

# TGTP - Thousands of Geometric problems for geometric Theorem Provers

Pedro Quaresma

CISUC, Mathematics Department  
University of Coimbra

ADG 2010, Munich, July 22, 2010

# TGTP Goal

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>1</sup><http://hilbert.mat.uc.pt/TGTP>

## 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.

# State of the Art & TGTP

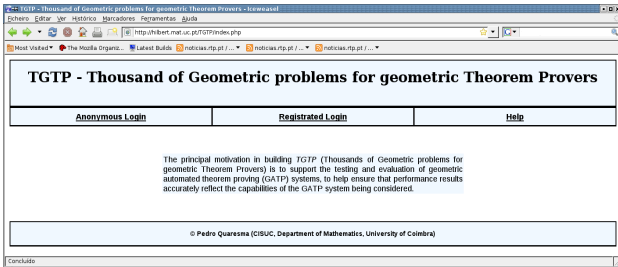
A common library of problems is necessary for meaningful system evaluations and comparisons, its size is important if the production of statistically significant results is intended.

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.

# Relevant Issues: Web-Based

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

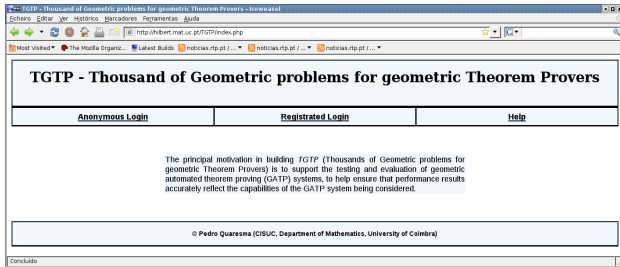
- is Web-based and is thus easily available to the research community.



# 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.

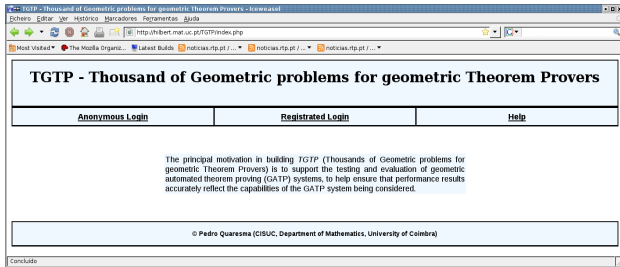


- Anonymous Login — see/get all the info;

# 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.



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

- 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				
<a href="#">Documents/Help</a>	<a href="#">Problems List</a>	<a href="#">Workbench</a>	<a href="#">Downloads</a>	<a href="#">Logout</a>
Show	all	N. rows	10	Page 1 of 14 136 Problems
				Home Pg Up Pg Dn End
				Search by Name
<a href="#">reset to default values</a>				
Code	Name	Short Description	N. Proofs/N. Attempts	
GEO0251	Six Points Circle		1/3	<a href="#">See details</a>
GEO0281	Adams Circle		0/3	<a href="#">See details</a>
GEO0227	Brahmagupta s Theorem		1/3	<a href="#">See details</a>
GEO0268	Butterfly theorem		1/3	<a href="#">See details</a>
GEO0272	Circumscribed Circle Exists Theorem		3/3	<a href="#">See details</a>
GEO0228	Dual Altitude Theorem		3/3	<a href="#">See details</a>



## 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				
<a href="#">Documents/Help</a>	<a href="#">Problems List</a>	<a href="#">Workbench</a>	<a href="#">Downloads</a>	<a href="#">Logout</a>

**Problem: GEO0315**

File	GEO0315
Problem	Euler Line Theorem
Short Description	
Bibliographic Reference	
Contributer	Goran Predovic
Status	Proved
Rating (n.proofs/n.attempts)	3/4
Code XML	XML file <a href="#">Aux/GEO0315_20100713111712.xml</a> (to download)
Proof Attempt 1	GCLC Area Method: 9.00, <a href="#">Proved</a>
Proof Attempt 2	GCLC Wu Method: 9.00, <a href="#">Proved</a>
Proof Attempt 3	GCLC Gröbner Basis Method: 9.00, <a href="#">Proved</a>
Proof Attempt 4	COQareaMethod: 1.0, <a href="#">Time-out: Failed to prove the conjecture</a>
Download Report File (text format)	<a href="#">report GEO0315</a>

## 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				
<a href="#">Documents/Help</a>	<a href="#">Problems List</a>	<a href="#">Workbench</a>	<a href="#">Downloads</a>	<a href="#">Logout</a>

**Proof Attempt: GEO0315 by GCLC Area Method, v9.00**

Problem: GEO0315

Command: gclc proof0315.code proof0315.pic -a

Computer: hbwt.mat.uc.pt

Model: x86\_64 unknown

CPU: Intel(R) Pentium(R) 4 CPU 3.00GHz

# CPU cores: 2

Memory: 2075840 kB

BigMem: 5893 B

Operating System: 2.6.22-3-686-bigmem GNU/Linux

Result: Proved

CPU Time Limit: 600s

Statistics:

Time CPU: 11.581

Elimination steps: 386

Geometrics steps: 876

Algebraic steps: 5944

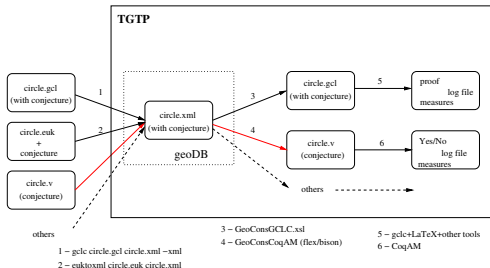
ATP log file

GC Language (R) -> LaTeX Converter Version 8.00  
 Written by Predrag Janjic, University of Belgrade,  
 Copyright (c) 1996-2008. Not for commercial use.

Objects:

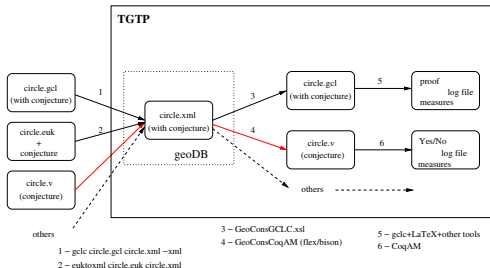
# Relevant Issues: Automatic Conversion

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



# Relevant Issues: Automatic Conversion

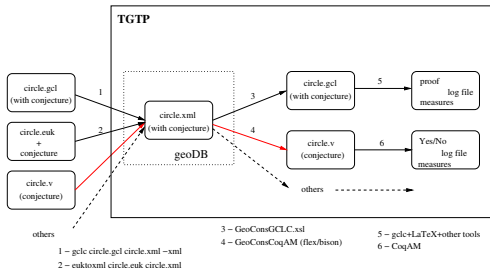
- Automatic conversion to other known formats is being provided, thus eliminating the necessity for any other transcription.



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

# 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.

# All Forms of GATP

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

# 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).

# Statistically Significant Testing

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

## TGTP - Thousand of Geometric problems for geometric Theorem Provers

[Documentation](#) | [Problems List](#) | [Workbench](#) | [Downloads](#) | [Logout](#)

### 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

Theoid	Coq - Area Method		GCLCprover - Area Method		GCLCprover - Wu's Method		GCLCprover - Gröbner Basis Method	
	status	CPU time	status	CPU time	status	CPU time	status	CPU time
GEO0001	1	3.32	1	0.004				
GEO0002			1	0.028				
GEO0003			1	0.003				
GEO0004			1	0.002				
GEO0005			1	0.037				
GEO0006			1	0.002				



# Statistically Significant Testing

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

## TGTP - Thousand of Geometric problems for geometric Theorem Provers

<a href="#">Documenta/Help</a>	<a href="#">Problems List</a>	<a href="#">Workbench</a>	<a href="#">Downloads</a>	<a href="#">Logout</a>
--------------------------------	-------------------------------	---------------------------	---------------------------	------------------------

### 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

Theoid	Coq - Area Method		GCLCprover - Area Method		GCLCprover - Wu's Method		GCLCprover - Gröbner Basis Method	
	status	CPU time	status	CPU time	status	CPU time	status	CPU time
GEO0001	1	3.32	1	0.004				
GEO0002			1	0.028				
GEO0003			1	0.003				
GEO0004			1	0.002				
GEO0005			1	0.037				
GEO0006			1	0.002				

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



# Well Documented

- 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.

## Adding New Problems

- a workbench to test/add new problems.

Select a Conjecture from the Database

Problems Listing: **CE00500 - New Centenary** Choose a Conjecture

User's Scrapbook

Save/Update **name (max 20 char)** 0 passed Select 0 passed Delete

Workbench

```

Require Import area_method.

Theorem a16_40 :
  forall A B C D F G : Prop,
  (a_midpoint F A B =>
   (a_midpoint E B C =>
    (infer_11_0 A E C F =>
     G => E
    )
   )
  )
  -> para1(A) A D B =>
  (a M D / G => G = 2.
  )
Proof.
area_method.
Qed.

```

**0 - CoqAreaMethod - 1.0** Introduce the code

**Theorem - Centroid Theorem**

Proved. Coq Area Method

**Measures of efficiency**

**Time spent (seconds)** 1.940000

**Computer** hilbert.mat.uc.pt x86\_64 unknown  
Intel(R) Pentium(R) 4 CPU 3.00GHz  
M. Cores: 2 SogoMIPS: 5669.8  
OS: 2.6.22-3-686-bigmem GNU/Linux.

Platform: Ln 13, Ch 5 | Total: Ln 13, Ch 259

# Adding New Problems

- a workbench to test/add new problems.

Select a Conjecture from the Database

Problems Listing: OE00000 - New Centenary Choose a Conjecture

User's Scrapbook

Event/Date	Name (max 20 char)	Passwd	Select	Passwd	Delete

Workbench

```
Require Import area_method.

Theorem a16_40 :
  forall A B C D F G : Prop,
  (a_midpoint F A B =>
  (a_midpoint E B C =>
  (infer_11_0 A E C F =>
  (D => E)
  => para1(A) A D B =>
  (A * D / G == B - 2.
  Proof.
  area_method.
  Qed.
```

CC0AreaMethod - 1.0 Introduce the code

**Theorem - Centroid Theorem**

Proved. Coq Area Method

**Measures of efficiency**

Time spent (seconds) 1.940000

Computer  
 hilbert.mat.uc.pt x86\_64 unknown  
 Intel(R) Pentium(R) 4 CPU 3.00GHz  
 N: Cores: 2 SogoMIPS: 5669.8  
 OS: 2.6.22-3-686-bigmem GNU/Linux.

Position Ln 13, Ch 5 Total Ln 13, Ch 219

- access to the list of problems in the database;

## Adding New Problems

- a workbench to test/add new problems.

Select a Conjecture from the Database

Problems Listing: CE05000 - New Centurion

User's Scrapbook

name (max 20 char)	B. pascol	Delete

Workbench

```
Require Import area_method.
Theorem a16_40 :
  forall A B C D F O : Point,
  is_midpoint F A B =>
  is_midpoint E B C =>
  Infer (1) O A E C F =>
  O = E.
paraM (x) A D B E =>
  A * D / O == B * E - 2.
Proof.
area_method.
Qed.
```

Theorem - Centroid Theorem

Proved, Coq Area Method

Measures of efficiency

Time spent (seconds): 1.940000

Computer

hilbert.mat.uc.pt x86\_64 unknown  
 Intel(R) Pentium(R) 4 CPU 3.00GHz  
 M. Coren: 2 BogoMIPS: 5669.8  
 OS: 2.6.22-3-686-bigmem GNU/Linux.

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

## Adding New Problems

- a workbench to test/add new problems.

Select a Conjecture from the Database

Problems Listing **CE00000 - New Centenary** Choose a Conjecture

User's Scrapbook

Save/Update	name (max 20 char)	Passwd	Select	Passwd	Delete

Workbench

```

Require Import area_method.

Theorem a16_40 :
  forall A B C D P Q : Point,
  is_midpoint P A B =>
  is_midpoint B B C =>
  is_midpoint C C D =>
  is_midpoint P A C =>
  is_midpoint P B D =>
  is_midpoint P Q D.
Proof.
  apply method.
Qed.
  
```

0 - CoqAreaMethod - 1.0

Interpolate the code

**Theorem - Centroid Theorem**

Proved. Coq Area Method

**Measures of efficiency**

Time spent (seconds) 1.940000

Computer

hilbert.mat.uc.pt	x86_64 unknown
Intel(R) Pentium(R) 4 CPU 3.00GHz	
4 CPU	
2 Cores	
BogoMIPS: 5069.8	
OS: 2.6.22-3-686-bigmem GNU/Linux.	

Platform: Ln 18, Ch 5 | total: Ln 18, Ch 219

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

# Distribution

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



# Distribution

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

- The Database contents, and the ERD.

# Distribution

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

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

# Distribution

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.

## Distribution

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.

# Open Problems

- Common Format.

The extension of the i2g format for the GATPs.

# Open Problems

- Common Format.

The extension of the i2g format for the GATPs.

⇒ a script to apply all the GATPs to all the problems.

# Open Problems

- Common Format.

The extension of the i2g format for the GATPs.

⇒ 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

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.



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

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

# Thank You