SAT Race 2019

Overview and Results

Marijn Heule[1] Matti Järvisalo[2] Martin Suda[3]

[1] Department of Computer Science, The University of Texas at Austin, USA
[2] HIIT, Department of Computer Science, University of Helsinki, Finland
[3] Czech Technical University, Prague, Czech Republic

July 12, 2019 @ SAT 2019, Lisbon

SAT Solver Competitions

Goals

- identify new challenging benchmarks
- promote SAT solvers & their development
- "snapshot" evaluation of current solvers

Long tradition, starting from 1992

- 3 competitions in the 90s (1992, 1993, 1996)
- 12 SAT Competitions (2002–)
- 5 SAT Races (2006, 2008, 2010, 2015, 2019)
- 1 SAT Challenge (2012)

Key rules

- Certified UNSAT using DRAT proof logging
- Disqualification of buggy solvers
  - Provided model incorrect
  - Report UNSAT on know-to-be-satisfiable instance
  - Proof check fails on UNSAT instance $\rightarrow$ "timeout"
    transition-period rule, will likely be changed
- Mandatory solver descriptions + open source

Other Rules and Changes

- Ranking scheme: PAR-2
  - Favors solvers that are faster (not only count solved instances)
- Select at most 20 instances from same source
  - Submitters cannot influence the competition too much
- Different this year:
  - No BYOB (Bring Your Own Benchmarks): submissions 50% down
  - (Slightly) scrambling all old benchmarks
  - No timeout for proof checking

Solvers and Benchmarks

55 solvers participated in the SAT Race 2019

From the submitted benchmarks, we were able to select 200 interesting ones for the race. Additionally 200 instances (not solvable by MiniSAT in 600 seconds) from prior competitive events were added.

New Benchmarks: Several new benchmark domains/sets were submitted, Bounded Model Checking, Cryptography, Edge-Matching Puzzles, Integer factoring, Miters, Matrix Multiplication, Knight Hamiltonian Problem, SHA-1 Pre-image, Tensor Rank

All old instances were scrambled by scranfilize (default settings)

Results

SAT Track: Top-3

SAT Track: Top-3

3. **smallsat** default (3505.78)
   by Jingchao Chen

SAT Track: Top-3

2. **MapleLCMDistChronoBT-DL** v2.1 (3436.32)
   MapleLCMDistChronoBT-DL v2.2 (3441.61)
   MapleLCMDiscChronoBT-DL v3 (3448.87)
   by Stepan Kochemazov, Oleg Zaikin, Victor Kondratiev, and Alexander Semenov
3. **smallsat** default (3505.78)
   by Jingchao Chen

SAT Track: Top-3

1. **CaDiCaL** sat (3176.29)
   CaDiCaL default (3322.23)
   by Armin Biere
2. **MapleLCMDistChronoBT-DL** v2.1 (3436.32)
   MapleLCMDistChronoBT-DL v2.2 (3441.61)
   MapleLCMDiscChronoBT-DL v3 (3448.87)
   by Stepan Kochemazov, Oleg Zaikin, Victor Kondratiev, and Alexander Semenov
3. **smallsat** default (3505.78)
   by Jingchao Chen

UNSAT Track: Top-3

UNSAT Track: Top-3

3. **PSIDS_MapleLCMDistChronoBT** (5640.05)
   by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

UNSAT Track: Top-3

2. **expMaple_CM_GCBumpOnlyLRB** (5616.55)
   by Md Solimul Chowdhury, Martin Müller, and Jia-Huai You
3. **PSIDS_MapleLCMDistChronoBT** (5640.05)
   by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

UNSAT Track: Top-3

1. **MapleLCMDiscChronoBT-DL** v3 (5601.41)
   by Stepan Kochemazov, Oleg Zaikin, Victor Kondratiev, and Alexander Semenov
2. **expMaple_CM_GCBumpOnlyLRB** (5616.55)
   by Md Solimul Chowdhury, Martin Müller, and Jia-Huai You
3. **PSIDS_MapleLCMDistChronoBT** (5640.05)
   by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

SAT + UNSAT Track: Top-3

SAT + UNSAT Track: Top-3

3. **MapleLCMdistCBTcoreFirst** default (4610.86)
   by Jingchao Chen

SAT + UNSAT Track: Top-3

2. **CaDiCaL** default (4582.97)
   by Armin Biere
3. **MapleLCMdistCBTcoreFirst** default (4610.86)
   by Jingchao Chen

SAT + UNSAT Track: Top-3

1. **MapleLCMDistChronoBT-DL** v3 (4525.14)
   MapleLCMDistChronoBT-DL v2.2 (4563.33)
   MapleLCMDiscChronoBT-DL v2.1 (4563.67)
   by Stepan Kochemazov, Oleg Zaikin, Victor Kondratiev, and Alexander Semenov
2. **CaDiCaL** default (4582.97)
   by Armin Biere
3. **MapleLCMdistCBTcoreFirst** default (4610.86)
   by Jingchao Chen

Impact of PAR-2

Penalized average runtime (PAR)

- PAR-x: penalized timeouts by $x\cdot$ TIMEOUT
- SCR, solution-count ranking: PAR-x as $x \to \infty$.
- x balances average successful runs and number of solved instances

In 2019: some differences between PAR-2 and SCR.

Impact of PAR-2

Penalized average runtime (PAR)

- PAR-x: penalized timeouts by $x\cdot$ TIMEOUT
- SCR, solution-count ranking: PAR-x as $x \to \infty$.
- x balances average successful runs and number of solved instances

In 2019: some differences between PAR-2 and SCR.
Determined winner in the SAT parallel track:

PAR-2

1. **MapleLCMDistCBT-DL**
   by Stepan Kochemazov et al.
2. **CaDiCaL**
   by Armin Biere
3. **MapleLCMdistCBTcoreFirst**
   by Jingchao Chen

SCR

(1). **CaDiCaL** (244)
   by Armin Biere

(2). **MapleLCMDistCBT-DL** (241)
   by Stepan Kochemazov et al.

(3). **expMaple_BumpOnly** (240)
   by Md Solimul Chowdhury et al.

Final Remarks

Full details (to be available) at
`http://sat-race-2019.ciirc.cvut.cz/`

- Detailed per-instance per-solver results
- Proceedings of solvers and benchmarks
- These slides

Final Remarks

Full details (to be available) at
`http://sat-race-2019.ciirc.cvut.cz/`

- Detailed per-instance per-solver results
- Proceedings of solvers and benchmarks
- These slides

Many thanks to

- all solver submitters and developers
- all benchmark submitters
- Aaron Stump and StarExec

SAT GRAND CHALLENGES

email suggestions to marijn@heule.nl

Thank you for your attention!