0195	0.00	0.00	0.03	0.45	—	—
GEO0196	—	—	—	1.20	—	—
GEO0197	—	—	—	2.36	—	—
GEO0198	—	—	—	1.59	—	—
GEO0199	0.00	0.06	0.08	0.90	—	—
GEO0200	—	—	—	8.07	—	—
GEO0201	0.00	0.05	0.29	4.02	—	—
GEO0202	0.00	1.90	3.19	2.33	—	—
GEO0203	0.00	0.02	0.11	2.31	—	—
GEO0204	0.00	0.16	59.99	0.69	—	—
GEO0205	—	—	—	0.51	—	—
GEO0206	—	—	—	1.96	—	—
GEO0207	—	—	—	3.36	—	—
GEO0208	—	—	—	0.93	—	—
GEO0209	—	—	—	0.93	—	—
GEO0210	—	—	—	10.06	—	—
GEO0211	—	—	—	6.85	—	—
GEO0212	—	—	—	0.48	—	—
GEO0213	—	—	—	0.74	—	—
GEO0214	—	—	—	0.57	—	—
GEO0215	—	—	—	0.53	—	—
GEO0216	—	—	—	0.59	—	—
GEO0217	—	—	—	0.70	—	—
GEO0218	—	—	—	1.46	—	—
GEO0219	—	—	—	0.58	—	—
GEO0220	—	—	—	5.16	—	—
GEO0221	—	—	—	3.38	—	—
GEO0222	0.00	0.09	60.00	—	—	—
GEO0223	0.00	0.04	0.00	—	—	—
GEO0224	0.00	0.00	0.04	—	—	—
GEO0225	0.00	0.10	59.99	—	—	—
GEO0226	0.26	0.54	60.00	—	0.204	2.185
GEO0227	0.60	0.11	60.05	—	—	—
GEO0228	0.00	0.00	0.03	—	0.032	0.265
GEO0230	0.60	1.05	60.00	—	—	—

	GCLC			Coq	GeoGebra	
Problem	AM	WM	GBM	AM	RM	BM
GEO0231	0.00	0.03	0.07	4.37	—	—
GEO0232	0.00	0.00	0.00	—	—	—
GEO0233	0.08	0.01	0.45	—	—	—
GEO0234	0.00	0.00	0.04	0.36	—	—
GEO0235	0.49	0.00	0.04	20.02	—	—
GEO0236	0.60	0.54	1.60	20.12	—	—
GEO0237	0.28	0.06	60.00	20.02	0.007	0.261
GEO0238	0.01	0.04	0.32	20.01	0.039	0.255
GEO0239	0.00	0.09	59.99	—	—	—
GEO0240	0.00	0.08	60.03	—	0.008	0.273
GEO0241	0.09	0.03	17.91	—	—	—
GEO0242	0.69	0.12	59.99	—	—	—
GEO0243	0.65	0.06	0.00	17.59	—	—
GEO0244	0.00	0.09	0.05	—	—	—
GEO0245	0.03	0.01	8.53	—	—	—
GEO0246	0.00	0.00	0.05	—	—	—
GEO0247	0.04	0.00	0.00	1.71	—	—
GEO0248	0.00	0.00	0.00	—	—	—
GEO0249	0.00	0.00	0.01	—	—	—
GEO0250	0.00	0.05	0.17	3.18	0.008	5.025
GEO0251	0.02	0.19	60.06	—	0.008	0.389
GEO0252	0.09	0.04	0.04	—	0.007	0.359
GEO0253	0.12	0.00	0.04	—	—	—
GEO0254	0.00	0.18	60.00	—	0.008	5.024
GEO0255	0.07	0.00	0.02	—	—	—
GEO0256	0.00	0.49	26.15	0.50	—	—
GEO0257	0.63	0.00	0.05	20.01	—	—
GEO0258	0.00	0.17	59.99	—	0.06	5.025
GEO0259	0.00	0.04	0.08	—	—	—
GEO0260	0.62	1.13	60.07	20.01	0.249	0.334
GEO0261	0.00	0.04	60.05	20.02	0.007	0.295
GEO0262	0.00	0.01	0.09	—	—	—
GEO0263	0.00	0.00	0.00	—	—	—
GEO0264	0.00	1.34	16.18	—	—	—
GEO0265	0.00	0.00	0.01	—	0.006	0.364
GEO0266	0.00	0.04	0.07	—	0.006	0.379
GEO0267	0.60	0.00	0.09	—	—	—
GEO0268	0.02	1.93	60.00	—	—	—
GEO0269	0.00	0.00	0.00	—	—	—
GEO0270	0.00	0.00	0.01	—	—	—
GEO0271	0.01	0.00	0.02	—	—	—
GEO0272	0.01	0.00	0.00	—	0.007	0.247
GEO0273	0.00	1.64	3.57	20.01	0.008	0.245
GEO0274	0.00	0.22	59.99	—	0.007	0.251

	GCLC			Coq	GeoGebra	
Problem	AM	WM	GBM	AM	RM	BM
GEO0275	0.60	0.04	0.20	—	—	—
GEO0276	0.00	0.00	0.04	—	0.008	0.248
GEO0277	0.00	0.00	0.10	—	—	—
GEO0278	0.00	0.12	60.00	—	0.008	0.3
GEO0279	0.00	0.88	60.04	—	0.008	5.026
GEO0280	0.00	0.02	60.07	—	0.007	5.023
GEO0281	0.00	0.86	59.99	—	0.009	5.024
GEO0282	0.00	0.03	0.10	—	0.009	0.252
GEO0283	0.00	0.04	0.08	—	0.007	0.248
GEO0284	0.00	0.09	60.19	—	—	—
GEO0285	0.00	0.01	0.57	—	0.015	0.253
GEO0286	0.00	0.00	0.11	—	—	—
GEO0287	0.00	0.00	0.00	—	0.007	0.328
GEO0288	0.00	0.70	59.99	—	0.008	0.41
GEO0289	0.60	0.04	0.06	—	—	—
GEO0290	0.00	0.05	0.05	—	—	—
GEO0291	0.00	0.04	0.00	20.02	0.022	0.238
GEO0292	0.00	0.00	0.17	—	—	—
GEO0293	0.00	0.00	0.01	—	—	—
GEO0294	0.00	1.25	59.99	—	—	—
GEO0295	0.02	0.03	60.09	—	0.007	0.417
GEO0296	0.00	0.00	0.08	10.03	0.006	0.398
GEO0297	0.00	0.06	60.00	—	—	—
GEO0298	0.00	0.00	0.00	—	0.006	0.582
GEO0299	0.00	0.04	0.05	—	—	—
GEO0300	0.00	0.07	60.06	—	—	—
GEO0301	0.00	0.00	0.01	0.83	—	—
GEO0302	0.02	0.00	8.18	2.41	0	0
GEO0303	0.01	0.09	3.92	—	0.008	0.268
GEO0304	0.00	0.04	0.12	—	—	—
GEO0305	0.00	0.03	0.03	—	—	—
GEO0306	0.00	0.07	60.06	—	—	—
GEO0307	0.00	0.00	0.00	—	—	—
GEO0308	0.00	0.00	0.00	—	—	—
GEO0309	0.00	0.04	1.67	—	—	—
GEO0310	0.00	0.03	60.05	—	—	—
GEO0311	0.00	0.00	0.32	—	—	—
GEO0312	0.00	0.00	0.46	—	0.006	0.35
GEO0313	0.00	0.00	0.08	0.48	0.006	0.24
GEO0314	0.00	0.00	0.09	—	—	—
GEO0315	0.60	0.01	0.19	20.02	—	—
GEO0316	0.00	0.35	60.06	—	0.008	0.25
GEO0317	0.00	0.04	0.05	—	—	—
GEO0318	0.00	0.00	0.00	—	—	—

	GCLC			Coq	GeoGebra	
Problem	AM	WM	GBM	AM	RM	BM
GEO0319	0.08	0.01	0.10	—	0.14	0.272
GEO0320	0.00	0.03	2.16	—	—	—
GEO0321	10.06	0.06	12.68	—	0.018	0.271
GEO0322	0.00	0.04	9.25	—	—	—
GEO0323	0.00	0.00	0.00	—	0.028	0.295
GEO0324	0.00	0.00	0.03	—	0.007	0.268
GEO0325	0.00	0.06	55.08	—	0.008	0.292
GEO0326	0.00	60.00	59.18	—	0.009	5.036
GEO0327	0.00	1.22	60.12	—	0.009	5.021
GEO0328	0.00	0.00	0.00	—	0.008	5.026
GEO0329	0.06	0.02	0.05	—	0.021	0.273
GEO0330	0.00	0.00	21.08	—	0	0
GEO0331	0.00	0.00	0.00	—	0.027	0.337
GEO0332	0.12	0.00	0.60	—	0.155	0.287
GEO0333	0.00	0.00	0.00	—	0.01	5.024
GEO0334	0.00	0.00	0.00	—	0.01	5.022
GEO0335	0.00	0.00	0.00	—	—	—
GEO0336	0.00	0.00	0.06	—	—	—
GEO0337	0.06	0.31	59.12	—	0.008	0.298
GEO0338	0.00	0.00	0.00	—	0.008	0.232
GEO0339	0.00	0.00	0.06	—	0.023	0.261
GEO0340	60.63	0.00	0.07	—	0	0
GEO0341	0.06	0.00	0.10	—	0.022	0.276
GEO0343	0.06	0.00	0.10	—	0.01	0.264
GEO0344	0.00	0.00	0.00	—	0.021	0.266
GEO0345	0.00	0.00	0.00	—	0.009	2.892
GEO0346	0.00	0.11	59.99	—	0.009	5.018
GEO0347	0.00	0.00	0.00	—	0.008	1.135
GEO0349	0.00	1.72	59.10	—	0.006	5.027
GEO0350	0.00	0.00	0.00	—	0.008	0.305
GEO0351	0.00	0.00	0.00	—	0.009	0.481
GEO0352	0.00	0.01	0.25	—	0.008	0.37
GEO0353	0.00	59.15	59.47	—	0	0
GEO0354	0.00	0.00	0.00	—	0.007	0.371
GEO0355	0.00	0.00	0.00	—	0.007	0.294
GEO0356	0.06	0.04	0.13	—	0.01	0.279
GEO0357	0.06	0.04	0.06	—	0.01	0.257
GEO0358	0.06	20.79	59.17	—	0.007	0.427
GEO0359	0.00	0.22	59.11	—	0.009	0.42
GEO0360	0.56	0.00	0.35	—	0.007	0.269
GEO0361	0.02	0.06	59.11	—	0.008	0.319
GEO0362	0.00	0.00	0.12	—	0.008	0.272
GEO0364	0.00	0.00	0.07	—	0.008	1.97
GEO0365	—	—	—	—	0.008	0.283

	GCLC			Coq	GeoGebra	
Problem	AM	WM	GBM	AM	RM	BM
GEO0366	—	—	—	—	0.008	2.247
GEO0368	—	—	—	—	0.007	5.028
GEO0369	—	—	—	—	0.006	0.249
GEO0370	0.00	0.00	0.22	—	0.01	0.251
GEO0372	0.90	0.38	59.99	—	0.006	0.265
GEO0373	0.00	0.00	0.00	—	0.008	0.271
GEO0374	0.00	0.00	0.00	—	0.006	0.245
GEO0375	9.06	0.23	60.05	—	0.007	0.29

2.2 Results by GATP

GASC 0.2 - 2019-11-14T15:49:05+01:00

GATP	CPU time			Success proved/attempted
	Average	min	max	
GCLC Area Method	0.34832402234637	0.00	10.25	179/224
GCLC Wu's Method	1.3553631284916	0.00	60.00	179/224
GCLC Gröbner Basis Method	13.460304878049	0.00	60.00	164/224
Coq Area Method	5.114025974026	0.36	20.12	77/224
GeoGebra Recio's Method	0.019552941176471	0	0.249	85/224
GeoGebra Botana's Method	1.1715764705882	0	5.036	85/224

Chapter 3

Future Work

The organization of the competition, in the long run, will require the support of the geometry automated deduction community: by entering the competition; by setting a problems committee that would choose the set of problems to be solved by the GATP, and maybe the more important point, by using its outcomes to their research and/or applications.

It is plan that the first edition of GASC would occur at ADG2020, the 13th International Workshop on Automated Deduction in Geometry, Hagenberg, Austria, 13-15 July, 2020.

Bibliography

- [1] Nuno Baeta, Pedro Quaresma, and Zoltán Kovács. Towards a geometry automated provers competition. In *Proceedings 8th International Workshop on Theorem proving components for Educational software*, 2020. (in preparation).
- [2] Yves Bertot and Pierre Castran. *Interactive Theorem Proving and Program Development*. Texts in Theoretical Computer Science. An EATCS Series. Springer, 2004.
- [3] Dirk Beyer, Marieke Huisman, Fabrice Kordon, and Bernhard Steffen, editors. *Tools and Algorithms for the Construction and Analysis of Systems: 25 Years of TACAS: TOOLympics*, volume 11429 of *LNCS*. Springer, 2019. Held as Part of ETAPS 2019, Prague, Czech Republic, April 611, 2019.
- [4] Francisco Botana and Zoltán Kovács. A Singular web service for geometric computations. *Annals of Mathematics and Artificial Intelligence*, pages 1–12, November 2014. <http://dx.doi.org/10.1007/s10472-014-9438-2>.
- [5] Francisco Botana, Zoltán Kovács, and Simon Weitzhofer. Implementing theorem proving in GeoGebra by using a Singular webservice. In *Proceedings EACA 2012, Libro de Resúmenes del XIII Encuentro de Álgebra Computacional y Aplicaciones*, pages 67–70. Universidad de Alcalá, 2012.
- [6] Shang-Ching Chou, Xiao-Shan Gao, and Jing-Zhong Zhang. *Machine Proofs in Geometry*. World Scientific, 1994.
- [7] Predrag Janičić. GCLC A tool for constructive euclidean geometry and more than that. In Andrés Iglesias and Nobuki Takayama, editors, *Mathematical Software - ICMS 2006*, volume 4151 of *Lecture Notes in Computer Science*, pages 58–73. Springer, 2006.
- [8] Predrag Janičić. Geometry constructions language. *J. Autom. Reasoning*, 44(1-2):3–24, 2010.
- [9] Predrag Janičić, Julien Narboux, and Pedro Quaresma. The Area Method: a recapitulation. *Journal of Automated Reasoning*, 48(4):489–532, 2012.
- [10] Zoltán Kovács. The portfolio prover in GeoGebra 5. In Francisco Botana and Pedro Quaresma, editors, *Proceedings of the 10th International Workshop on Automated Deduction in Geometry (ADG 2014), 9-11 July 2014*, pages 191–205. University of Coimbra, Portugal, 2014.

- [11] Zoltán Kovács, Tomás Recio, and Simon Weitzhofer. Implementing theorem proving in GeoGebra by using exact check of a statement in a bounded number of test cases. In *Proceedings EACA 2012, Libro de Resúmenes del XIII Encuentro de Álgebra Computacional y Aplicaciones*, pages 123–126. Universidad de Alcalá, 2012.
- [12] Julien Narboux. Formalization of the area method. Coq user contribution, 2009. http://dpt-info.u-strasbg.fr/~narboux/area_method.html.
- [13] Pedro Quaresma. Thousands of Geometric problems for geometric Theorem Provers (TGTP). In Pascal Schreck, Julien Narboux, and Jürgen Richter-Gebert, editors, *Automated Deduction in Geometry*, volume 6877 of *Lecture Notes in Computer Science*, pages 169–181. Springer, 2011.