The Results of SAT Competition 2020

Tomáš Balyo, Nils Froleyks, **Marijn Heule**,
Markus Iser, Matti Järvisalo, and Martin Suda

SAT 2020 Conference, Alghero, Italy (virtually)

July 8, 2020

SAT Solver Competitions

Goals

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

Long tradition, starting from 1992

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

Key rules

- Certified results of unsatisfiability using DRAT proof logging
- Disqualification of buggy solvers
  - Producing an incorrect model
  - Report UNSAT on a known satisfiable instance
  - Proof checker finds inconsistency (demoted to no-limit)
- Mandatory solver descriptions + open source
- Ranking scheme: PAR-2
  - Favors solvers that are faster (not only count solved instances)
- BYOB (Bring Your Own Benchmarks)
  - At most 20 instances per participant are used

What is New This Year

- We have two new tracks
  - Cloud Track – evaluate distributed solvers on the Amazon cloud. Solvers are run on 1600 virtual cores for 1000 seconds. Sponsored by Amazon. Participants received AWS credit to develop their solvers.
  - Planning Track – dedicated benchmark suite on 200 planning instances. Future competitions will have special benchmark suites for other applications.

- New formally-verified checker
  - **cake_lpr_array** by Yong Kiam Tan: very easy to install

Benchmark Instance Selection

GBD Benchmark Database (GBD)

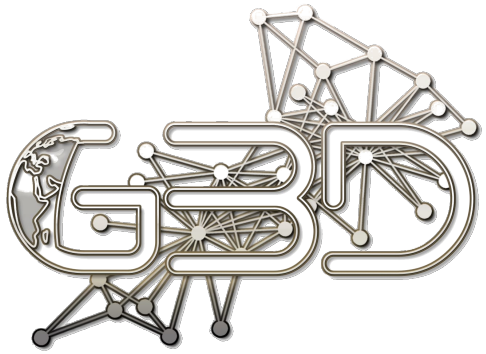
- Collaborative Management of Attributes of Benchmark Instances

  `https://pypi.org/project/global-benchmark-database-tool`

- Retrieval of Benchmark Instances by their Attributes

  `https://gbd.iti.kit.edu`

M. Iser and C. Sinz, "A Problem Meta-Data Library for Research in SAT", Proceedings of Pragmatics of SAT 2018, pp. 144–152, 2018



Tracks part 1

- Main (Sequential) Track (50 solvers)
  - 400 benchmarks, a combination of “application” and “crafted”
  - 5,000 sec limit for solving and 40,000 sec for proof checking
  - Solvers run on a single core
  - UNSAT proof logging required

Tracks part 1

- Main (Sequential) Track (50 solvers)
  - 400 benchmarks, a combination of "application" and "crafted"
  - 5,000 sec limit for solving and 40,000 sec for proof checking
  - Solvers run on a single core
  - UNSAT proof logging required
- Parallel Track (14 solvers)
  - The same 400 benchmarks from Main track
  - 5,000 sec limit for solving
  - 1 AWS m4.16xlarge: 64 virtual CPU cores, 256GB RAM

Tracks part 1

- Main (Sequential) Track (50 solvers)
  - 400 benchmarks, a combination of "application" and "crafted"
  - 5,000 sec limit for solving and 40,000 sec for proof checking
  - Solvers run on a single core
  - UNSAT proof logging required
- Parallel Track (14 solvers)
  - The same 400 benchmarks from Main track
  - 5,000 sec limit for solving
  - 1 AWS m4.16xlarge: 64 virtual CPU cores, 256GB RAM
- Cloud Track (6 solvers)
  - The same 400 benchmarks from Main track
  - 1,000 sec limit for solving
  - 100 AWS m4.4xlarge: total of 1600 virtual CPU cores

Tracks part 2

- Incremental Library Track (5 solvers)
  - benchmarks are SAT based applications (bones, essentials, lsp, max, ijtihad, pasar), we used same applications but with different inputs
  - combined rank for each application determines winner
  - 2,000 sec limit for solving

Tracks part 2

- Incremental Library Track (5 solvers)
  - benchmarks are SAT based applications (bones, essentials, lsp, max, ijtihad, pasar), we used same applications but with different inputs
  - combined rank for each application determines winner
  - 2,000 sec limit for solving
- Planning Track (49 solvers)
  - 200 benchmarks, all coming from planning problems
  - 5,000 sec limit for solving

Tracks part 2

- Incremental Library Track (5 solvers)
  - benchmarks are SAT based applications (bones, essentials, lsp, max, ijtihad, pasar), we used same applications but with different inputs
  - combined rank for each application determines winner
  - 2,000 sec limit for solving
- Planning Track (49 solvers)
  - 200 benchmarks, all coming from planning problems
  - 5,000 sec limit for solving
- No-Limit Track (64 solvers, superset of Main track participants)
  - 300 brand new benchmarks (subset of the Main Track benchmarks)
  - 5,000 sec limit for solving
  - Most of the solvers provided source codes and models, but not all
  - No awards: top solvers were open source and proof producing

Planning Track – Results

The Top 3 solvers of the Planning Track are:

Planning Track – Results

The Top 3 solvers of the Planning Track are:

3 **Kissat-sc2020-unsat** (PAR-2: 3472, 74 solved)
by Armin Biere

Planning Track – Results

The Top 3 solvers of the Planning Track are:

2 **Cryptominisat-ccnr-lsids** (PAR-2: 3441, 79 solved)
Cryptominisat-ccnr (PAR-2: 3446, 79 solved)
by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel

3 **Kissat-sc2020-unsat** (PAR-2: 3472, 74 solved)
by Armin Biere

Planning Track – Results

The Top 3 solvers of the Planning Track are:

1. **CaDiCaL-alluip-trail** (PAR-2: 3406, 80 solved)
   CaDiCaL-alluip (PAR-2: 3409, 80 solved)
   by Randy Hickey, Nick Feng, and Fahiem Bacchus
2. **Cryptominisat-ccnr-lsids** (PAR-2: 3441, 79 solved)
   Cryptominisat-ccnr (PAR-2: 3446, 79 solved)
   by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel
3. **Kissat-sc2020-unsat** (PAR-2: 3472, 74 solved)
   by Armin Biere

Planning Track – Results

The Top 3 solvers of the Planning Track are:

1. **CaDiCaL-alluip-trail** (PAR-2: 3406, 80 solved)
   CaDiCaL-alluip (PAR-2: 3409, 80 solved)
   by Randy Hickey, Nick Feng, and Fahiem Bacchus
2. **Cryptominisat-ccnr-lsids** (PAR-2: 3441, 79 solved)
   Cryptominisat-ccnr (PAR-2: 3446, 79 solved)
   by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel
3. **Kissat-sc2020-unsat** (PAR-2: 3472, 74 solved)
   by Armin Biere

Unfortunately, no planning specific solvers

Incremental Library Track

- 6 applications (bones, essentials, lsp, max, ijtihad, pasar)
- 50 benchmark instances per application
- Ranking by PAR-2 (2000 seconds timeout)
- Final Ranking: Number of Won Categories

	abcdsat-i20		CaDiCaL-sc2020		**Cryptominisat5**		Riss-7.1.2	
bones	513 (46)	2	631 (43)	3	**390 (46)**	**1**	903 (40)	4
essentials	1333 (35)	4	1210 (37)	2	**1200 (36)**	**1**	1241 (36)	3
lsp	2495 (21)	4	1959 (26)	3	**1789 (29)**	**1**	1881 (27)	2
max	**1987 (27)**	**1**	2021 (25)	2	2024 (25)	3	2021 (25)	2
ijtihad	3238 (10)	4	**3002 (13)**	**1**	3079 (12)	2	3145 (11)	3
pasar	471 (45)	2	506 (45)	3	969 (38)	4	**386 (46)**	**1**
final	1		1		**3**		1	

Winner: **Cryptominisat5**

Parallel Track SAT – Results

The Top 3 solvers of the Parallel Track SAT are:

Parallel Track SAT – Results

The Top 3 solvers of the Parallel Track SAT are:

3 **abcd-para-scavel** (PAR-2: 3405, 143 solved)
by Zhihui Li, Guanfeng Wu, Yanh Xu, and Qingshan Chen

Parallel Track SAT – Results

The Top 3 solvers of the Parallel Track SAT are:

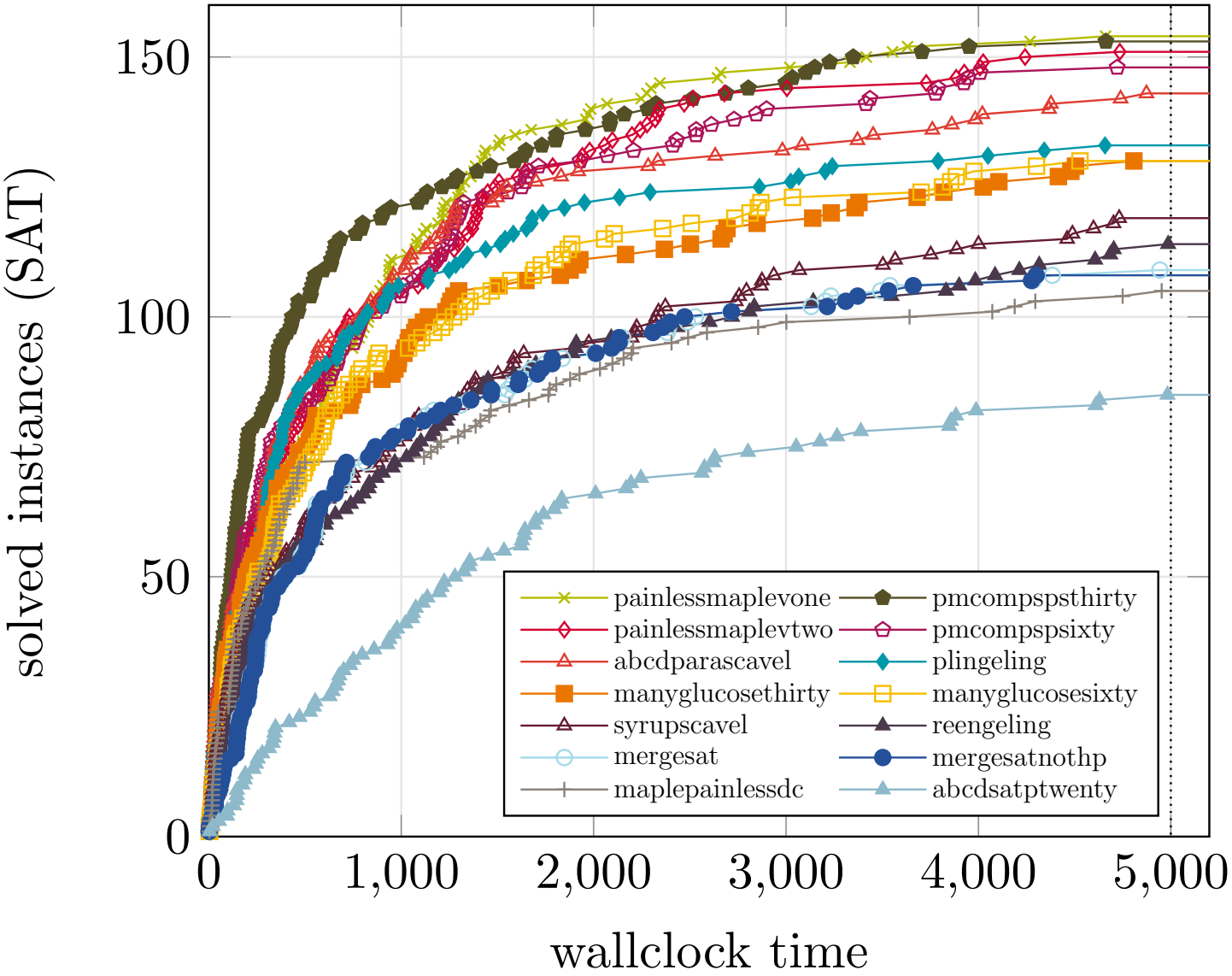
2. **ExMapleLCMDistChronoBT** (PAR-2: 2913, 154 solved)
   by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni
3. **abcd-para-scavel** (PAR-2: 3405, 143 solved)
   by Zhihui Li, Guanfeng Wu, Yanh Xu, and Qingshan Chen

Parallel Track SAT – Results

The Top 3 solvers of the Parallel Track SAT are:

1. **P-MCOMSPS-STR-32** (PAR-2: 2853, 153 solved)
   by Vincent Vallade, Ludovic Le Frioux, Souheib Baarir, Julien Sopena, and Fabrice Kordon
2. **ExMapleLCMDistChronoBT** (PAR-2: 2913, 154 solved)
   by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni
3. **abcd-para-scavel** (PAR-2: 3405, 143 solved)
   by Zhihui Li, Guanfeng Wu, Yanh Xu, and Qingshan Chen

Parallel Track SAT – Plot



Parallel Track UNSAT – Results

The Top 3 solvers of the Parallel Track UNSAT are:

Parallel Track UNSAT – Results

The Top 3 solvers of the Parallel Track UNSAT are:

3 **ManyGlucose-32** (PAR-2 3844, 131 solved)
ManyGlucose-64 (PAR-2: 3974, 129 solved)
by Hidetomo Nabeshima and Katsumi Inoue

Parallel Track UNSAT – Results

The Top 3 solvers of the Parallel Track UNSAT are:

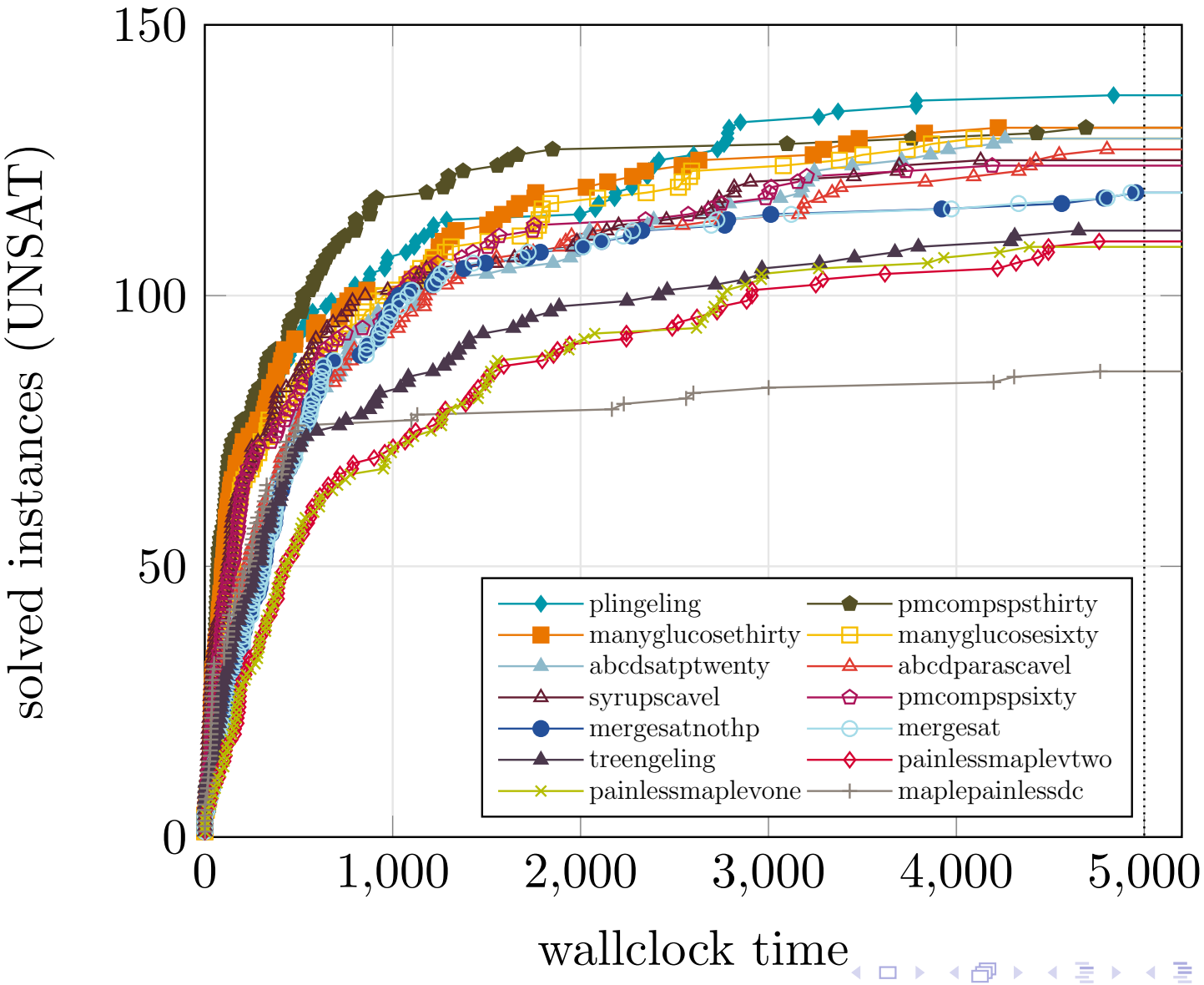
2. **P-MCOMSPS-STR-32** (PAR-2: 3729, 131 solved)
   by Vincent Vallade, Ludovic Le Frioux, Souheib Baarir, Julien Sopena, and Fabrice Kordon
3. **ManyGlucose-32** (PAR-2 3844, 131 solved)
   ManyGlucose-64 (PAR-2: 3974, 129 solved)
   by Hidetomo Nabeshima and Katsumi Inoue

Parallel Track UNSAT – Results

The Top 3 solvers of the Parallel Track UNSAT are:

1. **Plingeling** (PAR-2: 3630, 137 solved)
   by Armin Biere
2. **P-MCOMSPS-STR-32** (PAR-2: 3729, 131 solved)
   by Vincent Vallade, Ludovic Le Frioux, Souheib Baarir, Julien Sopena, and Fabrice Kordon
3. **ManyGlucose-32** (PAR-2 3844, 131 solved)
   ManyGlucose-64 (PAR-2: 3974, 129 solved)
   by Hidetomo Nabeshima and Katsumi Inoue

Parallel Track UNSAT – Plot



Parallel Track ALL – Results

The Top 3 solvers of the Parallel Track ALL are:

Parallel Track ALL – Results

The Top 3 solvers of the Parallel Track ALL are:

3 **abcd-para-scavel** (PAR-2: 3797, 270 solved)
by Zhihui Li, Guanfeng Wu, Yanh Xu, and Qingshan Chen

Parallel Track ALL – Results

The Top 3 solvers of the Parallel Track ALL are:

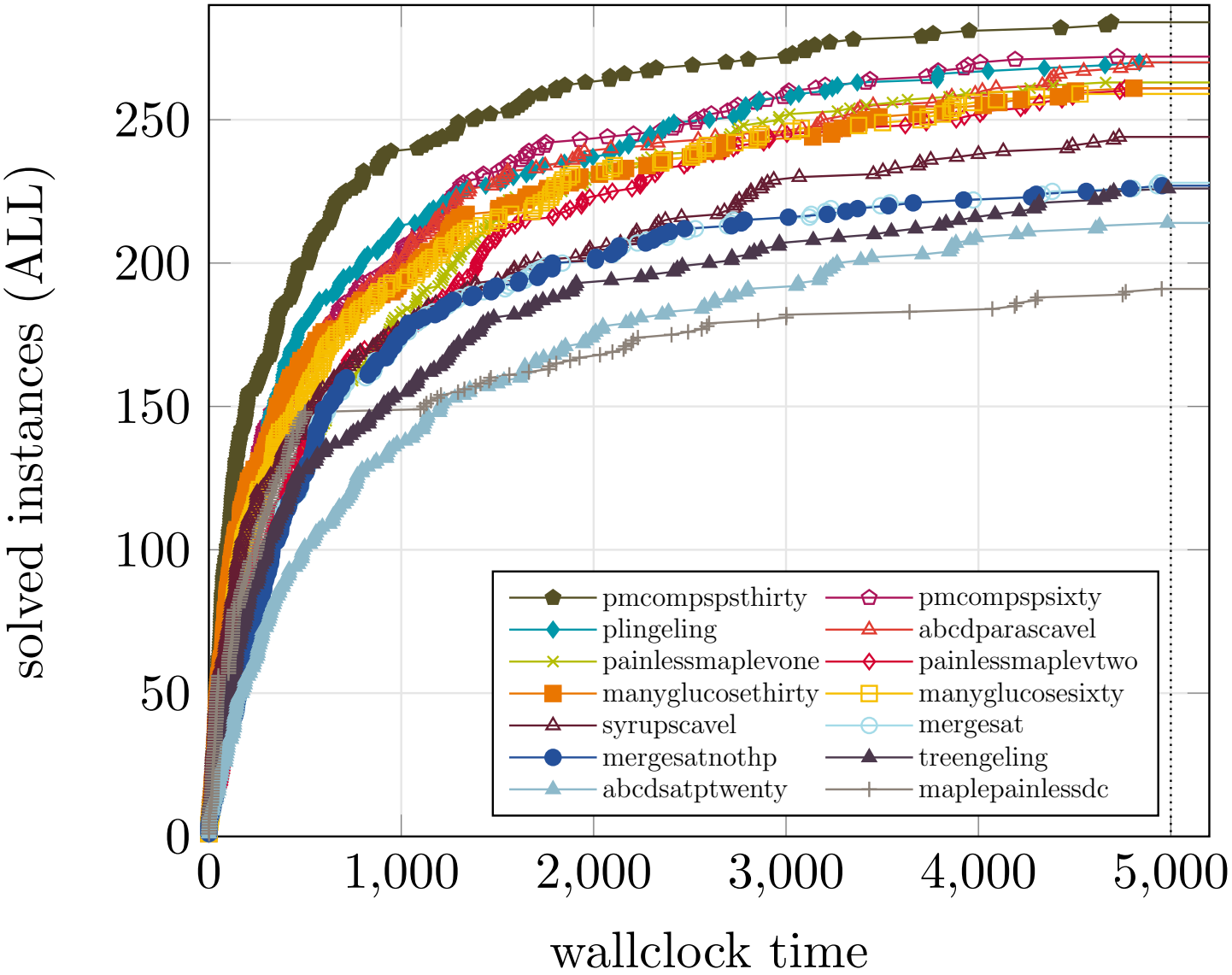
2. **Plingeling** (PAR-2: 3718, 270 solved)
   by Armin Biere
3. **abcd-para-scavel** (PAR-2: 3797, 270 solved)
   by Zhihui Li, Guanfeng Wu, Yanh Xu, and Qingshan Chen

Parallel Track ALL – Results

The Top 3 solvers of the Parallel Track ALL are:

1. **P-MCOMSPS-STR-32** (PAR-2: 3291, 284 solved)
   P-MCOMSPS-STR-64 (PAR-2: 3689, 272 solved)
   by Vincent Vallade, Ludovic Le Frioux, Souheib Baarir, Julien Sopena, and Fabrice Kordon
2. **Plingeling** (PAR-2: 3718, 270 solved)
   by Armin Biere
3. **abcd-para-scavel** (PAR-2: 3797, 270 solved)
   by Zhihui Li, Guanfeng Wu, Yanh Xu, and Qingshan Chen

Parallel Track ALL – Plot



Cloud Track – Results

The Top 3 solvers of the Cloud Track are:

Cloud Track – Results

The Top 3 solvers of the Cloud Track are:

3 **Slime** (PAR-2 4208, 239 solved)
by Oscar Riveros

Cloud Track – Results

The Top 3 solvers of the Cloud Track are:

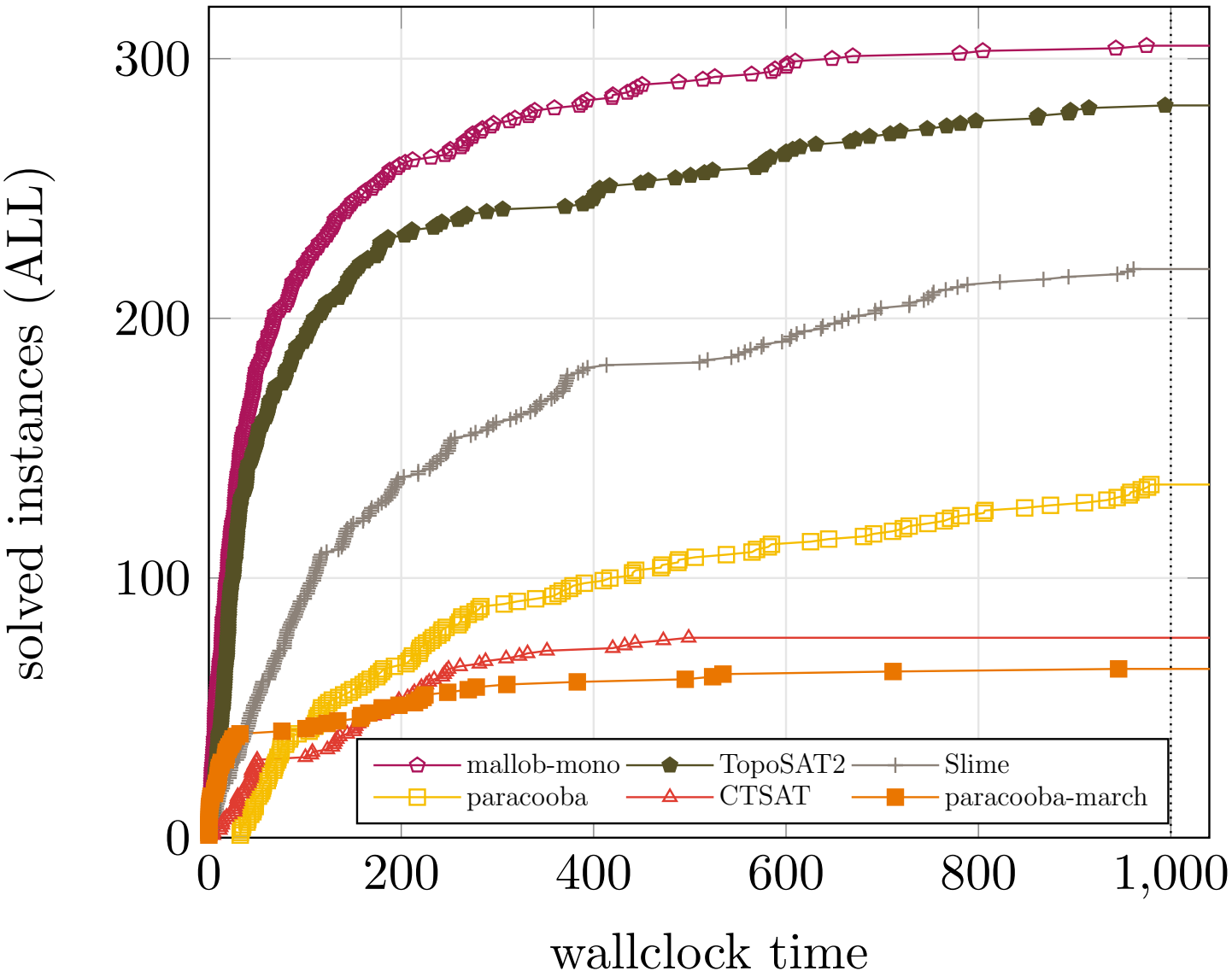
2. **TopoSAT2** (PAR-2: 3024, 283 solved)
   by Thorsten Ehlers, Mitja Kulczynski, Dirk Nowotka, and Philipp Sieweck
3. **Slime** (PAR-2 4208, 239 solved)
   by Oscar Riveros

Cloud Track – Results

The Top 3 solvers of the Cloud Track are:

1. **mallob-mono** (PAR-2: 2429, 306 solved)
   by Dominik Schreiber
2. **TopoSAT2** (PAR-2: 3024, 283 solved)
   by Thorsten Ehlers, Mitja Kulczynski, Dirk Nowotka, and Philipp Sieweck
3. **Slime** (PAR-2 4208, 239 solved)
   by Oscar Riveros

Cloud Track – Plot



Main Track SAT – Results

The Top 3 solvers of the Main Track SAT are:

Main Track SAT – Results

The Top 3 solvers of the Main Track SAT are:

3 **Cryptominisat-ccnr-lsids** (PAR-2: 3263, 144 solved)
Cryptominisat-ccnr (PAR-2: 3317, 145 solved)
by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel

Main Track SAT – Results

The Top 3 solvers of the Main Track SAT are:

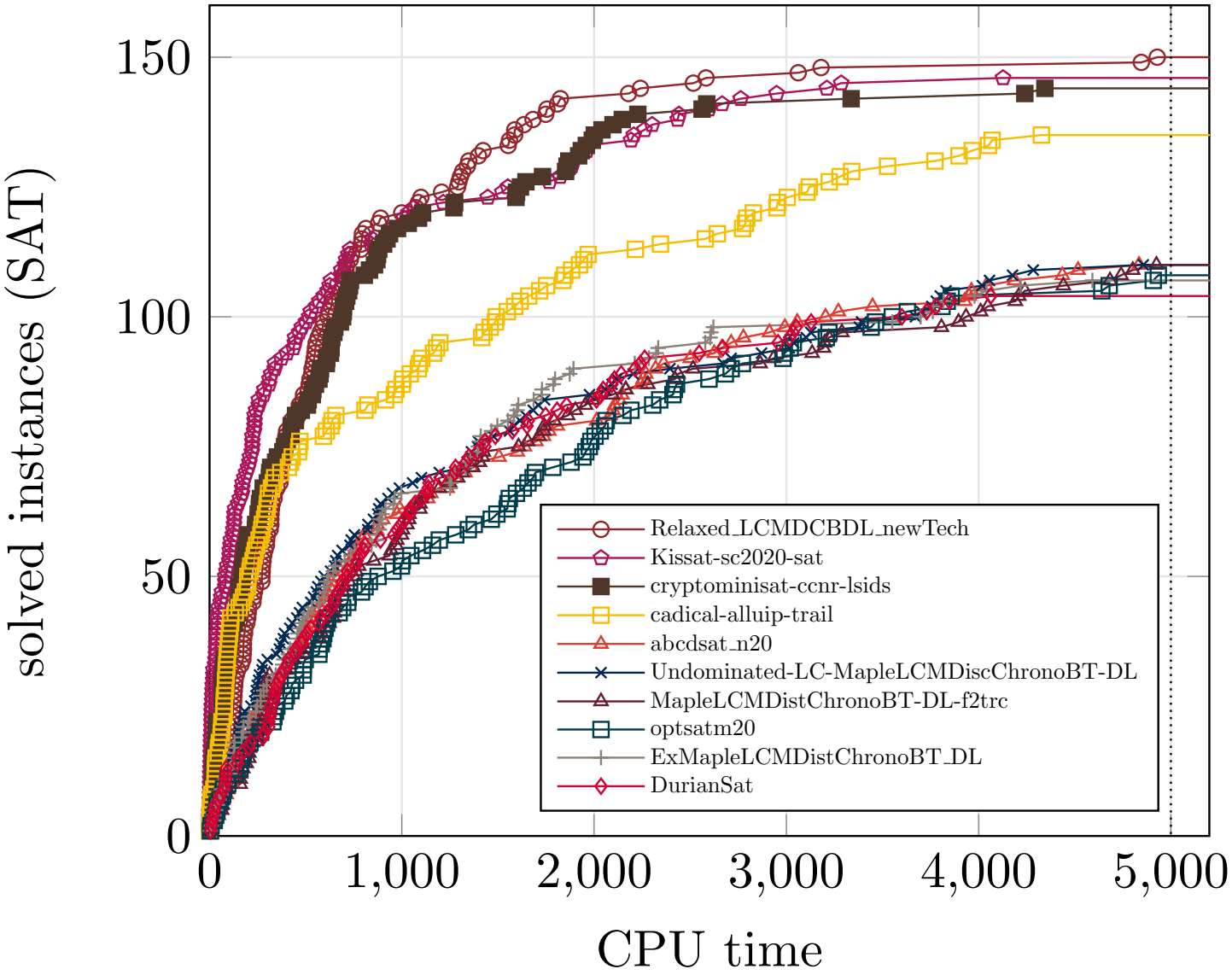
2. **Kissat-sc2020-sat** (PAR-2: 3128, 146 solved)
   by Armin Biere
3. **Cryptominisat-ccnr-lsids** (PAR-2: 3263, 144 solved)
   Cryptominisat-ccnr (PAR-2: 3317, 145 solved)
   by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel

Main Track SAT – Results

The Top 3 solvers of the Main Track SAT are:

1. **Relaxed_LCMDCBDL_newTech** (PAR-2: 2997, 150 solved)
   by Xindi Zhang and Shaowei Cai
2. **Kissat-sc2020-sat** (PAR-2: 3128, 146 solved)
   by Armin Biere
3. **Cryptominisat-ccnr-lsids** (PAR-2: 3263, 144 solved)
   Cryptominisat-ccnr (PAR-2: 3317, 145 solved)
   by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel

Main Track SAT – Top 10 Plot



Main Track UNSAT – Results

The Top 3 solvers of the Main Track UNSAT are:

Main Track UNSAT – Results

The Top 3 solvers of the Main Track UNSAT are:

3 **MapleLCMDistChronoBT-f2trc-s** (PAR-2: 4991, 110 solved)
by Stepan Kochemazov

Main Track UNSAT – Results

The Top 3 solvers of the Main Track UNSAT are:

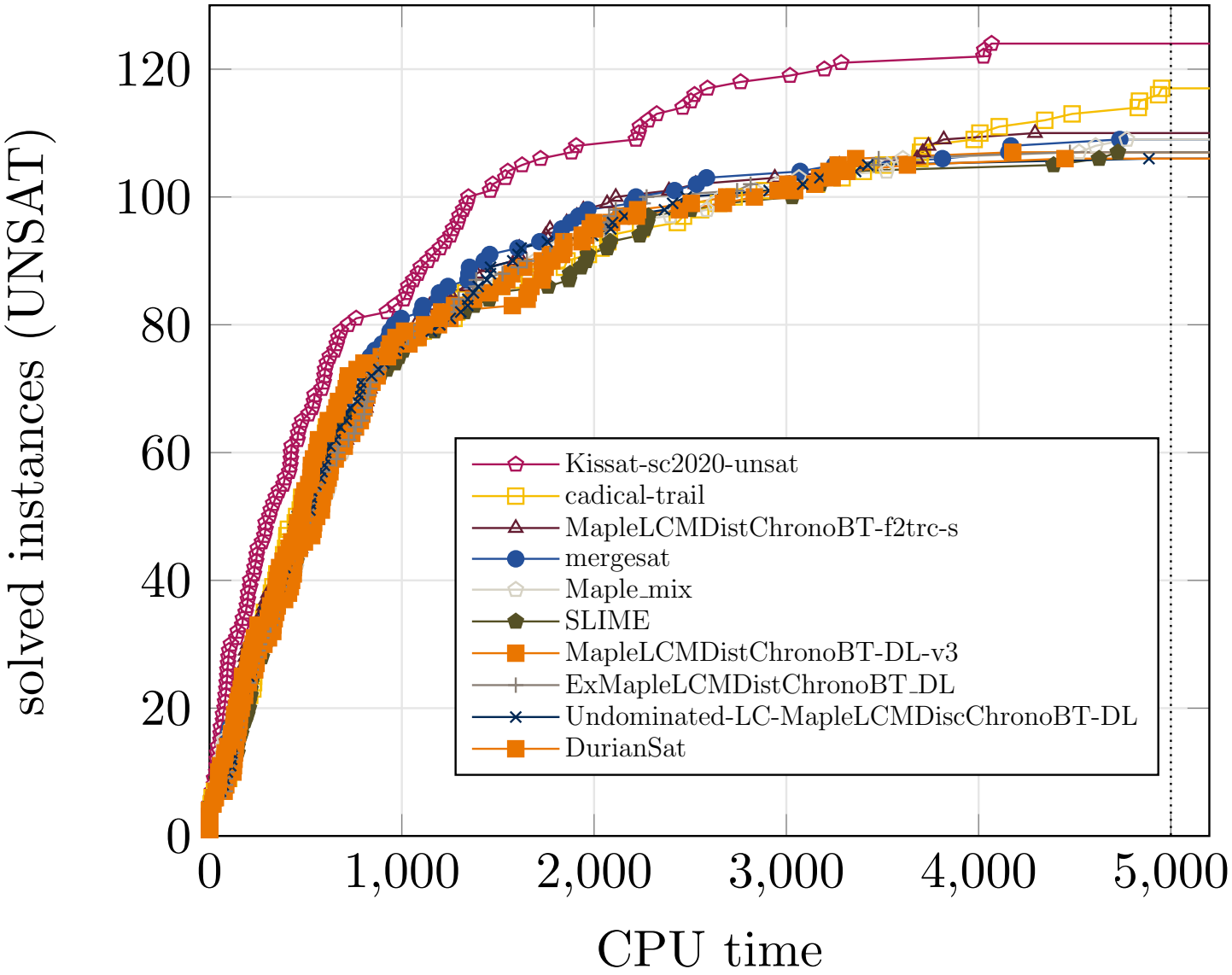
2. **CaDiCaL-trail** (PAR-2: 4842, 117 solved)
   by Randy Hickey, Nick Feng, and Fahiem Bacchus
3. **MapleLCMDistChronoBT-f2trc-s** (PAR-2: 4991, 110 solved)
   by Stepan Kochemazov

Main Track UNSAT – Results

The Top 3 solvers of the Main Track UNSAT are:

1. **Kissat-sc2020-unsat** (PAR-2: 4315, 124 solved)
   Kissat-sc2020-default (PAR-2: 4336, 126 solved)
   Kissat-sc2020-sat (PAR-2: 4725, 118 solved)
   by Armin Biere
2. **CaDiCaL-trail** (PAR-2: 4842, 117 solved)
   by Randy Hickey, Nick Feng, and Fahiem Bacchus
3. **MapleLCMDistChronoBT-f2trc-s** (PAR-2: 4991, 110 solved)
   by Stepan Kochemazov

Main Track UNSAT – Top 10 Plot



Main Track ALL – Results

The Top 3 solvers of the Main Track ALL are:

Main Track ALL – Results

The Top 3 solvers of the Main Track ALL are:

3 **Cryptominisat-ccnr-lsids** (PAR-2: 4267, 248 solved)
Cryptominisat-ccnr (PAR-2: 4278, 250 solved)
by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel

Main Track ALL – Results

The Top 3 solvers of the Main Track ALL are:

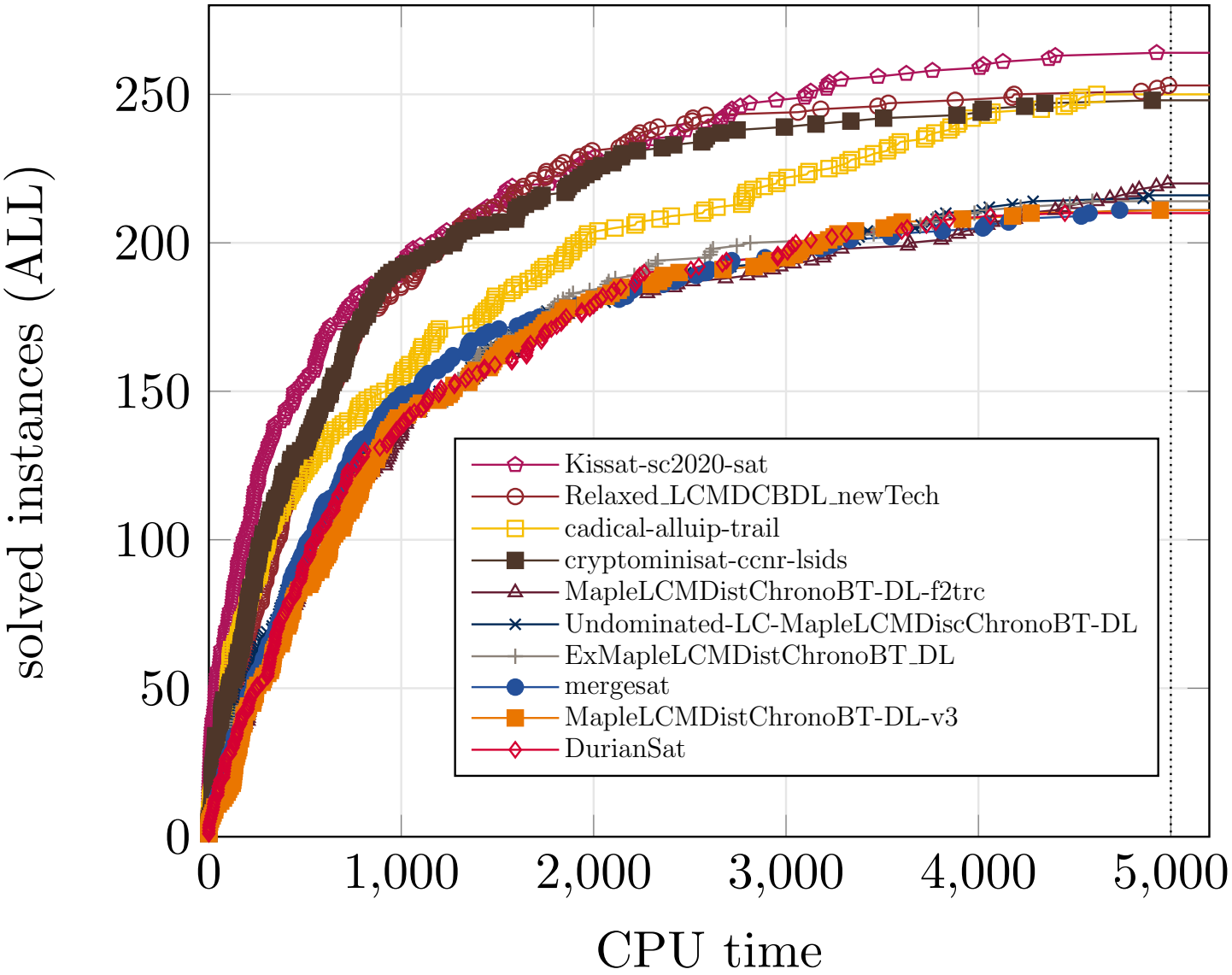
2. **Relaxed_LCMDCBDL_newTech** (PAR-2: 4179, 253 solved)
   by Xindi Zhang and Shaowei Cai
3. **Cryptominisat-ccnr-lsids** (PAR-2: 4267, 248 solved)
   Cryptominisat-ccnr (PAR-2: 4278, 250 solved)
   by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel

Main Track ALL – Results

The Top 3 solvers of the Main Track ALL are:

1. **Kissat-sc2020-sat** (PAR-2: 3926, 264 solved)
   Kissat-sc2020-default (PAR-2: 4083, 260 solved)
   by Armin Biere
2. **Relaxed_LCMDCBDL_newTech** (PAR-2: 4179, 253 solved)
   by Xindi Zhang and Shaowei Cai
3. **Cryptominisat-ccnr-lsids** (PAR-2: 4267, 248 solved)
   Cryptominisat-ccnr (PAR-2: 4278, 250 solved)
   by Mate Soos, Shaowei Cai, Jo Devriendt, Stephan Gocht, Arijit Shaw, and Kuldeep Meel

Main Track ALL– Top 10 Plot



More information and Acknowledgments

Additionals Information

- The Competition Proceedings (solver and benchmark descriptions) will soon be available at https://satcompetition.github.io/2020/
- For the detailed competition results see the SAT Competition website

Acknowledgments

- Thanks to all the participants
- Thanks for all the benchmarks
- Thanks to Mike Whalen, Jonathan Eidelman, and Frankie Botero at AWS
- Thanks to Aaron Stump and StarExec
- Thanks to CAS Software Karlsruhe for the medals
- Thank You for Your attention