

# SAT Competition 2018

## Overview and Results

Marijn Heule<sup>1</sup>   Matti Järvisalo<sup>2</sup>   Martin Suda<sup>3</sup>

<sup>1</sup> Department of Computer Science, The University of Texas at Austin, USA

<sup>2</sup> HIIT, Department of Computer Science, University of Helsinki, Finland

<sup>3</sup> TU Wien, Vienna, Austria

July 12 12, 2018 @ SAT'18, Oxford

# 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–)
- 4 SAT Races (2006, 2008, 2010, 2015)
- 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 → “timeout”  
*transition-period rule, will likely be changed*
- Mandatory solver descriptions + open source

## Recent changes (since 2017)

- Ranking scheme: PAR-2
  - Favors solvers that are faster (not only count solved instances)
- BYOB — Bring your own ~~beer~~ benchmarks
  - Each submitter *must* submit 20 benchmarks
- Proofs of unsatisfiability certified by a theorem prover
  - Proofs were converted into LRAT and checked with ACL2
- No longer in 2018:
  - Incremental track due to the lack of solver submissions
  - Agile track due to the absence of benchmark submissions
- New in 2018:
  - Only new benchmarks (400 in the main track)
  - At most 20 instances selected from one source.

# Tracks

# Tracks

Track	Benchmarks	Solvers	Limits	Cluster
Main (sequential)	400 main app + crafted	41	5000 s, 1 core, 24 GB 20 000 s DRAT	StarExec
Parallel	400 main	21	5000 s / 64 GB 24 cores / 48 threads	TACC
Random SAT	(planted) $k$ -SAT	10	5000 s / 24 GB	StarExec
No-limits	400 main	34	5000 s / 24 GB	StarExec

Total number of solvers (solver versions) submitted: 106

# Benchmarks

**Main:** Several new benchmark domains/sets submitted, Bitcoin Mining, Tree Decompositions, Cryptanalysis, Collatz Conjecture, Chromatic Number of the Plane, Graph Coloring, Polynomial Multiplication, GrandTour Puzzles, Floating-Point Verification, Cellular Automata, Scheduling

**Random:** Satisfiable  $k$ -SAT. Three types: medium size close to the phase transition, huge and somewhat below the phase transition, hard planted SAT, submitted  $q$ -planted solutions.

# Results



# Random Track: Top-3

## Random Track: Top-3

- 
- 
3. **glucose-3.0\_PADC\_10** (902011.58)  
by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

## Random Track: Top-3

- 
2. **gluHack** (901550.62)  
by Aolong Zha
3. **glucose-3.0\_PADC\_10** (902011.58)  
by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

## Random Track: Top-3

1. **Sparrow2Riss** (687420.74)  
by Adrian Balint and Norbert Manthey
2. **gluHack** (901550.62)  
by Aolong Zha
3. **glucose-3.0\_PADC\_10** (902011.58)  
by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

## Random Track: Top-3

1. **Sparrow2Riss** (687420.74)  
by Adrian Balint and Norbert Manthey
2. **gluHack** (901550.62)  
by Aolong Zha
3. **glucose-3.0\_PADC\_10** (902011.58)  
by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

Enormous gap between winner (188 solved) and the rest ( $\leq 165$  solved)

## Random Track: Top-3

1. **Sparrow2Riss** (687420.74)  
by Adrian Balint and Norbert Manthey
2. **gluHack** (901550.62)  
by Aolong Zha
3. **glucose-3.0\_PADC\_10** (902011.58)  
by Rodrigue Konan Tchinda and Clémentin Tayou Djamegni

Enormous gap between winner (188 solved) and the rest ( $\leq 165$  solved)

CDCL solvers are outperforming local search... end of random track?

## Parallel Track (SAT only): Top-3

## Parallel Track (SAT only): Top-3

- 
- 
3. **CryptoMiniSat 5.5** (571149.21)  
by Mate Soos



## Parallel Track (SAT only): Top-3

2. **Painless** (547991.46)  
by Ludovic Le Frioux, Hakan Metin, Souheib Baarir, Maximilien Colange, Julien Sopena, and Fabrice Kordon
3. **CryptoMiniSat 5.5** (571149.21)  
by Mate Soos

## Parallel Track (SAT only): Top-3

1. **Plingeling** (523727.54)  
by Armin Biere
2. **Painless** (547991.46)  
by Ludovic Le Frioux, Hakan Metin, Souheib Baarir, Maximilien Colange, Julien Sopena, and Fabrice Kordon
3. **CryptoMiniSat 5.5** (571149.21)  
by Mate Soos

## Parallel Track (UNSAT only): Top-3

## Parallel Track (UNSAT only): Top-3

- 
- 
3. **abcdSAT** (1110054.28)  
by Jingchao Chen

## Parallel Track (UNSAT only): Top-3

- 
2. **Plingeling** (886430.79)  
by Armin Biere
3. **abcdSAT** (1110054.28)  
by Jingchao Chen

## Parallel Track (UNSAT only): Top-3

1. **Painless** (849533.21)  
by Ludovic Le Frioux, Hakan Metin, Souheib Baarir, Maximilien Colange, Julien Sopena, and Fabrice Kordon
2. **Plingeling** (886430.79)  
by Armin Biere
3. **abcdSAT** (1110054.28)  
by Jingchao Chen

## Parallel Track (UNSAT only): Top-3

1. **Painless** (849533.21)  
by Ludovic Le Frioux, Hakan Metin, Souheib Baarir, Maximilien Colange, Julien Sopena, and Fabrice Kordon
2. **Plingeling** (886430.79)  
by Armin Biere
3. **abcdSAT** (1110054.28)  
by Jingchao Chen

Out of memory on UNSAT instances

Many parallel solvers performed worse on UNSAT (compared to sequential)

# Parallel Track: Top-3



## Parallel Track: Top-3

- 
- 
3. **abcdSAT** (1736011.66)  
by Jingchao Chen

## Parallel Track: Top-3

- 
2. **Plingeling** (1410158.33)  
by Armin Biere
3. **abcdSAT** (1736011.66)  
by Jingchao Chen

## Parallel Track: Top-3

1. **Painless** (1397524.67)  
by Ludovic Le Frioux, Hakan Metin, Souheib Baarir, Maximilien Colange, Julien Sopena, and Fabrice Kordon
2. **Plingeling** (1410158.33)  
by Armin Biere
3. **abcdSAT** (1736011.66)  
by Jingchao Chen

## Parallel Track: Top-3

1. **Painless** (1397524.67)  
by Ludovic Le Frioux, Hakan Metin, Souheib Baarir, Maximilien Colange, Julien Sopena, and Fabrice Kordon
2. **Plingeling** (1410158.33)  
by Armin Biere
3. **abcdSAT** (1736011.66)  
by Jingchao Chen

Performance on UNSAT dominates track

Regardless of PAR-2, which favors performance on SAT

## No-Limits Track: Top-3, but no awards

## No-Limits Track: Top-3, but no awards

- (3.) **CryptoMiniSat 5.5 V20** (1915985.64)  
by Mate Soos

## No-Limits Track: Top-3, but no awards

- (2.) **Maple\_CM** (1890452.09)  
by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li
- (3.) **CryptoMiniSat 5.5 V20** (1915985.64)  
by Mate Soos

## No-Limits Track: Top-3, but no awards

- (1.) **ReasonLS** (1875448.53)  
by Shaowei Cai and Xindi Zhang
- (2.) **Maple\_CM** (1890452.09)  
by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li
- (3.) **CryptoMiniSat 5.5 V20** (1915985.64)  
by Mate Soos



## No-Limits Track: Top-3, but no awards

- (1.) **ReasonLS** (1875448.53)  
by Shaowei Cai and Xindi Zhang
- (2.) **Maple\_CM** (1890452.09)  
by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li
- (3.) **CryptoMiniSat 5.5 V20** (1915985.64)  
by Mate Soos

No “real” no-limits solvers? (spoiler: main track solvers are faster)

# Main Track (SAT only): Top-3

# Main Track (SAT only): Top-3

- 
- 
3. **CryptoMiniSat 5.5** (791841.46)  
by Mate Soos

## Main Track (SAT only): Top-3

- 
2. **Maple\_LCM\_Scavel** (754380.86)  
by Yang Xu, Guanfeng Wu, Qingshan Chen, and Shuwei Chen
3. **CryptoMiniSat 5.5** (791841.46)  
by Mate Soos

## Main Track (SAT only): Top-3

1. **Maple\_LCM\_Dist\_ChronoBT** (750794.78)  
by Vadim Ryvchin and Alexander Nadel
2. **Maple\_LCM\_Scavel** (754380.86)  
by Yang Xu, Guanfeng Wu, Qingshan Chen, and Shuwei Chen
3. **CryptoMiniSat 5.5** (791841.46)  
by Mate Soos

# Main Track (UNSAT only): Top-3

## Main Track (UNSAT only): Top-3

### 3. **Maple\_CM** (1080196.80)

by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li

## Main Track (UNSAT only): Top-3

2. **Maple\_LCM\_M1** (1076075.19)  
by Zhen Li and Kun He
3. **Maple\_CM** (1080196.80)  
by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li



## Main Track (UNSAT only): Top-3

1. **CaDiCaL** (1035209.89)  
by Armin Biere
2. **Maple\_LCM\_M1** (1076075.19)  
by Zhen Li and Kun He
3. **Maple\_CM** (1080196.80)  
by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li

# Main Track: Top-3

### 3. **Maple\_CM** (1908304.62)

by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li

## Main Track: Top-3

- 
2. **Maple\_LCM\_Scavel** (1872489.47)  
by Yang Xu, Guanfeng Wu, Qingshan Chen, and Shuwei Chen
3. **Maple\_CM** (1908304.62)  
by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li

## Main Track: Top-3

1. **Maple\_LCM\_Dist\_ChronoBT** (1857321.82)  
by Vadim Ryvchin and Alexander Nadel
2. **Maple\_LCM\_Scavel** (1872489.47)  
by Yang Xu, Guanfeng Wu, Qingshan Chen, and Shuwei Chen
3. **Maple\_CM** (1908304.62)  
by Mao Luo, Fan Xiao, Chu-Min Li, Felip Manyà, Zhipeng Lü, Yu Li

# Glucose Hack Track: Top-3

# Glucose Hack Track: Top-3

1. **GHackCOMSPS** (2205133.64)  
by Chanseok Oh

3. **glu\_mix** (2262847.55)  
by Jingchao Chen

# Glucose Hack Track: Top-3

1. **GHackCOMSPS** (2205133.64)  
by Chanseok Oh
2. **inIDGlucose** (2246060.45)  
by Jo Devriendt
3. **glu\_mix** (2262847.55)  
by Jingchao Chen



# Virtual Best Solver

Virtual Best Solver would solve 302 instances

Best time-slicing the 5000 seconds:

k=1: 231 with Maple\_LCM\_Dist\_ChronoBT

k=2: 242

k=3: 252

k=4: 256

k=5: 258

k=6: 260 with Sparrow2Riss, Maple\_LCM\_Dist\_ChronoBT, CryptoMiniSat, YaISAT, CaDiCaL, smallsat

k=7: 260

k=8: 258

# Impact of PAR-2

Penalized average runtime (PAR)

- PAR- $x$ : penalized timeouts by  $x \cdot \text{TIMEOUT}$
- SCR, solution-count ranking: PAR- $x$  as  $x \rightarrow \infty$ .
- $x$  balances average succesful runtimes and number of solved instances

In 2018: some differences between PAR-2 and SCR.

# Impact of PAR-2

Penalized average runtime (PAR)

- PAR- $x$ : penalized timeouts by  $x \cdot \text{TIMEOUT}$
- SCR, solution-count ranking: PAR- $x$  as  $x \rightarrow \infty$ .
- $x$  balances average succesful runtimes and number of solved instances

In 2018: some differences between PAR-2 and SCR.

Determined winner in the SAT parallel track:

## PAR-2

1. **Plingeling** (523727.54)  
by Armin Biere
2. **Painless** (547991.46)  
by Ludovic Le Frioux et al.
3. **CryptoMiniSat** (571149.21)  
by Mate Soos

## SCR

- (1). **CryptoMiniSat** (171)  
by Mate Soos
- (2). **Plingeling** (157)  
by Armin Biere
- (3). **Painless** (153)  
by Ludovic Le Frioux et al.

## What is next?

Get your awards at the FLoC Olympic Games ceremony

- Saturday July 14 at 2pm
- Room L3 in the Mathematical Institute

## What is next?

Get your awards at the FLoC Olympic Games ceremony

- Saturday July 14 at 2pm
- Room L3 in the Mathematical Institute

Participants and organizers of SAT-related competitions and evaluations are invited to submit to a special issue of JSAT.

- Editors: Luca Pulina and Martina Seidl
- Deadline: End of October 2018

## Final Remarks

Full details (to be available) at  
<http://sat2018.forsyte.tuwien.ac.at/>

- Detailed per-instance per-solver results
- Proceedings at <http://hdl.handle.net/10138/237063>
- These slides

## Final Remarks

Full details (to be available) at  
<http://sat2018.forsyte.tuwien.ac.at/>

- Detailed per-instance per-solver results
- Proceedings at <http://hdl.handle.net/10138/237063>
- These slides

### **Many thanks to**

- all solver submitters and developers
- all benchmark submitters
- Aaron Stump and StarExec
- TACC for the Lonestar5 resources
- SAT Association for support for awards

Thank you for your attention!