Mathematics in the TPTP Problem Library (v1.2.1)

Geoff Sutcliffe
Department of Computer Science
James Cook University, Australia
Phone: +61 77 814622 Fax: +61-77-814029
Email: geoff@cs.jcu.edu.au

Christian B. Suttner
Institut für Informatik
TU München, Germany
Phone: +49-89-521098 Fax: +49-89-526502
Email: suttner@informatik.tu-muenchen.de

Abstract

The TPTP (Thousands of Problems for Theorem Provers) is a library of problems for automated theorem proving (ATP) systems, for the clause normal form (CNF) of 1st order logic. A large portion of the TPTP is concerned with various mathematical domains. This paper documents the mathematical domains of TPTP v1.2.1, and provides some insight into the performance of state-of-the-art ATP systems on the problems in these domains. For those new to the field this information provides an established starting point for research. For those already involved, this paper tells what has been achieved so far.

1 Introduction

The TPTP (Thousands of Problems for Theorem Provers) [SSY94, SS96] is a library of problems for automated theorem proving (ATP) systems, for the clause normal form (CNF) of 1st order logic. The TPTP is comprehensive and up-to-date, and thus provides an overview of the current application of ATP systems.

A large portion of the TPTP is concerned with various mathematical domains. This paper documents the mathematical domains of TPTP v1.2.1. TPTP v1.2.1 is the last TPTP release containing only CNF problems (TPTP v2.0.0 will contain problems in First Order Form (FOF)), and in this form the TPTP has already provided a link between the ATP and mathematics communities. This paper provides descriptions of the mathematical domains and some insight into the performance of state-of-the-art ATP systems on the problems. For those new to the field this information provides an established starting point for research. For those already involved, this paper tells what has been achieved so far. The references in this paper form a useful bibliography for those areas of mathematics that have been investigated using ATP techniques.

Section 2 provides an introduction to the TPTP. Section 3 gives an overview of each of the mathematical domains in the TPTP, and Section 4 looks at the performance of several ATP systems on the problems in these domains. Section 5 concludes the paper.

2 Introduction to the TPTP

The principal motivation for the TPTP project is to provide a basis for the testing and evaluation of ATP systems. This became necessary, as results being published often do not accurately reflect the capabilities of the ATP system being considered. A common library of problems is necessary for meaningful system evaluations, meaningful system comparisons, repeatability of testing, and the production of statistically significant results. The TPTP is such a library.

The goal for building the TPTP has been to overcome previous drawbacks, and to centralize the burden of problem collection and maintenance. The TPTP tries to address all relevant issues. In particular, the TPTP is easy to discover and obtain, is easy to use, spans a diversity of subject matters, is comprehensive, and is up-to-date. It also provides a simple, unambiguous reference mechanism for ATP problems. The TPTP problems are stored in a specifically designed, easy to understand format. A utility for manipulating the problems, and for converting the problems to other known ATP formats, is provided.

2.1 The TPTP Domain Structure

An attempt has been made to classify the totality of the TPTP problems in a systematic and natural way. The resulting domain scheme reflects the natural hierarchy of scientific domains, as presented in standard subject classification literature. The current classification is based mainly on the Dewey Decimal Classification [Dew89] and the Mathematics Subject Classification [MSC92] used for the Mathematical Reviews by the American Mathematical Society. Five main fields are defined: logic, mathematics, computer science, engineering, and other. Each field contains further subdivisions, called *domains*. Each domain is identified by a three-letter mnemonic. These mnemonics are also part of the TPTP problem naming scheme. The TPTP domains constitute the basic units of the classification. The full classification scheme is shown in Figure 1.

It is evident from Figure 1 that the mathematics field constitutes a significant part of the TPTP. Mathematics has been a traditional area of application for ATP, due to the availability of complete and understandable axiomatizations of various theories in mathematics. These theories are expressed formally in mathematics texts, and can often be easily translated to the logical form required for use in ATP. A classic example is Art Quaife's translation [Qua92a, Qua92b] of parts of Mendelson's text on mathematical logic [Men87].

2.2 Problem Versions and Standard Axiomatizations.

There are often many ways to formulate a problem for presentation to an ATP system. Thus, in the TPTP, there are often alternative presentations of a problem. The alternative presentations are called *versions* of the underlying *abstract problem*. As the problem versions are the objects that ATP systems must deal with, they are referred to simply as problems, and the abstract problems are referred to explicitly as such. Each problem is stored in a separate physical file. The primary reason for different versions of an abstract problem is the use of different axiomatizations. This issue is discussed below. A secondary reason is different formulations of the theorem to be proved.

TPTP				# of Problems abstract/versions/generic
Logic	Combinatory logic		COL	81 / 141 / 0
	Logic calculi		LCL	256 / 278 / 0
	Henkin models		HEN	12 / 64 / 0
Mathematics	Set Theory		SET	562 / 695 / 0
	Graph Theory		GRA	1 / 1 / 0
	Algebra	Boolean Algebra	BOO	19 / 52 / 0
		Robbins Algebra	ROB	27 / 36 / 0
		Left distributive	LDA	14 / 23 / 0
		Lattices	LAT	5 / 10 / 0
		Groups	GRP	193 / 313 / 50
		Rings	RNG	40 / 100 / 0
		General Algebra	ALG	2 / 4 / 0
	Number Theory		NUM	285 / 309 / 2
	Topology		TOP	19 / 24 / 0
	Analysis		ANA	5 / 19 / 0
	Geometry		GEO	78 / 165 / 0
	Category theory		CAT	19 / 58 / 0
Computer science	Computing Theory		COM	4 / 6 / 0
	Planning		PLA	23 / 30 / 0
	Program Verification		PRV	9 / 9 / 0
Engineering	Circuit Design		CID	3 / 4 / 0
	Circuit Verification		CIV	4 / 4 / 0
Other	Syntactic		SYN	340 / 350 / 27
	Puzzles		PUZ	34 / 45 / 3
	Miscellaneous		MSC	9 / 12 / 3

Figure 1: The domain structure of the TPTP.

Commonly, many different axiomatizations of a theory exist. In the TPTP an axiomatization is *standard* if it is complete (in the sense that it captures some closed theory) and it has not had any lemmas added. (Note: A standard axiomatization may be redundant, because some axioms are dependent on others. In general, it is not known whether or not an axiomatization is minimal, and thus the possibility of redundancy in standard axiomatizations must be tolerated.) In the TPTP, standard axiomatizations are kept in separate axiom files, and are included in problems as appropriate. By using different standard axiomatizations of a particular theory, different versions of a problem can be formed. Some of the standard axiomatizations can be specialized by adding extra axioms; these specialization axioms are also stored in separate files.

Within the ATP community, some problems have been created with *non-standard* axiomatizations. A non-standard axiomatization is formed by modifying a standard axiomatization. The standard axiomatization may be *reduced* (i.e., axioms are removed) and the result is an *incomplete* axiomatization, or it may be *augmented* (i.e., lemmas are added) and the result is a *redundant* axiomatization. Non-standard axiomatizations are typically used to find a proof of a theorem (based on the axiomatization) using a particular ATP system. In any 'real' application of an ATP system, a standard axiomatization of the application domain would typically have to be used, at least initially. Thus the use of standard axiomatizations is desirable, because it reflects such 'real' usage. In the TPTP, for each collected problem that uses a non-standard axiomatization, a new version of the problem is created with a standard axiomatization.

The notion of equality is particularly important in the mathematical domains of the TPTP, and some of the axiomatizations are formulated using only the equality relation. These are called *equality* axiomatizations. In the TPTP `equal`/2 is the only predicate used for equality, and is used iff the equality axiomatization is complete. A problem is said to "contain equality" if the `equal`/2 predicate is used at all in the problem, and a problem is said to be "pure equality" if only the `equal`/2 predicate is used in the problem.

2.3 Obtaining the TPTP

The TPTP may be obtained by ftp from Australia:
`ftp.cs.jcu.edu.au:pub/research/tptp-library/*`
or Germany:
`flop.informatik.tu-muenchen.de:pub/tptp-library/*`
or via the World Wide Web (WWW) from Australia:
`http://www.cs.jcu.edu.au/ftp/users/GSutcliffe/TPTP.html`
or Germany:
`http://wwwjessen.informatik.tu-muenchen.de/~tptp/`.

3 Mathematical Domains in the TPTP

The following subsections summarize the mathematical domains in the TPTP. A brief description of the domain is given, followed by a short discussion of the problems and standard axiomatizations in the TPTP, and their characteristics. At the end of each subsection non-ATP references for a general subject introduction, generic ATP references for prominent ATP research in the domain, and appropriate Dewey Decimal Classification (DDC) and Mathematics Subject Classification (MSC) numbers, are given.

At the end of this section there is a table summarizing properties of the domains' problems, that are important from an ATP perspective.

3.1 Algebra

An algebra is a set with a system of operations defined on it.

The Algebra domain is one of the smallest of the TPTP's mathematical domains. The first of the two abstract problems is a well known one in ATP, which is to prove that the composition of homomorphisms is a homomorphism [BLM+86, Qua92a, Qua92b].

The domain contains one standard axiomatization, based on Neumann-Bernays-Gödel Set Theory axioms [BLM+86] (see Section 3.13).

The problems in this domain are all non-Horn.

Indices : DDC 512; MSC 06XX, 20XX.
References : General [Bou89, BM65, BB70]; ATP –.

3.2 Analysis

Analysis is concerned with functions and limits. The main parts of analysis are differential calculus, integral calculus, and the theory of functions.

The Analysis domain contains problems that were proposed as challenges to ATP systems by Woody Bledsoe [Ble90]. They include the Intermediate Value Theorem and the problem to prove that the sum of two continuous functions is continuous.

The domain contains two standard axiomatizations. Both are in relational form, taken from [MOW76] and [Ble90].

This domain contains a mix of problems that are Horn, non-Horn, with equality, and without equality.

Indices : DDC 515; MSC 26XX.
References : General [Ros90]; ATP [Ble90].

3.3 Boolean Algebra

A Boolean algebra is a set of elements with two binary operations which are idempotent, commutative, and associative. These operations are mutually distributive, there exist universal bounds 0, 1, and there is a unary operation of complementation.

Most of the abstract problems in the Boolean Algebra domain were first brought to the attention of ATP research by McCharen, et al. [MOW76]. Bob Veroff later reformulated the problems using a pure equality axiomatization, and added some new problems [Ver94]. The problems prove identities in Boolean Algebra.

The domain contains four standard axiomatizations. The first is an equality axiomatization distributed with the Otter ATP system [McC94]. This axiomatization also appears in [Win82] where one of the axioms is shown to be dependent. The axiomatization further appears in [Wos88]. The second axiomatization is from [MOW76], the third is an equality axiomatization from the Argonne National Library [ANL], and the last is another equality axiomatization due to Bob Veroff [Ver94].

These problems are all Horn problems containing equality.

Indices : DDC 511.324, 512.89; MSC 06EXX.
References : General [Whi61, BM65, BB70]; ATP –.

3.4 Category Theory

A category is a structure together with the morphisms that preserve this structure.

The problems in the Category Theory domain were formalized by the ATP researchers at Argonne National Laboratory, and were taken from Chapter 1 of [Mit67]. They consider various morphisms, and relationships between the domain and codomain of a mapping.

The domain contains four standard axiomatizations, all from [ANL]. The first is based on [Mit67]. The second was proposed to Larry Wos by Art Quaife [Qua89a]. The third and fourth axiomatizations are based on the work of Scott [Sco79]. The third is a version of Scott's axioms as simplified by Art Quaife, and the fourth is a version with some redundant axioms removed.

Most of these problems contain equality.

Indices : DDC 512.55; MSC 18XX.
References : General [Mac71]; ATP [MOW76].

3.5 Geometry

Geometry deals with the measurement, properties, and relationships of points, lines, angles, surfaces, and solids. In the TPTP the Geometry domain deals mainly with plane geometry, based on Tarski's axiom system for Euclidean geometry.

The problems in the Geometry domain come mostly from [MOW76] and from Art Quaife's work [Qua89b]. These problems consider properties of points and lines, such as equidistance, betweeness, and colinearity. Better known examples of these are the Inner five-segment theorems and proof of the Outer Pasch Axiom. There are also some problems based on Hilbert geometry, from [Ben92].

The domain contains three standard axiomatizations. The first is a relational axiomatization from [MOW76], based on Tarski's axiom system for Euclidean geometry

[Tar59]. This axiomatization has a specialization for colinearity. The second axiomatization is by Art Quaife [Qua89b]. This axiomatization also has a specialization for colinearity, and also specializations for reflection and insertion. The differences between the [MOW76] and the [Qua89b] axiomatizations are explained in [Qua89b]. The third axiomatization is by Dan Benanav, based on Hilbert geometry [Ben92].

The problems are almost all non-Horn with equality.

Indices : DDC 516; MSC 51.
References : General [Tar51, Tar59]; ATP [Qua92b].

3.6 Graph Theory

A graph consists of a finite non-empty set of points together with a prescribed set of pairs of points.

Graph Theory is the smallest of the TPTP domains. A particular difficulty of dealing with graph theory is that many proofs in graph theory rely on induction, which cannot easily be encoded in 1st order logic. The single problem here is based on an observation about graphs made by Tseitin [Tse83] and was presented by Jeff Pelletier [Pel86].

This domain contains no standard axiomatizations.

The single problem in this domain is propositional, and is the only propositional problem in the mathematical domains of the TPTP.

Indices : DDC 510.09; MSC 05CXX, 68R10.
References : General [Har69, BB70]; ATP –.

3.7 Group Theory

A group is a system of elements which is closed under a single-valued binary operation which is associative, and relative to which there exists an element satisfying the identity law, and with each element another element (called its inverse) satisfying the inverse law.

Group Theory is one of the larger and more developed mathematical domains in the TPTP, and in ATP research in general. The abstract problems in this domain fall mainly into four areas: problems that prove basic properties of groups, problems that prove the dependence or sufficiency of group theory axioms, problems that establish the existence of certain groups, and problems dealing with lattice ordered groups. Of particular interest are the problems regarding quasigroups, investigated by Fujita, Slaney, et al. [FSB93, SFS95]. This research answered some open questions about the existence of certain quasigroups, posed by the mathematician Frank Bennett [Ben89, Ben92].

The domain contains seven standard axiomatizations. The first three are all relational, taken from the axioms presented in [MOW76]. They are axiomatizations for monoids, semigroups, and groups. The latter has a specialization for subgroups, and was also used in [Wos88]. The fourth axiomatization is also from [MOW76], but is an equality axiomatization. It has specializations for subgroups and for lattice ordered groups, the latter developed by Matthias Fuchs [Fuc94]. An extended version of this axiomatization appears in [Wos88]. The fifth axiomatization is a subset of the [MOW76] axioms but is still a complete axiomatization. The subset was selected by Bob Veroff [Ver92]. The last two axiomatizations deal with multiple groups and semi-groups, thus making it necessary to identify the group in question by an extra arguments in each of the predicates. The sixth axiomatization is for groups, from [MOW76], and the seventh is for semigroups, developed by Christoph Goller [Gol93]. As well as the standard TPTP axiomatizations, there are problems in this domain that establish the completeness of some single equality axioms for Group Theory [Kun92, McC93].

Group Theory is the best source of range restricted problems from the mathematical domains of the TPTP, as required by model generation ATP systems such as SATCHMO [MB88].

Indices : DDC 512.2; MSC 20
References : General [Bou89, BM65]; ATP [MOW76].

3.8 Lattice Theory

A lattice is a set of elements, with two binary operations which are idempotent, commutative, and associative, and which satisfy the absorption law.

Although a small mathematical domain in the TPTP, Lattice Theory played a particularly important role in ATP history, in the form of "SAM's lemma". The problems here were formulated by Bill McCune [McC88], based on problems that appear in [GOBL69]. [GOBL69] focuses on an interactive theorem prover SAM, that was used to help solve an open problem in Lattice Theory posed by Bumcroft [Bum65]. This problem is one of the problems in this TPTP domain.

The domain contains two standard axiomatizations. The first is an equality axiomatization from [McC88], and also appears in [Wos88] with four extra redundant axioms. It has specializations for complement and unique cimplement. The second axiomatization is from [MOW76].

Indices : DDC 512.865; MSC 06BXX.
References : General [BM65]; ATP [McC88].

3.9 Left Distributive Algebra

LD-algebras are related to large cardinals. Under a very strong large cardinal assumption, the free-monogenic LD-algebra can be represented by an algebra of elementary embeddings. Theorems about this algebra can be proved from a small number of properties, suggesting the definition of an embedding algebra.

The Left Distributive Algebra domain was developed for ATP by the mathematician Thomas Jech [Jec93]. Jech investigated this domain using the ATP system Otter [McC94], and posed several open questions in [Jec93].

The domain contains one standard axiomatization, from [Jec93].

All the Left Distributive Algebra problems involve equality.

Indices : DDC 512; MSC –.
References : General –, ATP [Jec93].

3.10 Number Theory

Number theory is the study of integers and their properties.

With the exception of a few older problems obtained from early ATP works [Cha70, LS74, WM76], the problems in the Number Theory domain are taken from [BLM+86] and [Qua92a, Qua92b]. The abstract problems deal with a wide range of Number Theory, including proofs of Peano's axioms, ordering properties, ordinals, transfinite induction, and transfinite recursion. A commonly known open problem in this domain is the Goldbach conjecture.

The domain contains four standard axiomatizations. The first two are from [LS74], and are relational and equality axiomatizations respectively. The first has specializations for less and division. The third and fourth axiomatizations are based on encodings of Gödel's treatment of the Neumann-Bernays-Gödel Set Theory axioms (see Section 3.13), taken from [BLM+86] and [Qua92a, Qua92b] repectively.

The problems are mostly large (in terms of the number of clauses), non-Horn, and contain equality. They are hard challenges for ATP systems.

Indices : DDC 512.7; MSC 11YXX.
References : General [HW92]; ATP [Qua92b].

3.11 Ring Theory

A ring is a system of elements which is an Abelian group under an operation of addition, and which is closed under a binary operation of multiplication, the latter being associative, and distributive with respect to addition.

The problems in the Ring Theory domain fall mainly into two groups. Problems from [LS74, WM76, MOW76] are formulated in a relational style, while the problems from the more recent work of Rick Stevens [Ste87, Ste88, Ste92] are formulated in a pure equality functional style. A particularly difficulty group of problems in this domain deals with proofs of the Moufang identities in alternative rings.

The domain contains five standard axiomatizations. The first is from [MOW76], in relational style. The second is a complete set of reductions for ring theory, determined by Peterson and Stickel [PS81]. The third and fourth axiomatizations are equality axiomatizations for alternative ring theory. The third is by Stevens [Ste87] and the fourth by Anantharaman and Hsiang [AH90]. The fifth axiomatization is an equality axiomatization by Lusk and Wos [LW91].

The problems are mostly Horn with equality.

Indices : DDC 512.4; MSC 13XX, 16XX.
References : General [Bou89, BB70]; ATP [MOW76].

3.12 Robbins Algebra

The Robbin Algebra domain revolves around the question "Is every Robbins algebra Boolean?". Most of the problems in this domain identify conditions that make a near-Boolean algebra Boolean.

Steve Winker has been dominant in applying ATP techniques in Robbins Algebra [Win90]. His work builds on research by Henkin et al. [HMT71].

The domain contains one standard axiomatization, an equality axiomatization from [Win90]. There is a specialization for integer arithmetic.

The problems in this domain are all Horn, and almost all contain equality.

Indices : DDC 512; MSC 03G15.
References : General [HMT71]; ATP [Win90].

3.13 Set Theory

Classically, a set is a totality of certain definite, distinguishable objects of our intuition or thought - called the elements of the set. Due to paradoxes that arise from such a naive definition, mathematicians now regard the notion of a set as an undefined, primitive concept [How72]. In this domain, Naive and Neumann-Bernays-Gödel axiom sets are used.

Set Theory is the largest mathematical domain in the TPTP. A large corpus of problems was provided by Art Quaife [Qua92a, Qua92b], who extended earlier work by Bledsoe et al. [BLM+86]. There are also some earlier problems from [LS74] and [MOW76]. The abstract problems deal with problems ranging from elementary Set Theory, e.g., showing that the intersection of sets is commutative, through to advanced concepts, e.g., Cantor's theorem that there can be no function mapping a set onto its power set.

The domain contains four standard axiomatizations. The first is an older axiomatization taken from [LS74], which has specializations for set membership, union, intersection, and difference. The second axiomatization is from [MOW76]. The third and fourth axiomatizations are encodings of Gödel's treatment of the Neumann-Bernays-Gödel Set Theory axioms [Göd40]. The third is by Bledsoe et al. [BLM+86], and the fourth is by Quaife [Qua92a, Qua92b]. Quaife's axiomatization has a specialization for Boolean algebra.

The problems based on Neumann-Bernays-Gödel Set Theory are all large, non-Horn, and contain equality. They, as with the large Number Theory problems, are hard challenges for ATP systems.

Indices : DDC 511.322, 512.817; MSC 03EXX, 04XX.
References : General [Neu25, Qui69]; ATP [Qua92b].

3.14 Topology

Topology is the study of properties of geometric configurations (e.g., point sets) which are unaltered by elastic deformations (homeomorphisms, i.e., functions that are 1-1 mappings between sets such that both the function and its inverse are continuous).

The axiomatization of Topology used here encodes elementary point-set topology [WM89]. However, there are believed to be some deficiencies in the axiomatization; firstly the axiomatization appears to be inconsistent, and secondly it may be necessary to add appropriate set theory axioms to obtain proofs of some of the conjectures. Very few proofs have been found of the conjectures in this domain.

The problems are all non-Horn, but without equality.

Indices : DDC 514; MSC 46AXX.
References : General [Kel55, Mun75]; ATP [WM89].

3.15 Domain Properties

Table 1 provides a synopsis of the problem properties for each domain. The table lists the number of abstract problems, the number of problems based on those abstract problems, the number of satisfiable problems (i.e., non-theorems), the number of standard problems (as discussed in Section 2.2), the number of non-Horn problems, the number of range restricted problems, the number of problems that contain equality, and the number of problems that contain only equality.

Dom	Ab	Pr	Sat		Std		non-Horn		R.R.		With Eq		Pure Eq	
ALG	2	4	0	(0%)	3	(75%)	4	(100%)	0	(0%)	3	(75%)	0	(0%)
ANA	5	19	0	(0%)	11	(58%)	13	(68%)	0	(0%)	6	(32%)	0	(0%)
BOO	19	52	2	(4%)	44	(85%)	0	(0%)	0	(0%)	52	(100%)	33	(63%)
CAT	19	58	2	(3%)	58	(100%)	18	(31%)	0	(0%)	57	(98%)	7	(12%)
GEO	78	165	0	(0%)	92	(56%)	162	(98%)	1	(1%)	162	(98%)	0	(0%)
GRA	1	1	0	(0%)	0	(0%)	1	(1%)	1	(1%)	0	(0%)	0	(0%)
GRP	193	313	1	(0%)	217	(69%)	75	(24%)	22	(7%)	238	(76%)	192	(61%)
LAT	5	10	0	(0%)	8	(80%)	2	(20%)	0	(0%)	8	(80%)	1	(10%)
LDA	14	23	0	(0%)	4	(17%)	18	(78%)	0	(0%)	23	(100%)	3	(13%)
NUM	285	309	1	(0%)	273	(88%)	295	(95%)	2	(1%)	287	(93%)	0	(0%)
RNG	40	100	0	(0%)	54	(54%)	3	(3%)	0	(0%)	91	(91%)	70	(70%)
ROB	27	36	0	(0%)	29	(81%)	0	(0%)	0	(0%)	35	(97%)	22	(61%)
SET	562	695	0	(0%)	583	(84%)	695	(100%)	0	(0%)	670	(96%)	0	(0%)
TOP	19	24	8	(33%)	0	(0%)	24	(100%)	0	(0%)	0	(0%)	0	(0%)

Table 1: Properties of the Domains

4 System Performances

Since the first release of the TPTP problem library in 1993, many researchers have used the TPTP as an appropriate and convenient basis for ATP system evaluation. Some researchers, who have tested their ATP systems over the entire TPTP, have made their result data available [SS95]. The more powerful general purpose ATP systems for which data is currently available, are CLIN-S 1.0 [CP94], LINUS 1.0.1 [Let95], Otter 3.0.4 [McC94], SETHEO 3.3 [LSBB92, LMG94, GLMS94], SNARK 960611 [Sti94], SPASS 0.45 [WGR96], and SPTHEO [Sut95]. Table 2 gives the performances of these ATP systems on the mathematical domains of the TPTP[1]. For each domain the table shows performances of the systems in order of number of problems solved. The number of problems solved, and that number relative to the highest number of problems solved, is given. If a system name is emboldened, that means that no other system solves a strict superset of the the problems solved by that system, i.e., the emboldened systems in combination solve all problems in the domain that have been solved by any of the ATP systems. Conversely, the problems solved by each non-emboldened system are a strict subset of those solved by one of the emboldened systems. The number of problems in each domain is given in ()s below the domain mnemmonic.

Several observations may be made from the figures in Table 2.

- For Horn problems with equality, as found in the Boolean Algebra, Category Theory, Group Theory, Ring Theory, and Robbins Algebra domains, Otter 3.0.4 and SNARK 960611 are the best performing systems.
- For non-Horn problems with equality, as found in the Geometry, Number Theory, and Set Theory domains, SNARK 960611, SPASS 0.45, and Otter 3.0.4 are the best performing systems.
- For non-Horn problems without equality, as found in the Topology domain, SPASS 0.45 is the best performing system.
- For the Number Theory and Set Theory domains, where there are very large problems (in terms of the number of clauses in the problem), SPTHEO m256n1 performs well.
- There are several domains, e.g., Geometry, Number Theory, Ring Theory, Set Theory, and Topology, where a variety of ATP techniques appear to be necessary to solve problems.
- There are some domains, e.g., Boolean Algebra and Robbin Algebra, where only a few techniques appear sufficient to solve problems.

5 Conclusion

The mathematical domains of the TPTP form an extensive and useful starting point for the investigation of mathematical theories using ATP techniques. Several ATP researchers and mathematicians have already contributed original results, some answering open questions in mathematics, to the TPTP. There are currently 9 open problems in the TPTP mathematical domains, which provide real challenges to the

[1]Some of the ATP systems have several variants formed from using different search strategies. Here the data for the variant with the best overall performance is given.

Dom	1st		2nd		3rd		4th		5th		6th		7th	
ALG	**Otter**		**SNARK**		CLIN-S		LINUS		SETHEO		SPASS		SPTHEO	
(4)	2	1.00	2	1.00	1	0.50	1	0.50	1	0.50	1	0.50	1	0.50
ANA	**CLIN-S**		**SETHEO**		SPASS		SPTHEO		LINUS		Otter		SNARK	
(19)	3	1.00	3	1.00	2	0.67	2	0.67	0	0.00	0	0.00	0	0.00
BOO	**Otter**		SNARK		**SPASS**		SETHEO		LINUS		SPTHEO		CLIN-S	
(52)	49	1.00	43	0.88	33	0.67	20	0.41	15	0.31	13	0.27	5	0.10
CAT	**Otter**		**SNARK**		**LINUS**		**SPASS**		**SPTHEO**		SETHEO		CLIN-S	
(58)	45	1.00	42	0.93	41	0.91	38	0.84	32	0.71	31	0.69	29	0.64
GEO	**SNARK**		**SPASS**		**Otter**		**SPTHEO**		**SETHEO**		**CLIN-S**		LINUS	
(165)	88	1.00	83	0.94	73	0.83	61	0.69	60	0.68	57	0.65	49	0.56
GRA	**CLIN-S**		**LINUS**		**Otter**		**SETHEO**		**SNARK**		**SPASS**		SPTHEO	
(1)	1	1.00	1	1.00	1	1.00	1	1.00	1	1.00	1	1.00	1	1.00
GRP	**SNARK**		**Otter**		**SPASS**		SETHEO		LINUS		CLIN-S		SPTHEO	
(313)	219	1.00	216	0.99	207	0.95	135	0.62	51	0.23	49	0.22	48	0.22
LAT	**Otter**		**SNARK**		CLIN-S		LINUS		SETHEO		SPASS		SPTHEO	
(10)	4	1.00	4	1.00	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00
LDA	**Otter**		**SPASS**		SNARK		LINUS		SPTHEO		SETHEO		CLIN-S	
(23)	10	1.00	7	0.70	6	0.60	3	0.30	2	0.20	1	0.10	0	0.00
NUM	**SPTHEO**		**Otter**		**SETHEO**		**SNARK**		**SPASS**		**LINUS**		**CLIN-S**	
(309)	26	1.00	25	0.96	25	0.96	24	0.92	23	0.88	23	0.88	21	0.81
RNG	**Otter**		**SNARK**		**LINUS**		**SETHEO**		**SPASS**		**SPTHEO**		CLIN-S	
(100)	54	1.00	39	0.72	25	0.46	24	0.44	22	0.41	19	0.35	6	0.11
ROB	**Otter**		SNARK		SPASS		LINUS		SPTHEO		SETHEO		CLIN-S	
(36)	14	1.00	10	0.71	10	0.71	5	0.36	5	0.36	4	0.29	1	0.07
SET	**SPTHEO**		**SPASS**		**Otter**		**SNARK**		**SETHEO**		**LINUS**		**CLIN-S**	
(695)	161	1.00	159	0.99	132	0.82	130	0.81	127	0.79	120	0.75	95	0.59
TOP	**SPASS**		**CLIN-S**		**Otter**		**SNARK**		**LINUS**		**SPTHEO**		**SETHEO**	
(24)	12	1.00	5	0.42	5	0.42	5	0.42	5	0.42	5	0.42	5	0.42

Table 2: System Performances

ATP community. The modular format of the TPTP makes it easy to formulate new conjectures based on the existing axiomatizations in the TPTP.

In the future the TPTP will be extended to cover further mathematical domains; in the immediate future the theory of ordered fields will be added. Contributions to the TPTP are encouraged.

Acknowledgement. Thanks to Thilo Trapp for extracting some data from the TPTP, for this paper.

References

[AH90] S. Anantharaman and J. Hsiang. Automated Proofs of the Moufang Identities in Alternative Rings. *Journal of Automated Reasoning*, 6(1):79–110, 1990.

[ANL] ANL. Argonne National Laboratory Problem Library. Available by anonymous ftp from info.msc.anl.gov, Maths and Computer Science Division, Argonne National Laboratory, Argonne, IL.

[BB70] G. Birkhoff and T. Bartee. *Modern Applied Algebra*. McGraw-Hill, 1970.

[Ben89] F.E. Bennett. Quasigroup Identities and Mendelsohn Designs. *Canadian Journal of Mathematics*, 41(2):341–368, 1989.

[Ben92] D. Benanav. Recognising Unnecessary Clauses in Resolution Based Systems. *Journal of Automated Reasoning*, 9(1):43–76, 1992.

[Ble90] W.W. Bledsoe. Challenge Problems in Elementary Calculus. *Journal of Automated Reasoning*, 6:341–359, 1990.

[BLM+86] R. Boyer, E. Lusk, W.W. McCune, R. Overbeek, Stickel M., and L. Wos. Set Theory in First-Order Logic: Clauses for Godel's Axioms. *Journal of Automated Reasoning*, 2(3):287–327, 1986.

[BM65] G. Birkhoff and S. MacLane. *A Survey of Modern Algebra*. Macmillan, 1965.

[Bou89] N. Bourbaki. *Algebra I – Chapters 1-3*. Springer Verlag, 1989.

[Bum65] R. Bumcroft. In *Proceedings of the Glasgow Mathematical Association 7(1)*, 1965.

[Cha70] C-L. Chang. The Unit Proof and the Input Proof in Theorem Proving. *Journal of the ACM*, 17(4):698–707, 1970.

[CP94] H. Chu and D. Plaisted. Semantically Guided First-order Theorem Proving using Hyper-linking. In A. Bundy, editor, *Proceedings of the 12th International Conference on Automated Deduction*, number 814 in Lecture Notes in Artificial Intelligence, pages 192–206. Springer-Verlag, 1994.

[Dew89] M. Dewey. *Dewey Decimal Classification and Relative Index.* Forest Press, 20 edition, 1989.

[FSB93] M. Fujita, J. Slaney, and F. Bennett. Automatic Generation of Some Results in Finite Algebra. In R. Bajcsy, editor, *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, pages 52–57. Morgan Kaufman, 1993.

[Fuc94] M. Fuchs. The Application of Goal-Orientated Heuristics for Proving Equational Theorems via the Unfailing Knuth-Bendix Completion Procedure. A case study: lattice ordered groups. Technical Report SEKI-Report SR-94-02, Universität Kaiserslautern, Kaiserslautern, Germany, 1994.

[GLMS94] C. Goller, R. Letz, K. Mayr, and J. Schumann. SETHEO V3.2: Recent Developments. In A. Bundy, editor, *Proceedings of the 12th International Conference on Automated Deduction*, number 814 in Lecture Notes in Artificial Intelligence, pages 778–782. Springer-Verlag, 1994.

[GOBL69] J.R. Guard, F.C. Oglesby, J.H. Bennett, and Settle L.G. Semi-Automated Mathematics. *Journal of the ACM*, 16(1):49–62, 1969.

[Göd40] K. Gödel. *The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory.* Princeton University Press, 1940.

[Gol93] C. Goller. Anwendung des Theorembeweisers SETHEO auf die Theorie der Halbgruppen und Gruppen. Technical Report Technical Report FKI-174-93, Technische Universität München, München, Germany, 1993.

[Har69] F. Harary. *Graph Theory.* Addison-Wesley, 1969.

[HMT71] L. Henkin, J. Monk, and A. Tarski. *Cylindrical Algebras*, volume Part 1. North-Holland, 1971.

[How72] A.G. Howson. *A Handbook of Terms used in Algebra and Analysis.* Cambrige University Press, 1972.

[HW92] G.H. Hardy and E.M. Wright. *An Introduction to the Theory of Numbers.* Oxford UP, 5 edition, 1992.

[Jec93] T. Jech. LD-algebras. *Association for Automated Reasoning Newsletter*, 22:9–12, 1993.

[Kel55] J.L. Kelley. *General Topology.* D. Van Nostrand, 1955.

[Kun92] K. Kunen. Single Axioms for Groups. *Journal of Automated Reasoning*, 9(3):291–308, 1992.

[Let95] R. Letz. LINUS: A Link Instantiation Prover with Unit Support. In *Proceedings of the 2nd Seminar on Deduction*, 1995.

[LMG94] R. Letz, K. Mayr, and C. Goller. Controlled Integration of the Cut Rule into Connection Tableau Calculi. *Journal of Automated Reasoning*, 13(3):297–337, 1994.

[LS74] J.D. Lawrence and J.D. Starkey. Experimental tests of resolution based theorem-proving strategies. Technical Report Technical Report, Computer Science Department, Washington State University, Pullman, USA, 1974.

[LSBB92] R. Letz, J. Schumann, S. Bayerl, and W. Bibel. SETHEO: A High-Performance Theorem Prover. *Journal of Automated Reasoning*, 8(2):183–212, 1992.

[LW91] E. Lusk and L. Wos. Benchmark Problems in Which Equality Plays the Major Role. Technical Report Technical Report MCS-P275-1191, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, USA, 1991.

[Mac71] S. MacLane. *Categories for the Working Mathematician.* Springer Verlag, 1971.

[MB88] R. Manthey and F. Bry. SATCHMO: a theorem prover implemented in Prolog. In Overbeek R. Lusk E., editor, *Proceedings of the 9th International Conference on Automated Deduction* , number 310 in Lecture Notes in Computer Science, pages 415–434. Springer-Verlag, 1988.

[McC88] W. McCune. Challenge Equality Problems in Lattice Theory. In *Proceedings of the 9th International Conference on Automated Deduction - Argonne Ilinois, USA, May 1988*, pages 704 – 709, Berlin, 1988. Springer.

[McC93] W.W. McCune. Single Axioms for Groups and Abelian Groups with Various Operations. *Journal of Automated Reasoning*, 10(1):1–13, 1993.

[McC94] W.W. McCune. Otter 3.0 Reference Manual and Guide. Technical Report Technical Report ANL-94/6, Argonne National Laboratory, Argonne, USA, 1994.

[Men87] E. Mendelson. *Introduction to Mathematical Logic.* Wadsworth and Brooks/Cole, 3 edition, 1987.

[Mit67] B. Mitchell. *Theory of categories*, volume 17. Academic Press, 1967.

[MOW76] J.D. McCharen, R.A. Overbeek, and L.A. Wos. Problems and Experiments for and with Automated Theorem-Proving Programs. *IEEE Transactions on Computers*, C-25(8):773–782, 1976.

[MSC92] *Mathematical Subject Classification.* American Mathematical Society, Provedence, USA, 1992.

[Mun75] J.R. Munkres. *Topology: A First Course.* Prentice-Hall, 1975.

[Neu25] J. von Neumann. Eine Axiomatisierung der Mengenlehre. *Journal für die reine und angewandte Mathematik*, 154:219–240, 1925.

[Pel86] F.J. Pelletier. Seventy-five Problems for Testing Automatic Theorem Provers. *Journal of Automated Reasoning*, 2(2):191–216, 1986.

[PS81] G. Peterson and M.E. Stickel. Complete Sets of Reductions for Some Equational Theories. *Journal of the ACM*, 28(8):233–264, 1981.

[Qua89a] A. Quaife. 1989.

[Qua89b] A. Quaife. Automated Development of Tarski's Geometry. *Journal of Automated Reasoning*, 5(1):97–118, 1989.

[Qua92a] A. Quaife. Automated Deduction in von Neumann-Bernays-Godel Set Theory. *Journal of Automated Reasoning*, 8(1):91–147, 1992.

[Qua92b] A. Quaife. *Automated Development of Fundamental Mathematical Theories.* Kluwer Academic Publishers, 1992.

[Qui69] W.V. Quine. *Set Theory and its Logic.* Harvard University Press, 1969.

[Ros90] K.A. Ross. *Elementary Analysis: The Theory of Calculus.* Springer, 2nd edition, 1990.

[Sco79] D. Scott. Identity and Existence in Intuitionist Logic. In Fournam M.P., editor, *Durham Proceedings Applications of Sheaves*, pages 660–696. Lecture Notes in Mathematics Volume 753, Springer, 1979.

[SFS95] J. Slaney, M. Fujita, and M. Stickel. Automated Reasoning and Exhaustive Search: Quasigroup Existence Problems. *Computers and Mathematics with Applications*, 29(2):115–132, 1995.

[SS95] G. Sutcliffe and C.B. Suttner. ATP System Results for the TPTP Problem Library (upto TPTP v1.1.3), 1995. Technical Report 95/16, Department of Computer Science, James Cook University, Townsville, Australia; Technical Report AR-95-06, Institut für Informatik, Technische Universität München, Munich, Germany.

[SS96] C.B. Suttner and G. Sutcliffe. The TPTP Problem Library (TPTP v1.2.1), 1996. Technical Report AR-96-02, Institut für Informatik, Technische Universität München, Munich, Germany; Technical Report 96/09, Department of Computer Science, James Cook University, Townsville, Australia.

[SSY94] G. Sutcliffe, C.B. Suttner, and T. Yemenis. The TPTP Problem Library. In A. Bundy, editor, *Proceedings of the 12th International Conference on Automated Deduction*, number 814 in Lecture Notes in Artificial Intelligence, pages 252–266. Springer-Verlag, 1994.

[Ste87] R.L. Stevens. Some Experiments in Nonassociative Ring Theory with an Automated Theorem Prover. *Journal of Automated Reasoning*, 3(2):211–221, 1987.

[Ste88] R.L. Stevens. Challenge Problems from Nonassociative Rings for Theorem Provers. In Overbeek R. Lusk E., editor, *Proceedings of the 9th International Conference on Automated Deduction* , number 310 in Lecture Notes in Computer Science, pages 730–734. Springer-Verlag, 1988.

[Ste92] R.L. Stevens. Unpublished note. 1992.

[Sti94] M.E. Stickel. The Deductive Composition of Astronomical Softwared from Subroutine Libraries. In A. Bundy, editor, *Proceedings of the 12th International Conference on Automated Deduction*, number 814 in Lecture Notes in Artificial Intelligence, pages 341–355. Springer-Verlag, 1994.

[Sut95] C.B. Suttner. *Parallelization of Search-based Systems by Static Partitioning with Slackness*. PhD thesis, Institut für Informatik, Technische Universität München, Munich, Germany, 1995. and Dissertationen zur künstlichen Intelligenz 101, Infix-Verlag.

[Tar51] A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. University of California Press, 2ed edition, 1951.

[Tar59] A. Tarski. What is Elementary Geometry? In L. Henkin, editor, *Proceedings of an International Symposium. The axiomatic method with special reference to geometry and physics*. North-Holland, 1959.

[Tse83] G.S. Tseitin. On the Complexity of Derivations in Propositional Calculus. In J. Siekmann and G. Wrightson, editors, *The Automation of Reasoning*. Springer-Verlag, 1983.

[Ver92] R. Veroff. Group theory problems. 1992.

[Ver94] R. Veroff. Problem Set. *AAR Newsletter*, 24:8, 1994.

[WGR96] C. Weidenbach, B. Gaede, and G. Rock. SPASS & FLOTTER. In M. McRobbie and J. Slaney, editors, *Proceedings of the 13th International Conference on Automated Deduction*, number 1104 in Lecture Notes in Artificial Intelligence, pages 141–145. Springer-Verlag, 1996.

[Whi61] J.E. Whitesitt. *Boolean Algebra and Its Applications*. Addison-Wesley, 1961.

[Win82] S. Winker. Generation and Verification of Finite Models and Counterexamples Using an Automated Theorem Prover Answering Two Open Questions. *Journal of the ACM*, 29(2):273–284, 1982.

[Win90] S. Winker. Robbins Algebra: Conditions that make a Near-Boolean Algebra Boolean. *Journal of Automated Reasoning*, 6(4):465–489, 1990.

[WM76] G.A. Wilson and J. Minker. Resolution Strategies: A Comparative Study. *IEEE Transactions on Computers*, C-25(8):782–801, 1976.

[WM89] C. Wick and W. McCune. Automated Reasoning about Elementary Point-Set Topology. *Journal of Automated Reasoning*, 5:239 – 255, 1989.

[Wos88] L. Wos. *Automated Reasoning - 33 Basic Research Problems*. Prentice-Hall, 1988.