#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### **TGTP** - Thousands of Geometric problems for geometric Theorem Provers

Pedro Quaresma

CISUC, Mathematics Department University of Coimbra

ADG 2010, Munich, July 22, 2010

- 32

1/34

#### Pedro Quaresma

### TGTP Goal

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

The principal motivation in building **TGTP**<sup>1</sup> is to support the testing and evaluation of geometric automated theorem proving (GATP) systems, to help ensure that performance results accurately reflect the capabilities of the GATP system being considered.

**TGTP** is a library of problems, in geometry, for GATP systems.

**TGTP** aims to supply the automated reasoning in geometry community with a comprehensive library of GATP test problems, in order to provide an overview and a simple, unambiguous reference mechanism.

<sup>&</sup>lt;sup>1</sup>http://hilbert.mat.uc.pt/TGTP <□> <♂> <≧> <≧> ≥ ⊙<

Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### State of the Art

There are several systems integrating dynamic geometry softwares (DGS), GATPs, and a set of examples. For example:

- Java Geometry Expert (JGEX) (...). It contains a large set of examples of proofs.
- *GEOTHER* (...). It contains a collection of theorems in both elementary and differential geometry.
- GeoThms (...) a library of geometric problems (...).

Many of the DGSs (*GeoGebra, Cabri, Cinderella,* ...), DGSs/GATPs (*GCLC, GeoView, GeoProof, Geometry Explorer, MMP/Geometer, GEX, Discover,* ...), and also GATPs (*Theorema,* ...) come with a (some times, large) set of examples.

However none of them try to provide a common platform for meaningful system evaluations and comparisons.

Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### State of the Art & TGTP

A common library of problems is necessary for meaningful system evaluations and comparisons, its size is important if the production of statistically significant results is intended.

The goal for building **TGTP** is, in a similar spirit of *TPTP* and other libraries of problems, to provide the GATP community with a centralised problem collection with an easy access to all researchers.

Pedro Quaresma

#### TGTP Goa

State of the art

Relevant Issues

Conclusions

### Relevant Issues: Web-Based

**TGTP** tries to address all relevant issues. In particular:

• is Web-based and is thus easily available to the research community.

HEATTGTP - Thousand of Geometric problems for geometric Theorem	n Provers - Iceweasel	• 0
Echerio Editar yer Historico Barcadores Fejramentas guda	lindes php	<u>.</u>
Most Visited * P The Mozilia Organiz Vilatest Builds 🔂 noticias.rt	p.pt / * 🛐 noticias.rtp.pt / * 🔂 noticias.rtp.pt / *	
TGTP - Thousand of Geo	ometric problems for geor	netric Theorem Provers
Anonymous Login	Registrated Login	Help
The principal geometric The automated the accurately refi	motivation in building <i>TGTP</i> (Thousands of Geometric corem Provers) is to support the testing and evaluation corem proving (CATP) systems, to hole onsure that pedro end the capabilities of the GATP system being considered.	problems for of geometric mance results
© Pedr	o Quaresma (CISUC, Department of Mathematics, University of Co	oimbra)
oncluido		

Pedro Quaresma

#### TGTP Goa

State of the art

Relevant Issues

Conclusions

### Relevant Issues: Web-Based

**TGTP** tries to address all relevant issues. In particular:

• is Web-based and is thus easily available to the research community.

TGTP - Thousand of Geometric problems for geometric Theo	em Provers - Iceweasel	
teiro Editar Ver Histórico Marcadores Fegramentas Ajuda		1 / 1
i 🌳 🔹 📴 😰 🔛 🖂 🔤 http://hilbert.mat.uc.pt/TG	TP/Index.php	<u>≙_</u> [Ω•
Host Visited 🔻 🇭 The Mozilla Organiz 📱 Latest Builds 🔞 noticiae	.rtp.pt / 👻 🛐 noticials.rtp.pt / 👻 🔂 noticials.rtp.pt / 👻	
TGTP - Thousand of Ge	eometric problems for geor	metric Theorem Provers
Anonymous Login	Registrated Login	Help
The princip geometric T automated t accurately n	al motivation in building TGTP (Thousands of Geometric heorem Provers) is to support the testing and evaluation reorem proving (GATP) systems, to help ensure that perfor flect the capabilities of the GATP system being considered.	c problems for n of geometric mance results
© Pe	dro Quaresma (CISUC, Department of Mathematics, University of C	oimbra)

• Anonymous Login — see/get all the info;

Pedro Quaresma

#### TGTP Goa

State of the art

Relevant Issues

Conclusions

### Relevant Issues: Web-Based

**TGTP** tries to address all relevant issues. In particular:

• is Web-based and is thus easily available to the research community.

IGTP - Thousand of Geometric problems for geometric Theory	em Provers - Iceweasel	
iro Editar Ver Histórico Marcadores Ferramentas Ajuda		
🤶 🔹 😰 🕲 🏠 🚢 📫 🔞 http://filbert.mat.uc.pt/TG1	Prindex.php	<u> </u>
sst Visited 🔻 🎓 The Mozilia Organiz 🖉 Latest Builds 🔞 noticias.	rtp.pt / 🔻 🛐 noticias.rtp.pt / 🔻 🔯 noticias.rtp.pt / 👻	
TGTP - Thousand of Ge	ometric problems for geon	netric Theorem Provers
Anonymous Login	Registrated Login	Help
The principa geometric T automated th accurately re	I motivation in building TGTP (Thousands of Geometric reorem Provers) is to support the testing and evaluation eorem proving (GATP) systems, to help ensure that perfor lite: the capabilities of the GATP system being considered.	problems for to geometric mance results
0 Per	Iro Quaresma (CISUC, Department of Mathematics, University of Co	oimbra)

- Anonymous Login see/get all the info;
- Registered Login see/get/add/modify all the info.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### Easy to Use

• is easy to use. Problems (conjectures and proofs attempts) are presented in a small set of Web-pages with all the relevant information.

ΤG	TP - Thousand of	Geomet	ric problems for	geometric Theo	rem Prov	ers
Docu	ments/Help Proble	ems List	Workbench	Downloads	Logou	<u>ıt</u>
Show reset to defa	all • N. rows	10 • P	age 1 of 14 36 Problems	Pg Up Pg Dn End	Search	by Name 💌
Code	Name		Short Descrip	tion	N. Proofs/N. Attempts	
GEO0251	Six Points Circle				1/3	<u>See</u> <u>details</u>
GEO0281	Adams Circle				0/3	<u>See</u> <u>details</u>
GEO0227	Brahmagupta s Theorem				1/3	<u>See</u> <u>details</u>
GEO0268	Butterfly theorem				1/3	<u>See</u> details
GEO0272	Circumscribed Circle Exists Theorem				3/3	<u>See</u> <u>details</u>
GEO0228	Dual Altitude Theorem				3/3	<u>See</u> details

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### Easy to Use

• is easy to use. Problems (conjectures and proofs attempts) are presented in a small set of Web-pages with all the relevant information.

TGTP - Thousand of Geometric problems for geometric Theorem Provers					
Documents/Help	Problems List	Workbench	Downloads	Logout	

#### Problem: GEO0315

File	GE00315
Problem	Euler Line Theorem
Short Description	
Bibliographic Reference	
Contributer	Goran Predovic
Status	Proved
Rating (n.proofs/n.attempts)	3/4
Code XML	XML file Aux/GEO0315 20100713111712.xml (to download)
Proof Attempt 1	GCLC Area Method: 9.00, Proved
Proof Attempt 2	GCLC Wu Method: 9.00, Proved
Proof Attempt 3	GCLC Gröbner Basis Method: 9.00, Proved
Proof Attempt 4	COQareaMethod: 1.0, Time-out: Failed to prove the conjecture
Download Report File (text format)	report GE00315

#### Pedro Quaresma

#### TGTP Goa

State of the art

Relevant Issues

Conclusions

### Easy to Use

• is easy to use. Problems (conjectures and proofs attempts) are presented in a small set of Web-pages with all the relevant information.

DocumentsRielp	Problems List	Workbench	Downloads	Logour
Proof Attempt: GEO03	15 by GCLC Area Met	hod, v9.00		
Problem: GEO0315				
Command: gclc proof0315.code proof0315 pic -a				
Computer: hilbert.mat.uc.pt				
Model: x86_64 unknown				
CPU: Intel(R) Pentium(R) 4 CPU 3.00GHz				
# CPU cores: 2				
Memary: 2075840 kB				
BogoMips: 5989.8				
Operating System: 2.6.22-3-695-bigmern GNU/Lin	ĸ			
Result Proved				
CPU Time-Limit 600s				
Statistics:				
Time CPU: 11.581 Elimination steps: 386 Geometrics steps: 876 Algebraic steps: 5944				
ATP log file:				
OC language (M) -> LaTeX Converter Vers Written by Predrag Janicic, University of Copyright (c) 1996-2008. Not for commerci	ion 8.00 Belgrade. al use.			
Chinatan				
UD10CEB:				

Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

# Relevant Issues: Automatic Conversion

• Automatic conversion to other known formats is being provided, thus eliminating the necessity for any other transcription.



Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

# Relevant Issues: Automatic Conversion

• Automatic conversion to other known formats is being provided, thus eliminating the necessity for any other transcription.



• XML Suite (Pedro Quaresma et.al. in *Communicating* Mathematics in The Digital Era, 2008)

Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

# Relevant Issues: Automatic Conversion

• Automatic conversion to other known formats is being provided, thus eliminating the necessity for any other transcription.



- XML Suite (Pedro Quaresma et.al. in *Communicating* Mathematics in The Digital Era, 2008)
- i2g, InterGeo Common File Format + conjectures.

-

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### All Forms of GATP

• tries to cover the different forms of automated proving in geometry, e.g. synthetic proofs and algebraic proofs.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### All Forms of GATP

- tries to cover the different forms of automated proving in geometry, e.g. synthetic proofs and algebraic proofs.
- is independent of any particular GATP system.
  - GCLC Area Method (synthetic/algebraic);
  - Coq Area Method (synthetic/algebraic);
  - GCLC Wu's Method (algebraic);
  - GCLC Gröbner Bases Method (algebraic).

Pedro Quaresma

#### TGTP Goa

State of the art

Relevant Issues

Conclusions

### Statistically Significant Testing

• aims to become large enough for statistically significant testing. In its current version it contains already over 170 problems.



#### Statistics/Proof Status

	attempts	succeeded	%of success	min. time	max. time	avg. time
Coq • Area Method	76	68	0.89	0.73	213.71	17.698
GCLCprover - Area Method	123	62	0.5	0	360.235	9.194
GCLCprover - Wu's Method	96	88	0.92	0	6.404	0.422
GCLCprover - Gröbner Basis Method	96	56	0.58	0	112.319	5.393

Proved	1
Disproved	2
Failed to prove the conjecture	3
Time-out: Failed to prove the conjecture	4
Maximal number of proof steps reached: Failed to prove the conjecture	5
The conjecture out of scope of the prover	6

	Coq - A	rea Method	GCLCprove	er - Area Method	GCLCprove	r - Wu's Method	GCLCprover • 0	Gröbner Basis Method
TheoId	status	CPU time	status	CPU time	status	CPU time	status	CPU time
GE00001	1	3.32	1	0.004				
GE00002			1	0.028				
GE00003			1	0.003				
GE00004			1	0.002				
GEO0005			1	0.037				
GE00006			1	0.002				

Pedro Quaresma

#### TGTP Goa

State of the art

Relevant Issues

Conclusions

### Statistically Significant Testing

 aims to become large enough for statistically significant testing. In its current version it contains already over 170 problems.



#### Statistics/Proof Status

	attempts	succeeded	%of success	min. time	max. time	avg. time
Coq - Area Method	76	68	0.89	0.73	213.71	17.698
GCLCprover - Area Method	123	62	0.5	0	360.235	9.194
GCLCprover - Wu's Method	96	88	0.92	0	6.404	0.422
GCLCprover - Gröbner Basis Method	96	56	0.58	0	112.319	5.393

Proved	1
Disproved	2
Failed to prove the conjecture	3
Time-out: Failed to prove the conjecture	4
Maximal number of proof steps reached: Failed to prove the conjecture	5
The conjecture out of scope of the prover	6

	Coq - A	rea Method	GCLCprover	<ul> <li>Area Method</li> </ul>	GCLCprover	- Wu's Method	GCLCprover -	Gröbner Basis Method
TheoId	status	CPU time	status	CPU time	status	CPU time	status	CPU time
GE00001	1	3.32	1	0.004				
GE00002			1	0.028				
GE00003			1	0.003				
GE00004			1	0.002				
GE00005			1	0.037				
GE00006			1	0.002				

• aims to become a comprehensive, up-to-date library.

#### Pedro Quaresma

TGTP Goal

State of the art

Relevant Issues

Conclusions

### Well Structured

• it aims to be well structured and documented. This allows effective and efficient use of the library.



#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

## • documents each problem. This contributes to the unambiguous identification of each problem.

- bibliographic reference for each problem and (eventually) each proof.
- documentation for the different methods, and GATPs;
- List of bibliographic references.

### Well Documented

#### Pedro Quaresma

#### TGTP Goal

State of the art

#### Relevant Issues

Conclusions

### • a workbench to test/add new problems.

Proofs Listing 0200000 - New Carlecture	<ul> <li>Cheese a Cenjecture</li> </ul>
ar's Scrapbook	
Severing date name (max 20 chars)	8 : pascal ¥ Select Select
ridbench	
2 O II	4 CODerreiffethod 2.0
Theorem with a sequence of the	Individuals the code
is adjustic E B C -> Inter_11 O A E C F ->	Theorem - Centroid Theorem
3 0 ⇔ E -> 9 parallel A 0 0 € -> A +> 0 / 0 +> E = 2.	Proved, Cog Area Method
11 Proof. 13 area_sethod. 13 post.	Measures of efficiency
	Time spent (seconds) 1.940000
	hilbert.mat.uc.pt x895_64 unknown Intel(R) Pentium(R) 4 CPU 3.00GHz
	N. Cores: 2 BogoMIPS: 5989.8 OS: 2.6.22-3-696-bigmem GNU/Linux.

Adding New Problems

#### Pedro Quaresma

#### TGTP Goal

State of the art

#### Relevant Issues

Conclusions

### • a workbench to test/add new problems.

Proofs Listing GE00000 - New Conjecture	V Choose a Conjecture
ar's Scrapbook	
Severity date name (max 20 chars)	8 - poscal ¥ Select 8 - poscal ¥ Select
klench	
Benning Tempert area anthon	4 - C00ereaftethod - 1.0
Theorem evd 40 : forell 4 D C E P D : Point,	intercolute the code
Le adduite E B C -> Lefer 11 0 A E C F ->	Theorem - Centroid Theorem
parallel A 0 0 8 A ** 0 / 0 ** 5 - 2	Proved, Coq Area Method
l pros. Bres_sethod. 2 pd.	Measures of efficiency
	Time spent (seconds) 1.940000
	Competition         NOB_05         NO

Adding New Problems

• access to the list of problems in the database;

#### Pedro Quaresma

TGTP Goal

State of the art

Relevant Issues

Conclusions

### • a workbench to test/add new problems.

Proofs Listing GE00000 New Canjecture		×	Choose a Conjecture
Scrapbook			
SeverUpdate	name (max 20 chars)	8 pascal 💌 Select	8 rpascal * Selete
ench			
Benuire Tenert area atthed	Q 2 9 0 10 H × 9 2 2		4 CDQareaMethod 1.0
Theorem exe 40 : formil A B C E F O : Point, is midenic F E F O :		· · · · · · · · · · · · · · · · · · ·	Intervaluate the code
12_M1595177 F A 0 →   12_M1595177 F A 0 →   12_M15911 0 A E C F →   0 → E →			Theorem - Centroid Theorem
A ** 0 / 0 ** E = 2. Proof.			Proved, Coq Area Method
area_method. Ged.			Measures of efficiency
		Time spent (seconds)	1.940000
		hilbert.mat.uc Intel(R) Pentis	.pt x86_64 unknown m(R) 4 CPU 3.00GHz
		N. Cores: 2 OS: 2.6.22-3-6	BogoMIPS: 5989.8 96-bigmem GNU/Linux .

Adding New Problems

- access to the list of problems in the database;
- personal *scrapbook*;

#### Pedro Quaresma

#### TGTP Goal

State of the art

#### Relevant Issues

Conclusions

### • a workbench to test/add new problems.

Proofs Listing OC00000 - New Conjecture	* Choose a Cenjecture
's Scrapbook	
Sever/undate name (max 20 chars)	8 : pascal ¥ Select Select
dench	
45 Q 22 0 0 10 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	4 COQereaMethod 2.0
Theorem and 40 : farall A D C E F O : Point, in antidence F A R	indicablate the code
in a second seco	Theorem - Centroid Theorem
g parallel A 0 0 € -> A + 0 / 0 + € = 2. Front.	Proved, Cog Area Method
2 krea_wethod. 3 ged.	Measures of efficiency
	Time spent (seconds) 1.940000
	hilbert.mat.uc.pt x88_64 unknown Intel/RI Pantium/RI 4 CPU 3.00CHz
	N. Cores: 2 BogoMIPS: 5989.8 OS: 2.6.22-3-696-bigmem GNU/Linux.

Adding New Problems

- access to the list of problems in the database;
- personal scrapbook;
- addition of problems (referring|registered users) to the database;

#### Pedro Quaresma

#### TGTP Goa

State of th art

Relevant Issues

Conclusions

## The information, and the supporting programs will be freely available.

#### Pedro Quaresma

#### TGTP Goa

State of th art

Relevant Issues

Conclusions

The information, and the supporting programs will be freely available.

• The Database contents, and the ERD.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

The information, and the supporting programs will be freely available.

- The Database contents, and the ERD.
- The format conversion programs.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

The information, and the supporting programs will be freely available.

- The Database contents, and the ERD.
- The format conversion programs.
- The list of problems for each GATP.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

The information, and the supporting programs will be freely available.

- The Database contents, and the ERD.
- The format conversion programs.
- The list of problems for each GATP.
- Other supporting programs.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### **Open Problems**

• Common Format.

The extension of the i2g format for the GATPs.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### **Open Problems**

• Common Format.

The extension of the i2g format for the GATPs.  $\Rightarrow$  a script to apply all the GATPs to all the problems.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

### **Open Problems**

• Common Format.

The extension of the i2g format for the GATPs.  $\Rightarrow$  a script to apply all the GATPs to all the problems.

### • Querying.

The development of geometric/mathematic searching mechanism (semantic geo-contents, semantic math-contents).

### Conclusions

#### TGTP

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

In the *GeoThms* system the author and Predrag Janičić already addressed some of the issues that are now being laid down for TGTP, namely the XML common format, and the list of problems.

Where the *GeoThms* goal is to have a publicly accessible and widely used Internet based framework for constructive geometry with a strong integration of DGSs, GATPs and a library of problems to those tools.

The *TGTP* goal is to provide the GATP community with a centralised problem collection, independent of any particular GATP system.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

The development of TGTP problem library is an ongoing project, aiming to provide all of the desired properties described above.

#### Pedro Quaresma

#### TGTP Goal

State of the art

Relevant Issues

Conclusions

The development of *TGTP* problem library is an ongoing project, aiming to provide all of the desired properties described above.

## Thank You

イロト 不同下 イヨト イヨト

3

34 / 34