Termination Competition 2015

Jürgen Giesl   Frederic Mesnard
Albert Rubio   René Thiemann   Johannes Waldmann

RWTH Aachen   HTWK Leipzig   Universite de la Reunion
UPC BarcelonaTech   University of Innsbruck   HTWK Leipzig

August, 2015. CADE-25

The Halting Problem

The longer it keeps you waiting
the more you appreciate a termination analysis

History of the Termination Competition

- Started in 2003.

- From 2004 to 2009: executed online on all benchmarks

- From 2009 on: random selection of benchmarks

- From 2010 on: Live execution during a conference.

- 2014: First time running under StarExec

Competition areas

- Term Rewriting and Transition systems

  TRS (Standard, Context-Sensitive, Higher-Order, Integer, Cycles,...)
  String Rewrite Systems
  Certified categories
  Integer Transition Systems

- Complexity analysis

  Runtime complexity (TRS)
  Derivational complexity (TRS)
  Certified categories

- Programming Languages

  C, Integer C, Java, Haskell, Prolog.

Tools

15 tools from 13 teams

1. AProVE (Aachen, Germany)
2. AutoNon (Amsterdam, The Netherlands)
3. Ctrl (Innsbruck, Austria)
4. cycsrs (Frankfurt, Germany and Eindhoven, The Netherlands)
5. HipTNT+ (Singapore)
6. matchbox (Leipzig, Germany)
7. muterm (Valencia, Spain)
8. NaTT, (Nagoya, Japan)
9. T2 (Microsoft Cambridge, UK)
10. TCT2 and TCT3 (Innsbruck, Austria)
11. TTT2 (Innsbruck, Austria)
12. UltimateBuchiAutomizer (+Joogie) (Freiburg, Germany and Canberra, Australia)
13. Wanda (Innsbruck, Austria)

Running Competition

- Execution organizer:

  Johannes Waldmann

- Second time running under StarExec Platform.

- An important reimplementation effort needed.

- Unexpectedly many more problems appeared in this second use of StarExec

  Solving platform problems until the very last moment.

Running Competition

- Benchmarks taken form the Termination Problem Data Base (TPDB)

- Timeout 300 seconds

- Only categories with at least two participants (from different teams) are run in the competition.
  There is a full demonstration run afterwards with all categories.

Termination Competition 2015

General Information wc = 300 a = 1 b = 1 c = 0.1 (2015-08-05 18:48:18.60928 UTC) 51787 pairs, 12024641.5 / 6232857.3 s finished in 399686h 5m 50s

Termination of Term Rewriting (and Transition Systems) finished in 399686h 5m 50s, 33684 pairs, 6726808.6 / 3493948.5 s

Combined Ranking (Rules): 1. AProVE 2015 (20) 2. TTT2 (10) 3. NaTT 1.3 (10) 4. matchbox2015-07-26.1 (7) 5. muterm 5.17 (6) 6. AProVE certified (3) 7. T2 - 2015-07-09 - 13745bd6 (2) Wanda (2)
9. AProVE certified TRS Standard (1) 10. cycsrs-29-07-2015.5 (0) AutoNon 1.21 (0) Ctrl (0)

category	post-proc	rankings	statistics
TRS Standard	plain.3	AProVE 2015 (1310), NaTT 1.3 (1023), TTT2 (989), muterm 5.17 (834), Wanda (636), matchbox2015-07-26.1 (524), AutoNon 1.21 (228),	10486 pairs, 1666958.3 / 824051.3 s
SRS Standard	plain.3	AProVE 2015 (832), TTT2 (598), matchbox2015-07-26.1 (365), NaTT 1.3 (202), muterm 5.17 (135), AutoNon 1.21 (58),	7890 pairs, 2570235.3 / 1324912.6 s
Cycles	plain.3	matchbox2015-07-26.1 (646), cycsrs-29-07-2015.5 (422),	2630 pairs, 950125.9 / 453572.0 s
TRS Relative	plain.3	NaTT 1.3 (70), AProVE 2015 (55), TTT2 (41), matchbox2015-07-26.1 (40),	392 pairs, 77146.9 / 34711.5 s
SRS Relative	plain.3	AProVE 2015 (88), matchbox2015-07-26.1 (32), TTT2 (24), NaTT 1.3 (17),	820 pairs, 274225.6 / 145649.5 s
TRS Standard certified	ceta-2.20-2 AProVE certified TRS Standard (1223), TTT2 (962),		2996 pairs, 251319.4 / 124453.7 s
SRS Standard certified	ceta-2.20-2 AProVE certified (816), TTT2 (570),		2630 pairs, 486311.2 / 223843.9 s
TRS Relative certified	ceta-2.20-2 AProVE certified (51), TTT2 (41),		196 pairs, 29112.5 / 17898.6 s
SRS Relative certified	ceta-2.20-2 AProVE certified (88), TTT2 (20),		410 pairs, 76372.8 / 46591.8 s
TRS Equational	plain.3	AProVE 2015 (67), muterm 5.17 (63),	152 pairs, 3067.0 / 3466.0 s
TRS Conditional	plain.3	muterm 5.17 (101), AProVE 2015 (85),	234 pairs, 5576.4 / 5099.8 s
TRS Context Sensitive	plain.3	muterm 5.17 (98), AProVE 2015 (97),	216 pairs, 7100.2 / 5007.1 s
TRS Innermost	plain.3	AProVE 2015 (273), muterm 5.17 (203),	732 pairs, 102735.3 / 77628.8 s
Integer Transition Systems	plain.3	T2 - 2015-07-09 - 13745bd6 (1061), AProVE 2015 (1034), Ctrl (423),	3666 pairs, 212567.8 / 200512.7 s
Integer TRS	plain.3	AProVE 2015 (102), Ctrl (85),	234 pairs, 13954.0 / 6549.1 s

Complexity Analysis of Term Rewriting finished in 399683h 19m 8s, 14796 pairs, 5140915.1 / 2607632.9 s

Combined Ranking (Rules): 1. TCT3_2015 (6) 2. AProVE 2015 (4) 3. AProVE certified (1) matchbox2015-07-26.1 (1) 5. TCT2_20150725 (0)

category	post-proc	rankings	statistics
Derivational Complexity - Full Rewriting	plain.3	TCT3_2015 (853), matchbox2015-07-26.1 (369), TCT2_20150725 (0),	5427 pairs, 2534205.9 / 1205378.3 s
Runtime Complexity - Full Rewriting	plain.3	AProVE 2015 (1218), TCT3_2015 (414), TCT2_20150725 (0),	2877 pairs, 791627.1 / 439245.3 s
Runtime Complexity - Innermost Rewriting	plain.3	AProVE 2015 (2102), TCT3_2015 (769), TCT2_20150725 (0),	3246 pairs, 916945.1 / 491516.2 s
Runtime Complexity - Innermost Rewriting certified	ceta-2.20-2	TCT3_2015 (689), AProVE certified (495), TCT2_20150725 (0),	3246 pairs, 898136.9 / 471493.2 s

Termination of Programming Languages finished in 399683h 22m 25s, 3307 pairs, 156917.9 / 131275.9 s

Combined Ranking (Rules): 1. UltimateBuchiAutomizer (3) 2. HipTNT+ v3 (2) AProVE 2015 (2)

category	post-proc	rankings	statistics
C	plain.3	UltimateBuchiAutomizer (277), AProVE 2015 (252), HipTNT+ v3 (249),	1416 pairs, 84389.0 / 85616.8 s
C Integer Programs	plain.3	HipTNT+ v3 (305), UltimateBuchiAutomizer (295), AProVE 2015 (289),	1005 pairs, 40026.8 / 22753.0 s
Java Bytecode	plain.3	AProVE 2015 (410), UltimateBuchiAutomizer (141),	886 pairs, 32502.0 / 22906.2 s

Termination Competition 2015 data is produced on StarExec at U Iowa, and aggregated on star-exec-presenter at F-IMN, HTWK Leipzig.

Competition Data

- 15 tools

- \> 15,000 problems from the TPDB (benchmarks library)

- 120 execution nodes (StarExec).

- ∼ 14 hours of live execution (would be 10 weeks in single node!)

- CeTA is the certifier in use (Christian Sternagel and René Thiemann)

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

SRS Standard:

1st	AProVE
2nd	TTT2
3rd	matchbox

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

SRS Standard:

1st	AProVE
2nd	TTT2
3rd	matchbox

Cycles:

1st	matchbox
2nd	cycsrs

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

TRS Relative:

1st	NaTT
2nd	AProVE
3rd	TTT2

SRS Standard:

1st	AProVE
2nd	TTT2
3rd	matchbox

Cycles:

| 1st | matchbox |
| 2nd | cycsrs |

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

TRS Relative:

1st	NaTT
2nd	AProVE
3rd	TTT2

SRS Standard:

1st	AProVE
2nd	TTT2
3rd	matchbox

SRS Relative::

1st	AProVE
2nd	matchbox
3rd	TTT2

Cycles:

1st	matchbox
2nd	cycsrs

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

TRS Relative:

1st	NaTT
2nd	AProVE
3rd	TTT2

SRS Standard:

1st	AProVE
2nd	TTT2
3rd	matchbox

SRS Relative::

1st	AProVE
2nd	matchbox
3rd	TTT2

Cycles:

| 1st | matchbox |
| 2nd | cycsrs |

TRS Equational:

| 1st | AProVE |
| 2nd | muterm |

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st AProVE
2nd NaTT
3rd TTT2

SRS Standard:

1st AProVE
2nd TTT2
3rd matchbox

Cycles:

1st matchbox
2nd cycsrs

TRS Relative:

1st NaTT
2nd AProVE
3rd TTT2

SRS Relative::

1st AProVE
2nd matchbox
3rd TTT2

TRS Equational:

1st AProVE
2nd muterm

TRS Conditional:

1st muterm
2nd AProVE

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

SRS Standard:

1st	AProVE
2nd	TTT2
3rd	matchbox

Cycles:

1st	matchbox
2nd	cycsrs

TRS Relative:

1st	NaTT
2nd	AProVE
3rd	TTT2

SRS Relative::

1st	AProVE
2nd	matchbox
3rd	TTT2

TRS Equational:

1st	AProVE
2nd	muterm

TRS Conditional:

1st	muterm
2nd	AProVE

TRS Context Sensitive:

1st	muterm
2nd	AProVE

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

TRS Standard:

1st	AProVE
2nd	NaTT
3rd	TTT2

SRS Standard:

1st	AProVE
2nd	TTT2
3rd	matchbox

Cycles:

| 1st | matchbox |
| 2nd | cycsrs |

TRS Relative:

1st	NaTT
2nd	AProVE
3rd	TTT2

SRS Relative::

1st	AProVE
2nd	matchbox
3rd	TTT2

TRS Equational:

| 1st | AProVE |
| 2nd | muterm |

TRS Conditional:

| 1st | muterm |
| 2nd | AProVE |

TRS Context Sensitive:

| 1st | muterm |
| 2nd | AProVE |

TRS Innermost:

| 1st | AProVE |
| 2nd | muterm |

Integer Transition Systems:

1st	T2
2nd	AProVE
3rd	Ctrl

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

Integer Transition Systems:

1st	T2
2nd	AProVE
3rd	Ctrl

Integer TRS:

1st	AProVE
2nd	Ctrl

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

Integer Transition Systems:

1st	T2
2nd	AProVE
3rd	Ctrl

Integer TRS:

1st	AProVE
2nd	Ctrl

TRS Standard certified:

1st	AProVE
2nd	TTT2

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

Integer Transition Systems:

1st	T2
2nd	AProVE
3rd	Ctrl

Integer TRS:

1st	AProVE
2nd	Ctrl

TRS Standard certified:

1st	AProVE
2nd	TTT2

SRS Standard certified:

1st	AProVE
2nd	TTT2

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

Integer Transition Systems:

1st	T2
2nd	AProVE
3rd	Ctrl

Integer TRS:

1st	AProVE
2nd	Ctrl

TRS Standard certified:

1st	AProVE
2nd	TTT2

SRS Standard certified:

1st	AProVE
2nd	TTT2

TRS Relative certified:

1st	AProVE
2nd	TTT2

TermComp 2015 Winners. Term Rewriting (and Transition Systems)

Integer Transition Systems:

1st   T2
2nd   AProVE
3rd   Ctrl

Integer TRS:

1st   AProVE
2nd   Ctrl

TRS Standard certified:

1st   AProVE
2nd   TTT2

SRS Standard certified:

1st   AProVE
2nd   TTT2

TRS Relative certified:

1st   AProVE
2nd   TTT2

SRS Relative certified:

1st   AProVE
2nd   TTT2

Runtime Complexity
Full Rewriting:

1st	AProVE
2nd	TCT3

Runtime Complexity
Full Rewriting:

1st	AProVE
2nd	TCT3

Runtime Complexity
Innermost Rewriting:

1st	AProVE
2nd	TCT3

TermComp 2015 Winners. Complexity Analysis

Runtime Complexity
Full Rewriting:
  1st   AProVE
  2nd   TCT3

Runtime Complexity
Innermost Rewriting:
  1st   AProVE
  2nd   TCT3

Runtime Complexity
Innermost Rewriting certified:
  1st   TCT3
  2nd   AProVE

TermComp 2015 Winners. Complexity Analysis

Runtime Complexity
Full Rewriting:
  1st   AProVE
  2nd   TCT3

Runtime Complexity
Innermost Rewriting:
  1st   AProVE
  2nd   TCT3

Runtime Complexity
Innermost Rewriting certified:
  1st   TCT3
  2nd   AProVE

Derivational Complexity
Full Rewriting:
  1st   TCT3
  2nd   matchbox

TermComp 2015 Winners. Programming Languages

- C:

1st	UltimateBuchiAutomizer
2nd	AProVE
3rd	HipTNT+

TermComp 2015 Winners. Programming Languages

- C:

  1st   UltimateBuchiAutomizer
  2nd   AProVE
  3rd   HipTNT+

- C Integer Programs:

  1st   HipTNT+
  2nd   UltimateBuchiAutomizer
  3rd   AProVE

TermComp 2015 Winners. Programming Languages

- C:
    - 1st   UltimateBuchiAutomizer
    - 2nd   AProVE
    - 3rd   HipTNT+

- C Integer Programs:
    - 1st   HipTNT+
    - 2nd   UltimateBuchiAutomizer
    - 3rd   AProVE

- Java Bytecode:
    - 1st   AProVE
    - 2nd   UltimateBuchiAutomizer+Joogie

TermComp 2015 Leaders

- Term Rewriting: AProVE

- Complexity Analysis: TCT3

- Programming Languages: UltimateBuchiAutomizer

But up to 8 tools out of 13 won at least one category!

Check complete results in
http://nfa.imn.htwk-leipzig.de/termcomp-2015/competitions/4

Thanks to all participants

Thanks to all participants

And thanks a lot to Johannes Waldmann