

System Description: SystemOnTPTP

Geoff Sutcliffe

School of Information Technology, James Cook University, Townsville, Australia
geoff@cs.jcu.edu.au

SystemOnTPTP is a WWW interface that allows an ATP problem to be easily and quickly submitted in various ways to a range of ATP systems. The interface uses a suite of currently available ATP systems, which are maintained in a database structure. The interface is generated directly from the database, and thus is as current as the database. The interface is a single WWW page, in three parts: the problem specification part, the mode specification part, and the system selection part. A user specifies the problem first, optionally selects systems for use, and then specifies the mode of use.

Figure 1 shows the problem specification part of the WWW page. There are three ways that the problem can be specified. First, a problem in the TPTP library [SS98] can be specified by name. The interface provides a tool for browsing the TPTP problems if desired. Second, a file, containing a problem written in TPTP syntax, on the same computer as the WWW browser can be specified for uploading. The interface provides an option for browsing the local disk and selecting the file. Third, the problem formulae can be provided directly in a text window. Links to example TPTP files are provided to remind the user of the TPTP syntax if required.

Fig. 1. Problem specification

The screenshot shows a Netscape browser window titled "Netscape: System on TPTP". The address bar displays the URL "http://www.cs.jcu.edu.au/cgi-bin/tptp/SystemOnTPTPFormMaker". The page has a yellow background and is titled "System on TPTP". Under the heading "Problem", there are three options for specifying a problem:

- ☒ TPTP Problem Name (e.g., SYN054-1)
CID003-1 [Browse TPTP](#)
- ☐ Local file to upload
Browse...
- ☐ TPTP Formulae (FOF example - [SYN054+1.p](#), CNF example - [SYN054-1.p](#))

A large text input area is provided for the TPTP Formulae option.

Figure 2 shows the start of the system selection part of the interface. There is one line for each system in the suite, indicating the system name and version, a default CPU time limit, the default tptp2X transformations for the system, the default tptp2X format for the system, and the default command line for the system. To select a system the user selects the corresponding tickbox. The default values in the text fields can be modified, if required.

Fig. 2. System selection

ATP System	Time Limit	Transform	Format	Command
<input type="checkbox"/> Bliksem 1.10A	100 seconds	rm_equality:rstfp	bliksem	bliksem %s
<input type="checkbox"/> CoDe 3.00	100 seconds	none	code	code3 -FE -c- -d- -e-
<input type="checkbox"/> E 0.51	100 seconds	rm_equality:rstfp	tptp	eprover --tptp-forma
<input type="checkbox"/> FDP 0.9	100 seconds	none	protein	fdp-casc %s %d
<input type="checkbox"/> Fiesta 2	100 seconds	rm_equality:rstfp	dedam	fiesta-wrapper %s
<input type="checkbox"/> GLiDeS 0.0	100 seconds	none,rm_equality:stf	glides	GLiDeS %s

Figure 3 shows the mode specification part of the interface. The lefthand side contains information and the RecommendSystems button for obtaining system recommendations for the specified problem. System recommendations are generated as follows: ATP problems have been classified into 14 disjoint specialist problem classes (SPCs) according to problem characteristics such as effective order, use of equality, and syntactic form. In a once off analysis phase for each SPC, performance data for ATP systems (not necessarily all in the suite), for some carefully selected TPTP problems in the SPC, is analyzed. Systems that solve a subset of the problems solved by another system are discarded. The remaining systems are recorded as recommended, in order of the number of problems solved in the SPC. Later at run time, when system recommendations for a specified problem (not necessary one of those used in the analysis phase, or not even from the TPTP) are requested, the problem is classified into its SPC and the corresponding system recommendations are returned.

Fig. 3. Mode specification

System Recommendations

System recommendations are free. The recommendations are based on the [systems' results for the TPTP](#). There is no guarantee that they are the best systems for this particular problem.

Note: Submitting your ATP system's latest performance data is the best way to get it recommended.

Solve the Problem at JCU

Solution attempts are limited to users with passwords plus a few users without passwords. If you submit without a valid password your job may be rejected if the machine is too busy.

Password:

Output mode

☐ Quiet
☒ Interesting
☐ Verbose

Parallel mode

☐ Selected
☐ Naive ...
☐ SSCPA ...
☒ Eager SSCPA ...

seconds
All selected systems

systems

Our server does not output results until all tasks are completed. Be patient while the provers do their thing.

The righthand side of the mode specification part of the interface gives information, options, and the submit buttons for using the ATP systems sequentially or in parallel. When a problem is submitted using either of these submit buttons, the ATP systems are executed on a server at James Cook University. Due to resource restrictions, only one public user may submit at a time. A password field is provided that allows privileged users to submit at any time. The Output mode options specify how much output is returned during processing. In Quiet mode only the final result is returned, giving the type of result and the time taken. The result is either that a proof was found, that it was established that no proof can be found, or that the systems gave up trying for an unknown reason. In Interesting mode information about the progress of the submission is returned; see below for an example. In Verbose mode a full trace of the submission is returned, including the problem in TPTP format, the problem after transformation and formatting for the systems, and all standard output produced by the systems. The RunSelectedSystems button sequentially gives the specified problem to each of the systems selected in the system selection part of the interface. For each selected system, the problem is transformed and formatted using the `tptp2X` utility as specified in the system selection. The transformed and formatted problem is given to the system using the specified command line, with a CPU time limit as specified for the system.

The Parallel mode options specify the type of parallelism to use when a problem is submitted using the RunParallel button. All of the parallel modes perform competition parallelism [SS94], i.e., multiple systems are run in parallel on the machine (using UNIX multitasking if there are less available CPUs than systems) and when any one gets a deciding result all of the systems are killed. The differences between the modes are which systems are used and the individual time limits imposed on each system's execution. A limit on the total CPU time that can be taken by the executing systems is specified in the seconds field of the interface. In Naive selected mode all of the systems selected in the system selection part of the interface are run in parallel with equal CPU time limits (the appropriate fraction of the total time limit). In Naive mode the specified number of systems, taken in alphabetical order from the selection list, are run in parallel with equal time limits. In SSCPA mode the system recommendation component is used to get system recommendations for the specified problem. The suite of systems is then checked for versions of the recommended systems, in order of recommendation, until the specified number of systems have been found, or the recommendations are exhausted. The systems are then run in parallel with equal time limits. In Eager SSCPA mode the systems used are the same as for SSCPA mode, but the individual system time limits are calculated by repeatedly dividing the total time limit by two, and allocating the values to the systems in order. In this manner the highest recommended system gets half the total time limit, the next system gets a quarter, and so on, with the last two systems getting equal time limits. The motivations and effects of these parallel modes are discussed in [SS99]. The effectiveness of SSCPA was demonstrated in the CADE-16 ATP System Competition [Sut00].

Figure 4 shows the interesting output from submission of the TPTP problem `CID003-1` in the Eager SSCPA mode. The execution is first transferred from the WWW server onto a SUN workstation where the ATP systems are installed. The system recommendation component is then invoked. The problem is identified as being real

1st order, having some equality, in CNF, and Horn. Five systems are recommended for the SPC: E 0.32, E-SETHEO 99csp, Vampire 0.0, OtterMACE 437, and Gandalf c-1.0d. The suite of systems is then checked for versions of these systems, and it is found that versions of four of them, E 0.51, Vampire 0.0, OtterMACE 437, and Gandalf c-1.0d, are in the suite. The submission required three systems, so E 0.51, Vampire 0.0, and OtterMACE 437 are used. The individual system time limits out of the total specified limit of 300 seconds are then computed, 150 seconds for E 0.51 and 75 seconds each for Vampire 0.0 and OtterMACE 437. The problem is then transformed and formatted for each of the systems, and the systems are run in parallel. E 0.51 finds a proof after 39.8 seconds CPU time, 62.2 seconds wall clock time, at which stage all of the systems are killed.

Fig. 4. Interesting output for CID003-1 in Eager SSCPA mode

```

Moving onto SUN machine atoll.cs.jcu.edu.au
Getting system recommendations for CID003-1
This is a THM_FFO_SEQ_CNF_HRN TPTP problem.
E---0.32 recommended
E-SETHEO---99csp recommended
Vampire---0.0 recommended
OtterMACE---437 recommended
Gandalf---c-1.0d recommended
E---0.51 can be used
Vampire---0.0 can be used
OtterMACE---437 can be used
Gandalf---c-1.0d can be used
E---0.51 will be used
Vampire---0.0 will be used
OtterMACE---437 will be used
E---0.51 is given 150 seconds in -cE mode
Vampire---0.0 is given 75 seconds in -cE mode
OtterMACE---437 is given 75 seconds in -cE mode
E---0.51 has started (PID 6771)
OtterMACE---437 has started (PID 6772)
Vampire---0.0 has started (PID 6773)
E---0.51's input being prepared by ttp2X
OtterMACE---437's input being prepared by ttp2X
Vampire---0.0's input being prepared by ttp2X
RESULT: Theorem for E---0.51 on CID003-1: 39.8 s CPU, 62.2 s WC
Kill E---0.51 (PID 6771)
Kill OtterMACE---437 (PID 6772)
Kill Vampire---0.0 (PID 6773)
OtterMACE---437 has ended (PID 6772)
Vampire---0.0 has ended (PID 6773)
E---0.51 has ended (PID 6771)
RESULT: Theorem for SSCPA---Alpha on CID003-1: 66.4 s CPU, 94 s WC

```

SystemOnTPTP is implemented by perl scripts that generate the WWW interface from the system database, accept the submission from the browser, extract the system recommendations, invoke the ttp2X utility to transform and reformat the problem, and control the execution of the systems. The interface is available at:

<http://www.cs.jcu.edu.au/cgi-bin/ttp/SystemOnTPTPFormMaker>

SystemOnTPTP makes it easy for users to easily and quickly submit a problem in TPTP syntax to an appropriate ATP system. The user is absolved of the responsibilities and chores of selecting systems to use, installing the systems, transforming and formatting the problem for the systems, and controlling their execution. This user friendly environment is particularly appropriate for ATP system

users who want to focus on the problem content rather than the mechanisms of ATP. The interface is not designed for, and is therefore not suitable for, users who wish to submit a batch of problems to a particular ATP system. Such users should obviously install that ATP system on their own computer, which would also allow use of the system's own input format rather than the TPTP format. ATP system developers are invited to submit their systems and performance data for inclusion and use in the interface.

References

- SS94 Suttner C.B., Schumann J. (1994), **Parallel Automated Theorem Proving**, Kanal L., Kumar V., Kitano H., Suttner C., *Parallel Processing for Artificial Intelligence 1*, pp.209-257, Elsevier Science.
- SS98 Sutcliffe G., Suttner C.B. (1998), **The TPTP Problem Library: CNF Release v1.2.1**, *Journal of Automated Reasoning* 21(2), pp.177-203.
- SS99 Sutcliffe G., Seyfang D. (1999), **Smart Selective Competition Parallelism ATP**, Kumar A., Russell I., *Proceedings of the 12th Florida Artificial Intelligence Research Symposium* (Orlando, USA), pp.341-345, AAAI Press.
- Sut00 Sutcliffe G. (To appear), **The CADE-16 ATP System Competition**, *Journal of Automated Reasoning*.