SMT-COMP 2022
17th International Satisfiability Modulo Theory Competition

Haniel Barbosa Jochen Hoenicke François Bobot

Universidade Federal de Minas Gerais, Brazil

Albert-Ludwigs-Universität Freiburg, Germany

CEA List, France

Aug 9, 2022

SMT-COMP

Annual competition for SMT solvers
on (a selection of) benchmarks from SMT-LIB

History

2005	first competition
2013	evaluation instead of competition
2014	since then hosted by StarExec

Goals:

- spur development of SMT solver implementations
- promote SMT solvers and their usage
- support the SMT-LIB project
  - to promote and develop the SMT-LIB format
    - model validation
    - proof checking
  - to collect relevant benchmarks
- engage and include new members

SMT Solvers and SMT-LIB

SMT Solver

- checks formulas in SMT-LIB format for satisfiability modulo theories

SMT-LIB is

1. a language in which benchmarks are written
2. a community effort to collect benchmarks

Non-incremental

391 363 instances
with 1 query each
in 81 logics.

Incremental

43 285 instances
with 33 998 935 queries
in 39 logics.

SMT Solvers and SMT-LIB

SMT Solver

- checks formulas in SMT-LIB format for satisfiability modulo theories

SMT-LIB is

1. a language in which benchmarks are written
2. a community effort to collect benchmarks

Non-incremental

391 363 instances
with 1 query each
in 81 logics.

Incremental

43 285 instances
with 33 998 935 queries
in 39 logics.

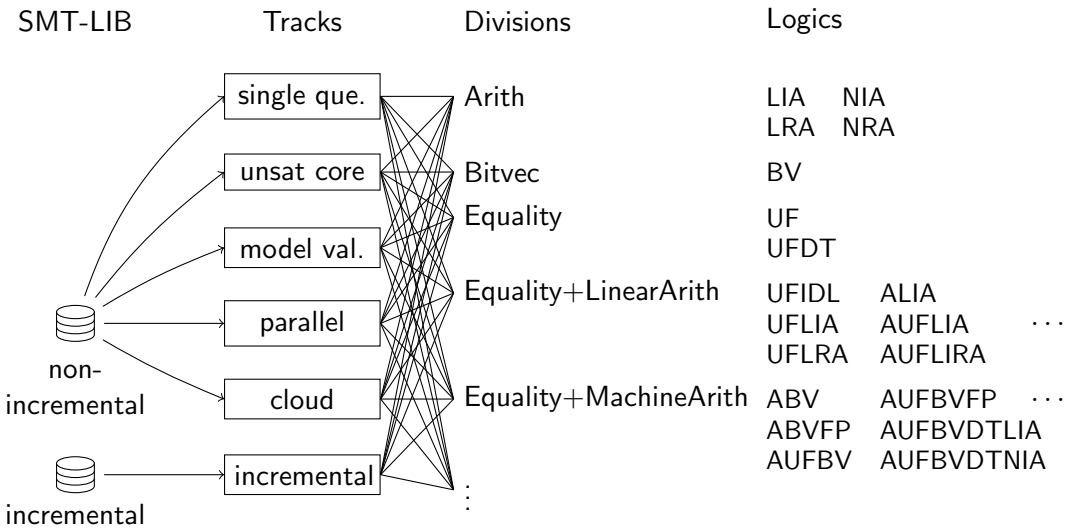
Selected Non-incremental

206 932 instances

Selected Incremental

22 300 instances

Competition Overview



SMT-COMP Tracks (traditional)

Single Query Track

- Determine satisfiability of one problem
- Solver answers sat/unsat/unknown

Unsat Core Track

- Find small unsatisfiable subset of input.
- Solver answers unsat + list of formulas.

Model Validation Track

- Find a model for a satisfiable problem.
- Solver answers sat + value for each non-logical symbol.

Incremental Track

- Solve many small problems interactively.
- Solver acks commands and answers sat/unsat for each check.

SMT-COMP Tracks (experimental)

Model Validation

- Division with quantifier-free floating-point logics
- Model validation with Dolmen (thanks to Gillaume Bury and François Bobot)

Cloud and Parallel Track (sponsored by AWS, led by Mike Whalen)

- Solve a large problem over the cloud (or a big computer)
  - 100 machines, 1600 cores, 6400 GB of memory (cloud)
  - 64 cores, 256 GB of memory (parallel)
- Solver answers sat/unsat/unknown

SMT-COMP Tracks (experimental)

Model Validation

- Division with quantifier-free floating-point logics
- Model validation with Dolmen (thanks to Gillaume Bury and François Bobot)

Cloud and Parallel Track (sponsored by AWS, led by Mike Whalen)

- Solve a large problem over the cloud (or a big computer)
  - 100 machines, 1600 cores, 6400 GB of memory (cloud)
  - 64 cores, 256 GB of memory (parallel)
- Solver answers sat/unsat/unknown

Proof Exhibition Track

- Solver submitted together with a checker for unsatisfiability proofs
- No predefined format or checker
- No ranking
- Qualitative assessment

Tracks, Solvers, Divisions, and Benchmarks

Teams: 21 (+3)

Track	Solvers	Divisions	Benchmarks
Single Query	22(+3)	19(+1)	93 945
Incremental	8(+1)	17(+2)	22 300
Unsat Core	6(-1)	18(+1)	57 245
Model Validation	8(+1)	7(+ 1 exp.)	32 766
Proof Exhibition	4	18 exp.	57 245
Parallel	4(+1)	14 exp.	400
Cloud	4(-1)	14 exp.	400

Number in parenthesis shows changes from 2021

Participants

SMT-COMP 2022 participants rely on multiple reasoning frameworks:

- CDCL(T)
- mcSAT
- saturation
- automata
- finite domain
- CP
- local search
- besides wrappers extending the scope of existing solvers

Six new solvers participated:

- NRA-LS (Liu et al.)
- OSTRICH (Chen et al.)
- Q3B (Jonáš and Strejček)
- Yices-ismt (Jia et al.)
- Z3++ (Cai et al.)
- solsmt (Reitwiessner and Soos)

Non-Competitive Solvers

Submitted by organisers

- z3-4.8.17
- MathSAT 5.6.8
- Division winners from previous years (23 Solvers)

Submitted by participants

- Fixed solvers (OpenSMT, STP, Yices-ismt, Z3++,smtinterpol)

Scoring

Computing scores:

- Single Query/Parallel/Cloud: number of solved instances
- Incremental: number of solved queries
- Unsat Core: number of top-level assertions removed
- Model Validation: number of solved instances with correct models

Error scores:

- All Tracks: given for sat reply for unsat instance, or vice versa
- Unsat Core: given if returned core is satisfiable.
- Model Validation: given if given model evaluates formula to false

Error scores are draconian.

Score and Ranking

In each track we collect different scores:

- Sequential score (SQ, UC, MV): all time limits apply to cpu time
- Parallel score (all): all time limits apply to wallclock time
- SAT score (SQ): parallel score for satisfiable instances
- UNSAT score (SQ): parallel score for unsatisfiable instances
- 24s (SQ): parallel score with time limit of 24s

Division ranking (for each score)

- For each division, one winner is declared

Two competition-wide rankings (for each score)

- Biggest lead: division winner with most score difference to second place
- Largest contribution: improvement each solver provided to a virtual best solver

FLoC Medalists

8 medals

- Biggest Lead
  - Single Query: Gold and Silver
  - Model Validation: Gold
  - Incremental: Gold
- Largest Contribution
  - Single Query: Gold and Silver
  - Model Validation: Gold
  - Incremental: Gold

FLoC Medalists

cvc5 wins three gold medals:

- Single Query, Biggest Lead
- Single Query, Largest Contribution
- Incremental, Largest Contribution

FLoC Medalists

Z3++ wins two gold medals:

- Model Validation, Biggest Lead
- Model Validation, Largest Contribution

FLoC Medalists

SMTInterpol wins one gold medal:

- Incremental, Biggest Lead

FLoC Medalists

Bitwuzla wins one silver medal:

- Single Query, Biggest Lead

FLoC Medalists

YicesQS wins one silver medal:

- Single Query, Largest Contribution

For full results please come to the SMT workshop Thursday 4pm.