Report on the Termination Competition 2008

Simon Bailey (Platform), Frederic Blanqui (Certification), Jürgen Giesl (LP/FP/SRS/TRS), Georg Moser (Complexity), Johannes Waldmann (editor).

Termination Competition

Termination Analyzers are applied to Termination Problems, results are presented on the web.

- encourage research and implementation efforts
- allow to measure progress
- show that implementations are mature, reliable and ready to be used in applications

Competitions since 2003, organized by Albert Rubio, Claude Marché, Hans Zantema.

since 2008: hosted by Computational Logic Group at U Innsbruck, chair: Aart Middeldorp

What’s new in 2008?

- complexity analysis
- more emphasis on certification
- automated submission and test runs of provers

What’s new in 2008?

- complexity analysis
- more emphasis on certification
- automated submission and test runs of provers

ongoing work (2009 . . .)

- flexible query interface
- flexible execution service

Termination Problem Semantics

- input:
  - (an effective description of a) binary relation $\rightarrow$ (a step of a computation)
- questions:
  - termination: $\rightarrow$ is well-founded
  - derivational complexity: length of $\rightarrow$-chains
- answers (yes, no, lower/upper bounds)
  - for human inspection
  - for automated verification

What Computations?

- functional program (Haskell)
- logic program (Prolog)

- rewrite system
  - unary/arbitrary signature (SRS/TRS)
  - strategy (none, inner/outermost, contextsens.)
  - start terms (all, constructor-based)
  - relative/equational: $\longrightarrow_1 \circ \longrightarrow_2^*$, $\longrightarrow_1 \circ \longleftrightarrow_2^*$
    instances: $\longleftrightarrow_2^*$ for A, C, I

What Computations?

- functional program (Haskell)
- logic program (Prolog)
- coming: imperative programs (Java Bytecode)
- rewrite system
    - unary/arbitrary signature (SRS/TRS)
    - strategy (none, inner/outermost, contextsens.)
    - start terms (all, constructor-based)
    - relative/equational: $\longrightarrow_1 \circ \longrightarrow_2^*$, $\longrightarrow_1 \circ \longleftrightarrow_2^*$ instances: $\longleftrightarrow_2^*$ for A, C, I

What Computations?

- functional program (Haskell)
- logic program (Prolog)
- coming: imperative programs (Java Bytecode)
- rewrite system
  - unary/arbitrary signature (SRS/TRS)
  - strategy (none, inner/outermost, contextsens.)
  - start terms (all, constructor-based)
  - relative/equational: $\longrightarrow_1 \circ \longrightarrow_2^*$, $\longrightarrow_1 \circ \longleftrightarrow_2^*$ instances: $\longleftrightarrow_2^*$ for A, C, I
- coming: background theories (integers, . . .)

What Problems?

Termination Problem Data Base

combined categories	FP	LP	SRS	TRS
number of problems	1676	351	777	2036

.

What Problems?

Termination Problem Data Base

combined categories	FP	LP	SRS	TRS
number of problems	1676	351	777	2036

.

in 2008: use complete set of problems

- good for archival purposes
- bad for excitement

What Problems?

Termination Problem Data Base

combined categories	FP	LP	SRS	TRS
number of problems	1676	351	777	2036

.

in 2008: use complete set of problems

- good for archival purposes
- bad for excitement

coming:

- get more entries for TPDB
- select subsets (benchmarks) for competition

Results: Term Rewriting

“standard” (no theory, no strategy)

total	1391 problems	
AProVE	995 Yes	231 No
$\mathrm{T_TT_2}$	792 Yes	178 No
Jambox	750 Yes	60 No
	558	

Results: Term Rewriting

“standard” (no theory, no strategy)

total	1391 problems	
AProVE	995 Yes	231 No
$\text{T}_\text{T}\text{T}_2$	792 Yes	178 No
Jambox	750 Yes	60 No
AProVE/CoLoR	558 Yes	

best from corresponding certified category

Term Rewriting (non-standard)

innermost: AProve: 241 Yes, 4 No, of 358.
outermost:

total	291 problems	
JamboxGoesOut	72 Yes	0 No
T_TT_2	0 Yes	158 No
TrafO	46 Yes	30 No
AProVE	27 Yes	37 No

modulo AC theory: AProVE: 57/2 of 71
relative: Jambox: 24/0 of 40
contextsensitive: AProVE: 94/0 of 109

String Rewriting

total	732 problems	
$\mathrm{T_TT_2}$	512 Yes	40 No
AProVE	501 Yes	22 No
Jambox	252 Yes	
nonloop		92 No

relative: Jambox: 32/0 of 42

String Rewriting

total	732 problems	
TTT$_2$	512 Yes	40 No
AProVE	501 Yes	22 No
Matchbox (cert.)	466 Yes	
Jambox	252 Yes	
nonloop		92 No

relative: Jambox: 32/0 of 42

Certified Termination

motivation: why do we believe the answers?

- SMT: sat $\longrightarrow$ print model, unsat $\longrightarrow$??
- Termination: no $\longrightarrow$ print loop, yes $\longrightarrow$??

solution: replace “??” by a formal proof,
and use mechanized proof checker.
approaches/implementations:

- A3PAT (Coccinelle)/Coq
- CoLoR/Coq

Certified Termination

motivation: why do we believe the answers?

- SMT: sat $\longrightarrow$ print model, unsat $\longrightarrow$??
- Termination: no $\longrightarrow$ print loop, yes $\longrightarrow$??

solution: replace “??” by a formal proof,
and use mechanized proof checker.
approaches/implementations:

- A3PAT (Coccinelle)/Coq
- CoLoR/Coq
- Isafor/Isabelle

Certified Termination

motivation: why do we believe the answers?

- SMT: sat $\longrightarrow$ print model, unsat $\longrightarrow$??
- Termination: no $\longrightarrow$ print loop, yes $\longrightarrow$??

solution: replace “??” by a formal proof,
and use mechanized proof checker.
approaches/implementations:

- A3PAT (Coccinelle)/Coq
- CoLoR/Coq
- Isafor/Isabelle
- extraction (Isafor $\longrightarrow$ Haskell)

Certified Termination (Results)

term rewriting:

total	1391 problems
AProVE/CoLoR	558
AProve/CoLoR∨A3PAT	520
Cime3/A3PAT	485

string rewriting:

total	732 problems
Matchbox/CoLoR	466
AProVE/CoLoR	406
AProve/CoLoR∨A3PAT	371

Derivational Complexity

- main focus: polynomial upper bounds
- methods:
  - upper triangular matrix interpretations
  - match bounds
  - arctic matrices

more detailed output (degree of polynomial)
requires more detailed scoring.
result: CAT $>$ TCT

Derivational Complexity

- main focus: polynomial upper bounds
- methods:
  - upper triangular matrix interpretations
  - match bounds
  - arctic matrices

more detailed output (degree of polynomial)
requires more detailed scoring.
result: CAT $>$ TCT
coming (?): lower bounds (cf. loops), more
functions (exp, ack), certification

Challenges for the Future

organization

- more visibility (termcomp while conference)
- define benchmarks (= problem sets)
- more efficient steering committee

Challenges for the Future

organization

- more visibility (termcomp while conference)
- define benchmarks ($=$ problem sets)

Challenges for the Future

organization

- more visibility (termcomp while conference)
- define benchmarks (= problem sets)

host/platform

- more flexibility (termexec)
- problems/results queries (also pre-2008)

Challenges for the Future

organization

- more visibility (termcomp while conference)
- define benchmarks (= problem sets)

host/platform

- more flexibility (termexec)
- problems/results queries (also pre-2008)

participants

- better use of hardware (multi-core)
- better re-use of software (modules, interfaces)